PARTIAL DIFFERENTIAL EQUATIONS OF APPLIED MATHEMATICS

PURE AND APPLIED MATHEMATICS

A Wiley-Interscience Series of Texts, Monographs, and Tracts

A complete list of the titles in this series appears at the end of this volume.

PARTIAL DIFFERENTIAL EQUATIONS OF APPLIED MATHEMATICS

Third Edition

ERICH ZAUDERER
Emeritus Professor of Mathematics
Polytechnic University
New York

A JOHN WILEY & SONS, INC., PUBLICATION

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.
Published simultaneously in Canada.

Library of Congress Cataloging-in-Publication Data:

ISBN-13: 978-0-471-69073-3
ISBN-10: 0-471-69073-2

Printed in the United States of America.

10 9 8 7 6 5 4 3 2

To my wife, Naomi,
my children,
and my grandchildren

CONTENTS

PREFACE

The study of partial differential equations (PDEs) of applied mathematics involves the formulation of problems that lead to partial differential equations, the classification and characterization of equations and problems of different types, and the examination of exact, approximate, and numerical methods for the solution of these problems. Each of these aspects is considered in this book.

The widespread availability of computers in the scientific community and the advent of mathematical software such as *Maple, Matlab,* and *Mathematica* has had the effect of eliminating the need for carrying out many routine symbolic and numerical calculations that arise when solving PDEs manually. Furthermore, it has become possible to create and employ fairly sophisticated numerical methods for the solution of PDEs without having to use lengthy computer codes created by professional numerical analysts. For example, *Maple* has built-in procedures or codes that can solve both ordinary differential equations (ODEs) and PDEs symbolically and numerically. The procedures available for the solution of initial and boundary value problems for ODEs greatly exceed those that are available for the solution of initial and boundary value problems for PDEs. For that reason we have created a number of *Maple* procedures that deal with problems arising in the solution of PDEs and are related to the material in each of the chapters in the book. These procedures generate solutions to problems using the methods developed in each chapter. A graphical representation of the results can often be generated. This has been done for the first ten chapters of

the book, whose material generally follows the presentation of the second edition of the text.

Two new chapters dealing with *finite difference* and *finite element methods* have been added for the third edition. For these two chapters a large number of *Maple procedures* have been created for the solution of various initial and boundary value problems for PDEs. Thus, not only are the ideas behind the numerical solution methods presented, but their implementation is made possible. It was not possible to do everything in triplicate, using *Maple*, *Matlab*, and *Mathematica*, because many new codes were created. As a result, it was decided to restrict our presentation to the use of *Maple*.

The first chapter is concerned with the formulation of problems that give rise to first- and second-order PDEs representative of the three basic types (parabolic, hyperbolic, and elliptic) considered in this book. These equations are all obtained as limits of difference equations that serve as models for discrete *random walk problems*. These problems are of interest in the theory of Brownian motion and this relationship is examined. A new section has been added that presents random walks that yield first order PDEs in the limit. Finally, a section that employs *Maple* procedures to simulate the various random walks and thereby generate approximate solutions of the related PDEs is included. These methods fall under the general heading of *Monte Carlo methods*. They represent an alternative to the direct numerical solution of the difference equations in the manner considered in Chapter 11. Only elementary concepts from probability theory are used in this chapter.

Chapter 2 deals with first order PDEs and presents the *method of characteristics* for the solution of initial value problems for these equations. Problems that arise or can be interpreted in a wave propagation context are emphasized. First order equations also play an important role in the methods presented in Chapters 9 and 10.

In Chapter 3, PDEs are classified into different types and simplified *canonical forms* are obtained for second order linear equations and certain first order systems in two independent variables. The concept of *characteristics* is introduced for higher-order equations and systems of equations, and its significance for equations of different types is examined. In addition, the question of what types of *auxiliary conditions* are to be placed on solutions of PDEs so that the resulting problems are reasonably formulated is considered. Further, some physical concepts, such as *energy conservation* and *dispersion*, which serve to distinguish equations of different types are discussed. Finally, the concept of *adjoint differential operators* is presented.

Chapter 4 presents the method of *separation of variables* for the solution of problems given in bounded spatial regions. This leads to a discussion of *eigenvalue problems* for PDEs and the one-dimensional version thereof, known as the *Sturm-Liouville problem*. *Eigenfunction expansions*, in general, and *Fourier series*, in particular, are considered and applied to the solution of homogeneous and inhomogeneous problems for linear PDEs of second order. It is also shown that eigenfunction expansions can be used for the solution of *nonlinear problems* by considering a nonlinear heat conduction problem.

In Chapter 5, the *Fourier*, *Fourier sine*, *Fourier cosine*, *Hankel*, and *Laplace transforms* are introduced and used to solve various problems for PDEs given over un-

bounded regions in space or time. As the solutions of these problems are generally obtained in an integral form that is not easy to evaluate, *approximation methods* for the evaluation of Fourier and Laplace integrals are presented.

Not all problems encountered in applied mathematics lead to equations with smooth coefficients or have solutions that have as many derivatives as required by the order of the PDEs. Consequently, Chapter 6 discusses methods whereby the concept of solution is weakened by replacing the PDEs by *integral relations* that reduce the number of derivatives required of solutions. Also, methods are presented for dealing with problems given over *composite media* that can result in singular coefficients. Finally, the method of *energy integrals* is discussed and shown to yield information regarding the uniqueness and dependence on the data of solutions of PDEs.

Green's functions, which are discussed in Chapter 7, depend on the theory of generalized functions for their definition and construction. Therefore, a brief but self-contained discussion of *generalized functions* is presented in this chapter. Various methods for determining Green's functions are considered and it is shown how initial and boundary value problems for PDEs can be solved in terms of these functions.

Chapter 8 contains a number of topics. It begins with a *variational characterization* of the eigenvalue problems considered in Chapter 4, and this is used to verify and prove some of the properties of eigenvalues and eigenfunctions stated in Chapter 4. Furthermore, the *Rayleigh-Ritz method,* which is based on the variational approach, is presented. It yields an approximate determination of eigenvalues and eigenfunctions in cases where exact results are unavailable. The classical *Riemann method* for solving initial value problems for second order hyperbolic equations is discussed briefly, as are *maximum* and *minimum principles* for equations of elliptic and parabolic types. Finally, a number of basic *PDEs of mathematical physics* are studied, among which the equations of fluid dynamics and Maxwell's equations of electromagnetic theory are discussed at length.

Chapters 9 and 10 deal with *perturbation* and *asymptotic methods* for solving both linear and nonlinear PDEs. In recent years these methods have become an important tool for the applied mathematician in simplifying and solving complicated problems for linear and nonlinear equations. *Regular* and *singular perturbation methods* and *boundary layer theory* are discussed in Chapter 9. Linear and nonlinear *wave propagation problems* associated with the reduced wave equation that contains a large parameter are examined in Chapter 10. These include the scattering and diffraction of waves from various obstacles and the problem of beam propagation in *linear* and *nonlinear optics*. It is also shown in Chapter 10 how singularities that can arise for solutions of hyperbolic equations can be analyzed without having to solve the full problem given for these equations. Finally, an *asymptotic simplification* procedure is presented that permits the replacement of linear and nonlinear equations and systems by simpler equations that retain certain essential features of the solutions of the original equations.

Chapter 11 presents a full discussion of *finite difference methods* for the numerical solution of initial and initial and boundary value problems for PDEs. Equations of all three types, as well as systems of PDEs, are considered. Linear and nonlinear problems are examined. A large number of difference schemes are introduced, and

questions of consistency and stability are examined. Specially created and built-in *Maple procedures* are presented for implementation of most of these difference schemes.

The *finite element method* for the approximate numerical solution of initial and initial and boundary value problems for a large class of PDEs in two spatial dimensions is presented in Chapter 12. It is developed from the *Galerkin integral representations* of the given problems and the *Galerkin method* for constructing approximate solutions of these problems. *Triangulations* of the spatial region over which the problem is formulated are created and *finite element solutions* are constructed. *Maple procedures* that carry out these processes are presented and their use is demonstrated.

The text includes a substantial number of figures. As we have indicated, not only do built-in and the newly constructed *Maple* procedures solve problems analytically and numerically, but they can also represent results graphically. The figures in Chapters 1, 11, and 12, were generated with the use of *Maple*, which was not the case for the remaining figures. This accounts for the difference in the representation of coordinate axes in some of the figures, for example.

The Bibliography contains a list of references as well as additional reading. The entries are arranged according to the chapters of the book and they provide a collection of texts and papers that discuss some or all of the material covered in each chapter, possibly at a more elementary or advanced level than that of the text.

This book is intended for advanced undergraduate and beginning graduate students in applied mathematics, the sciences, and engineering. The student is assumed to have completed a standard calculus sequence including elementary ODEs, and to be familiar with some elementary concepts from advanced calculus, vector analysis, and matrix theory. (For instance, the concept of uniform convergence, the divergence theorem, and the determination of eigenvalues and eigenvectors of a matrix are assumed to be familiar to the student.) Although a number of equations and problems considered are physically motivated, a knowledge of the physics involved is not essential for the understanding of the mathematical aspects of the solution of these problems.

In writing this book I have not assumed that the student has been previously exposed to the theory of PDEs at some elementary level and that this book represents the next step. Thus I have included such standard solution techniques as the separation of variables and eigenfunction expansions together with the more advanced methods described earlier. However, in contrast to the more elementary presentations of this subject, this book does not dwell at great length on the method of separation of variables, the theory of Fourier series or integrals, the Laplace transform, or the theory of Bessel or Legendre functions. Rather, the standard results and methods are presented briefly but from a more general and advanced point of view. Thus, even with the addition of the numerical finite difference and finite element methods, it has been possible to present a variety of approaches and methods for solving problems for linear and nonlinear equations and systems without having the length of the book become excessive.

There is more than enough material in the book to be covered in a year-long course. For a shorter course it is possible to use the first part of Chapter 3 and Chapters 4

and 5 as a core, and to select additional material from the other chapters, such as numerical methods, as time permits. The book contains many examples. Very often, new approaches or methods are brought out in the form of an example. Thus the examples should be accorded the same attention as the remainder of the text.

In preparing the third edition, the material contained in the second edition was retained, but rewritten, clarified, and revised where necessary, with corrections made as needed. In addition to the inclusion of two new chapters, some new material was added throughout the first ten chapters. In particular, *Maple* methods that deal with the material in each chapter are presented in a new section at the end of each chapter. Additionally, for example, Chapter 1 has a new discussion of random walks related to first order PDEs. To assist the reader, the sections of the book have been broken up into a collection of subsections that focus on specific topics and subtopics that are considered.

A number of new exercises have been created to supplement those of the second edition. The exercises are placed at the end of each section. With a few exceptions, no substantially new theories or concepts are introduced in the exercises. For the most part, the exercises are based on material developed in the text, and the student should attempt to solve as many of them as possible to test his or her mastery of the subject. Answers and solutions to selected exercises and all the Maple codes that were created for use in the book are available via the FTP site:

ftp://ftp.wiley.com/sci_tech_med/partial_differential/

A supplementary Instructor's Solutions Manual is also available.

I would like to thank Susanne Steitz and Steve Quigley, mathematics editors at Wiley-Interscience, for their support of this project. I acknowledge my gratitude to my wife, Naomi, for her assistance and understanding during the many hours that were spent in writing this book.

ERICH ZAUDERER

New Jersey
March 2006

CHAPTER 1

RANDOM WALKS AND PARTIAL DIFFERENTIAL EQUATIONS

It is traditional to begin a course on partial differential equations of applied mathematics with derivations of the basic types of equations to be studied based on physical principles. Conventionally, the problem of the vibrating string or the process of heat conduction is considered and the corresponding wave or heat equation is derived. (Such derivations and relationships are discussed later in the text.) We have chosen instead to use some elementary *random walk problems* as a means for deriving and introducing prototypes of the equations studied in the text: (1) the *diffusion equation*, (2) the *wave equation*, or more specifically, the *telegrapher's equation*, and (3) *Laplace's equation*. More general forms of each of these types of second order partial differential equations are also derived, and random walks that generate first order linear and nonlinear partial differential equations are also presented. The equations describing the random walks are *partial difference equations*, whose continuum limits are the aforementioned partial differential equations. This reverses the approach used in the numerical method of finite differences for the solution of differential equations, where derivatives in the given equation are replaced by difference quotients. Only elementary and basic concepts from probability theory are required for our discussion of the random walk problems and they are not used in the sequel.

Partial Differential Equations of Appplied Mathematics, Third Edition. By Erich Zauderer

Our discussion of the limiting processes will be somewhat heuristic and formal, but they can be rigorously justified. The discrete formulation of differential equations based on random walk problems yields insights into the usefulness and validity of a number of numerical approaches employed in the method of finite differences for the numerical solution of the partial differential equations considered. This relates specifically to questions of consistency, stability and convergence that arise in the analysis of finite difference methods and are addressed in Chapter 11, which deals with these methods. Also, some basic properties of each of the types of equations derived are elementary consequences of the discrete random walk formulations and are verified directly for each of the limiting partial differential equations.

1.1 THE DIFFUSION EQUATION AND BROWNIAN MOTION

Unrestricted Random Walks and Their Limits

We begin with the *unrestricted one-dimensional random walk problem.* A particle starts at the origin of the x-axis and executes random steps or jumps each of length δ to the right or to the left. Let x_i be a random variable that assumes the value δ if the particle moves to the right at the ith step and the value $-\delta$ if it moves to the left. Each step is assumed to be independent of the others, so that in the language of probability theory, the x_i are identically distributed, independent random variables. Let p or q be the probability that the particle moves to the right or to the left, respectively. These probabilities are identical for each step, so we have $Prob(x_i = \delta) \equiv P(x_i = \delta) = p$ and $P(x_i = -\delta) = q$. The particle must move either to the right or to the left, so that $p + q = 1$. The position of the particle at the nth step is given by

$$X_n = x_1 + x_2 + \cdots + x_n. \tag{1.1.1}$$

Using the *binomial distribution*, the probability that the particle is located at a fixed point after a given number of steps can be determined explicitly. Since we are more interested in the continuum limit of the random walk problem as the step length $\delta \to 0$ and the number of steps $n \to \infty$, we do not consider the exact solution of the problem above (see Exercise 1.1.11). Instead, we determine the *mathematical expectation* and the *variance* of the random variable X_n. They will be related below to certain physically significant constants when we connect the random walk model with the theory of the *Brownian motion* of a particle. The expectation of X_n yields the expected or mean location of the particle at the nth step, and the variance of X_n is a measure of how much the actual location of the particle varies around the expected location.

Example 1.1. Expectation and Variance: The Discrete Case. A discrete valued random variable x that assumes the values a_m with probability p_m, has the expected value or the *mathematical expectation* $E(x)$, given as

$$E(x) = \sum_m a_m p_m. \tag{1.1.2}$$

Thus for the random variable x_i we have

$$E(x_i) = (+\delta)P(x_i = \delta) + (-\delta)P(x_i = -\delta) = (p - q)\delta. \tag{1.1.3}$$

Since $E(x)$ is a linear function of x, (1.1.3) yields

$$E(X_n) = E\left(\sum_{i=1}^{n} x_i\right) = \sum_{i=1}^{n} E(x_i) = (p - q)\delta n. \tag{1.1.4}$$

In the physical literature the mathematical expectation is expressed as $E(x) = \langle x \rangle$ and this notation is used below. Thus (1.1.4) yields

$$\langle X_n \rangle = (p - q)\delta n. \tag{1.1.5}$$

The expected location of the particle after n steps is at $x = 0$, $x > 0$, or $x < 0$ according as $p = q$, $p > q$, or $p < q$.

The *variance* $V(x)$ of a random variable x is defined in terms of the expectation as $V(x) = \langle (x - \langle x \rangle)^2 \rangle = \langle x^2 \rangle - \langle x \rangle^2$, where the last equation follows from the properties of the expectation. Now, the variance of a sum of independent random variables equals the sum of the variances of the random variables. Thus, since $X_n = \sum_{i=1}^{n} x_i$ and the x_i are independent random variables, $V(X_n) = V(\sum_{i=1}^{n} x_i) = \sum_{i=1}^{n} V(x_i) = \sum_{i=1}^{n} \left[\langle x_i^2 \rangle - \langle x_i \rangle^2\right]$. But

$$\langle x_i^2 \rangle = (+\delta)^2 P(x_i = \delta) + (-\delta)^2 P(x_i = -\delta) = (p + q)\delta^2 = \delta^2 \tag{1.1.6}$$

since $p + q = 1$. Recalling (1.1.3) and $(p + q)^2 = 1$, we easily obtain

$$V(X_n) = \sum_{i=1}^{n} \left[\delta^2 - (p - q)^2 \delta^2\right] = 4pq\delta^2 n. \tag{1.1.7}$$

Brownian Motion

The foregoing results will now be applied to yield a mathematical description of one-dimensional *Brownian motion*. This motion refers to the ceaseless, irregular, and (apparently) random motion of small particles immersed in a liquid or gas. An individual particle is assumed to undergo one-dimensional motion due to random collisions with smaller particles in the fluid or gas. In a given unit of time, many collisions occur. Each collision is assumed to be independent of the others and to impart a displacement of the particle of length δ to the right or to the left. The apparent randomness of the motion is characterized by assuming that each collision moves the particle independently to the right or to the left with probability p or q, respectively. Clearly, the observed motion of the particle can be simulated by the random walk problem.

Suppose it is found experimentally, for a particle immersed in a fluid or gas undergoing one-dimensional motion, that the average or mean displacement of the particle per unit time equals c, while the variance of the observed displacement around the average equals $D > 0$. We assume that there are r collisions per unit time. (We note that r is not known a priori but is assumed to be large). Then, for our random walk model of the observed motion, we must have (approximately) after r steps

$$(p - q)\delta r \approx c, \qquad 4pq\delta^2 r \approx D, \tag{1.1.8}$$

in view of the expressions (1.1.5) and (1.1.7) for the mean and the variance.

Since the motion of the particle appears to be continuous, we examine the limit as the step length $\delta \to 0$ and the number of steps $r \to \infty$. This must occur in such a way that c and D [as given in (1.1.8)] remain fixed in the limit. Now if $p \neq q$ and $p - q$ does not tend to zero as $\delta \to 0$ and $r \to \infty$, we conclude from (1.1.8) that $\delta r \to c/(p - q)$, $4pq\delta^2 r \to (4cpq/p - q)\,\delta \to 0$. However, $4pq\delta^2 r$ must tend to $D \neq 0$ in the limit as $\delta \to 0$ and $r \to \infty$, in view of (1.1.8), so we must have $p - q \to 0$ in the limit. Combined with $p + q = 1$, this implies that both p and q tend to $\frac{1}{2}$ in the limit. If $p = q = \frac{1}{2}$ in the discrete model, we have $c = 0$. But if $p - q \neq 0$, so that $c \neq 0$, the particle exhibits a *drift* to the right or to the left depending on the sign of c. Additionally, if $D = 0$, there is no variation around the average displacement c and the motion of the particle must appear to be deterministic and not random or irregular.

These results can be realized if we set $p = (1 + b\delta)/2$, $q = (1 - b\delta)/2$ with the constant b (independent of δ) to be determined and chosen such that $0 \leq p, q \leq 1$. Then, $p + q = 1$ and $p - q = b\delta$, and the foregoing limits imply that as $\delta \to 0$ and $r \to \infty$, we have $\delta^2 r \to D$, $p \to \frac{1}{2}$, $q \to \frac{1}{2}$, and $b = c/D$. More generally, if we set

$$p = \frac{1}{2}(a + b\delta), \qquad q = \frac{1}{2}(a - b\delta), \tag{1.1.9}$$

with $0 < a \leq 1$, we find that $p + q = a$. Thus, if $a < 1$, there is a nonzero probability, $1 - p - q = 1 - a$, that the particle rests at each step (i.e., it takes a step of zero length). Since the additional value assumed by the random variable x_i is zero, the expectation and the variance of the random variable are again given by (1.1.3) and (1.1.6), except that $p + q = a$ in (1.1.6). As a result, (1.1.7) becomes $V(X_n) = (a - b^2\delta^2)\delta^2 n$. Finally, (1.1.8) (suitably modified) yields $a\delta^2 r \to D$ and $b\delta^2 r \to c$, so that $b = ac/D$. We see that a and b cannot be determined uniquely in terms of the observed quantities c and D. The more precise details regarding the probability that the particle rests or moves at each step cannot be determined from its observed motion.

For r steps to occur in unit time, each individual step of length δ must occur in $1/r = \tau$ units of time while n steps occur in $n/r = n\tau$ time units. To describe the motion of the particle that starts at the point $x = 0$ at the time $t = 0$, we use the random walk model to obtain the probability that the particle is at the position x at the time t. That is, after n steps we have approximately $X_n = x$ and $n\tau = t$. If $x > 0$,

for example, we must have

$$X_n = k\delta = x \qquad \text{and} \qquad n\tau = t, \tag{1.1.10}$$

where k equals the excess of the number of steps taken to the right over those taken to the left. A similar expression with $k < 0$ is valid if $x < 0$.

We define

$$v(x,t) = P(X_n = x) \qquad \text{at the time} \qquad t = n\tau \tag{1.1.11}$$

to be the probability that at the (approximate) time t, the particle is located (approximately) at the point x. If $p + q = 1$, we can determine an explicit expression for $v(x,t)$ on using the binomial distribution. However, since we are mainly interested in a *continuum limit* of the random walk problem as δ and τ tend to zero, we construct a partial difference equation satisfied by $v(x,t)$. Again, we do not solve the difference equation but show that in the continuum limit it tends to a partial differential equation that serves as a model for the Brownian motion of a particle. (See, however, Exercise 1.1.11.)

The probability distribution $v(x,t)$ satisfies the difference equation

$$v(x, t+\tau) = pv(x-\delta, t) + qv(x+\delta, t). \tag{1.1.12}$$

This states that *the probability that the particle is at x at the time $t + \tau$ equals the probability that it was at the point $x - \delta$ at the time t multiplied by the probability p that it moved to the right in the following step* plus *the probability that the particle was at the point $x + \delta$ at the time t multiplied by the probability q that it moved to the left in the following step.* The plausibility of (1.1.12), where $p + q = 1$, is apparent.

Expanding in a Taylor series with remainder, we have

$$\begin{cases} v(x, t+\tau) = v(x,t) + \tau v_t(x,t) + O(\tau^2), \\ v(x \pm \delta, t) = v(x,t) \pm \delta v_x(x,t) + \frac{1}{2}\delta^2 v_{xx}(x,t) + O(\delta^3). \end{cases} \tag{1.1.13}$$

where $O(y^k)$ means that $\lim_{y\to 0} y^{-k}O(y^k)$ is finite. Substituting (1.1.13) into (1.1.12) we readily obtain

$$v_t(x,t) = \left[(q-p)\frac{\delta}{\tau}\right] v_x(x,t) + \frac{1}{2}(p+q)\frac{\delta^2}{\tau} v_{xx}(x,t) + O\left(\tau + \frac{\delta^4}{\tau}\right). \tag{1.1.14}$$

In the limit as $\delta \to 0$ and $\tau \to 0$, assuming that $p + q = 1$, (1.1.14) tends to the partial differential equation

$$\frac{\partial v(x,t)}{\partial t} = -c\frac{\partial v(x,t)}{\partial x} + \frac{1}{2}D\frac{\partial^2 v(x,t)}{\partial x^2} \tag{1.1.15}$$

on using (1.1.11) with $r = 1/\tau$ and noting that p and q tend to $\frac{1}{2}$ in the limit. (Since δ^2/τ has a finite limit, δ^4/τ must tend to zero.) In view of (1.1.10), which

requires x and t to remain fixed in the limit, we must also have $|k| \to \infty$ and $n \to \infty$ such that $k\delta \to x$ and $n\tau \to t$. Further, $v(x,t)$ must now be interpreted as a *probability density* associated with the continuous random variable x at the time t, rather than as a *probability distribution* in the discrete random walk model. This means that the probability that x lies in a given interval $[a,b]$ is given by $P(a \le x \le b) = \int_a^b v(x,t)\,dx$ at the time t. If the interval $[a,b]$ is small, we have, $v(x,t) \approx P(a \le x \le b)/(b-a)$.

If p and q are defined as in (1.1.9), the difference equation (1.1.12) must be replaced by

$$v(x,t+\tau) = (1-p-q)v(x,t) + pv(x-\delta,t) + qv(x+\delta,t) \qquad (1.1.16)$$

since there is a nonzero probability that the particle was at the point x at the time t and did not move from there. However, if we expand v as in (1.1.13) and note the results following (1.1.9), we again obtain (1.1.15) in the limit, if we set $a\delta^2/\tau = D$ and $b\delta^2/\tau = c$. We note that δ^2/τ must have a finite limit as $\delta \to 0$ and $\tau \to 0$. Thus, the (truncation) error in (1.1.14) is $O(\tau + \delta^2)$.

The equation (1.1.15) is known as a *diffusion equation* and D is called the *diffusion coefficient*. We examine the significance of the coefficients c and D in (1.1.15) in Example 1.3, with calculations based directly on (1.1.15) rather than on the results obtained from the random walk model.

To determine a unique solution $v(x,t)$ of (1.1.15), we need to specify an *initial condition* for the density function. Since the particle was initially (i.e., at the time $t = 0$) assumed to be located at the point $x = 0$, it has unit probability of being at $x = 0$ at the time $t = 0$ and zero probability of being elsewhere at that time. This serves as initial data for the difference equation (1.1.12). In terms of the density function $v(x,t)$ that satisfies (1.1.15), we still must have $v(x,0) = 0$ for $x \neq 0$, but with the density concentrated at $x = 0$. Since the total probability at $t = 0$ satisfies

$$P(-\infty \le x \le \infty) = \int_{-\infty}^{\infty} v(x,0)\,dx = 1, \qquad (1.1.17)$$

we see that $v(x,0)$ behaves like the *Dirac delta function*; that is,

$$v(x,0) = \delta(x). \qquad (1.1.18)$$

[The delta function $\delta(x)$ is not to be confused with the step length δ introduced in the random walk problem.]

Example 1.2. The Dirac Delta Function. The Dirac delta function is discussed fully in Section 7.2. At present, we characterize it as the limit of a sequence of discontinuous functions $\delta_\epsilon(x)$ defined as $\delta_\epsilon(x) = \begin{cases} 1/2\epsilon, & |x| < \epsilon, \\ 0, & |x| > \epsilon. \end{cases}$ Each $\delta_\epsilon(x)$ has unit area under the curve and in the limit as $\epsilon \to 0$, $\delta_\epsilon(x) \to 0$ for all $x \neq 0$. However, on formally interchanging the limit process with integration, we obtain

$\int_{-\infty}^{\infty} \delta(x)\,dx = \lim_{\epsilon\to 0} \int_{-\infty}^{\infty} \delta_\epsilon(x)\,dx = 1$. It also follows for continuous functions $f(x)$, on using the mean value theorem for integrals before passing to the limit, that

$$\int_{-\infty}^{\infty} f(x)\delta(x)\,dx = f(0) \tag{1.1.19}$$

so that in effect $f(x)\delta(x) = f(0)\delta(x)$. Further, $\delta(x-\xi)$ vanishes for all $x \neq \xi$, so that if $\delta(x)$ is replaced by $\delta(x-\xi)$ in (1.1.19), we obtain $f(\xi)$ instead of $f(0)$, so that $f(x)\delta(x-\xi) = f(\xi)\delta(x-\xi)$.

The combined equations (1.1.15) and (1.1.18) constitute an initial value problem for the partial differential equation (1.1.15). It is derived later and it can be shown directly by substitution that the function

$$v(x,t) = \frac{1}{\sqrt{2\pi Dt}} \exp\left[-\frac{(x-ct)^2}{2Dt}\right] \tag{1.1.20}$$

is a solution of the initial value problem (1.1.15) and (1.1.18). For fixed t, the probability density function (1.1.20) is the density function of a *normal* or *Gaussian distribution* with *mean* $\langle x\rangle = ct$ and *variance* $\langle (x-\langle x\rangle)^2\rangle = Dt$. In view of the exponential decay of $v(x,t)$, we find that the density is concentrated around the curve $x - ct = 0$, so that the particle appears to move with the (drift) velocity $dx/dt = c$. This also follows from the equation $\langle x\rangle / t = c$. The variance of the particle location around the path $x = ct$ [in the (x,t)-plane] increases linearly in t and is given by Dt.

Although the foregoing results for the mean and the variance are well known for the normal distribution, they may be verified directly by considering the limit of the discrete random walk problem, and this is done in Example 1.3.

The foregoing random walk model yields (1.1.5) and (1.1.7) as the mean and the variance, respectively. As $n \to \infty$, $\delta \to 0$, and $1/r \to 0$, we obtain from (1.1.5) and (1.1.7)

$$\begin{cases} \langle x\rangle = \lim(p-q)\delta n = \lim(p-q)\delta r n/r = \lim[(p-q)\delta r]n\tau = ct, \\ \langle (x-\langle x\rangle)^2\rangle = \lim 4pq\delta^2 n = \lim 4pq\delta^2 r n\tau = Dt, \end{cases} \tag{1.1.21}$$

on using (1.1.11) and (1.1.10). [Although we have assumed that $p + q = 1$ in (1.1.21), the results remain valid if p and q are defined as in (1.1.9).] Using (1.1.21) and applying the central limit theorem of probability theory to X_n (which is a sum of n independent random variables) as $n \to \infty$, we can conclude directly that the limiting distribution for x is Gaussian with mean ct and variance Dt.

Example 1.3. Expectation and Variance: The Continuous Case. If x is a continuous random variable, the kth moment of x is defined as

$$\langle x^k\rangle = \int_{-\infty}^{\infty} x^k v(x,t)\,dx, \tag{1.1.22}$$

where $v(x,t)$ is the probability density function at the time t. We assume that $v(x,t)$ and its x-derivatives vanish sufficiently rapidly at infinity, so that all terms evaluated at infinity in the following integrations vanish. [In fact, (1.1.20) shows that $v(x,t)$ decays exponentially at infinity.]

First we show $\langle x^0 \rangle = 1$ for all t so that $v(x,t)$ is, in fact, a probability density for all $t > 0$. Integrating in (1.1.15) we have

$$\int_{-\infty}^{\infty} v_t \, dx = \frac{d\langle x^0 \rangle}{dt} = -c \int_{-\infty}^{\infty} v_x \, dx + \frac{1}{2} D \int_{-\infty}^{\infty} v_{xx} \, dx = 0, \tag{1.1.23}$$

since v and v_x are assumed to vanish at infinity. Thus $\langle x^0 \rangle$ is constant in time. But (1.1.17) shows that $\langle x^0 \rangle = 1$ at $t = 0$, so that $\langle x^0 \rangle = 1$ for all time.

Next, we consider the expected value $\langle x \rangle$ of the continuous random variable x. We multiply (1.1.15) by x and integrate from $-\infty$ to ∞. This gives $\int_{-\infty}^{\infty} x v_t \, dx = d/dt \int_{-\infty}^{\infty} xv \, dx = d\langle x \rangle / dt = -c \int_{-\infty}^{\infty} x v_x \, dx + (D/2) \int_{-\infty}^{\infty} x v_{xx} \, dx$ $= c \int_{-\infty}^{\infty} v \, dx = c \langle x^0 \rangle = c$ on integrating by parts, using (1.1.23) and $\langle x^0 \rangle = 1$. At the time $t = 0$, since $v(x,0) = \delta(x)$, the property (1.1.19) of the delta function implies that $\langle x \rangle = 0$ at $t = 0$. Therefore, we obtain an initial value problem for $\langle x \rangle$, $d\langle x \rangle / dt = c$, $\langle x \rangle \big|_{t=0} = 0$ with the solution

$$\langle x \rangle = ct, \tag{1.1.24}$$

which agrees with (1.1.21).

The variance of x is given by $V(x) = \langle x^2 \rangle - \langle x \rangle^2$. The second moment $\langle x^2 \rangle$ clearly vanishes at $t = 0$ in view of (1.1.18) and (1.1.19). Multiplying (1.1.15) by x^2 and integrating, we have $\int_{-\infty}^{\infty} x^2 v_t \, dx = d\langle x^2 \rangle / dt = -c \int_{-\infty}^{\infty} x^2 v_x \, dx + (D/2) \int_{-\infty}^{\infty} x^2 v_{xx} \, dx = 2c \int_{-\infty}^{\infty} xv \, dx + D \int_{-\infty}^{\infty} v \, dx = 2c \langle x \rangle + D \langle x^0 \rangle$ on integrating by parts. Using (1.1.24) and $\langle x^0 \rangle = 1$, we obtain the initial value problem $d\langle x^2 \rangle / dt = 2c^2 t + D$, $\langle x^2 \rangle \big|_{t=0} = 0$, which yields $\langle x^2 \rangle = c^2 t^2 + Dt$. Thus the variance is given by

$$V(x) = \langle x^2 \rangle - \langle x \rangle^2 = c^2 t^2 + Dt - c^2 t^2 = Dt \tag{1.1.25}$$

which agrees with (1.1.21).

Restricted Random Walks and Their Limits

The foregoing random walk problem and the limiting initial value problem provide a valid description of the motion of a Brownian particle in an unbounded region. If the region is bounded on one or both sides, boundary conditions must be added. We now consider two random walk problems in which the particle is restricted to move in the region $x < l$ with $l > 0$. The point $x = l$ is either an absorbing or a reflecting boundary.

At an *absorbing boundary*, once the particle reaches the boundary point at $x = l$, it is absorbed and can no longer move into the region $x < l$. With $v(x,t)$ as the probability that the particle is (approximately) at the point x at the time t, we have at $x = l$,

$$v(l, t+\tau) = pv(l-\delta, t), \tag{1.1.26}$$

since the particle cannot reach $x = l$ from the right. We assume that $p+q = 1$. Using the Taylor series (1.1.13) gives $(1-p)v(l,t) = O(\tau+\delta)$. Since $p \to \frac{1}{2}$ as δ and τ tend to zero, we have in the limit

$$v(l,t) = 0 \tag{1.1.27}$$

as the boundary condition for the density function at an absorbing boundary.

For the case of a *reflecting boundary*, the boundary is assumed to have the properties of an elastic barrier. When the particle reaches the barrier at the time t, there is a probability p that the particle moves to the right beyond $x = l$ and returns to $x = l$ at the time $t+\tau$. (Of course, it can also move to the left with probability q when it reaches the barrier.) Thus the boundary condition for the probability $v(x,t)$ is

$$v(l, t+\tau) = pv(l-\delta, t) + pv(l,t). \tag{1.1.28}$$

Again using (1.1.13) we have $v(l,t) + \tau v_t(l,t) + O(\tau^2) = 2pv(l,t) - p\delta v_x(l,t) + O(\delta^2)$. On multiplying across by δ/τ, noting that $2p - 1 = p - q$, and going to the limit in (1.1.28), we obtain $\lim(\delta/\tau)(p-q)v - \lim(\delta^2/\tau)pv_x - \lim \delta v_t + \lim O(\delta\tau + \delta^3/\tau) = cv - \frac{1}{2}Dv_x = 0$ on using (1.1.11). Therefore, the boundary condition at the reflecting boundary point $x = l$ is

$$cv(l,t) - \frac{1}{2}Dv_x(l,t) = 0. \tag{1.1.29}$$

Within the interval $-\infty < x < l$, $v(x,t)$ satisfies the diffusion equation (1.1.15) with the initial condition (1.1.18) at $t = 0$. At $x = l$, $v(x,t)$ satisfies either the boundary condition (1.1.27) or (1.1.29). The solutions of these initial and boundary value problems are considered in Chapter 5. If we study the motion of a Brownian particle in a finite interval, say $-l < x < l$, each endpoint can represent an absorbing or a reflecting boundary. The boundary conditions at $x = -l$ in the case of an absorbing or reflecting boundary are given as in (1.1.27) and (1.1.29), respectively, except that l is replaced by $-l$. In the interior of the interval, $v(x,t)$ satisfies (1.1.15) with the initial condition (1.1.18). Such problems are considered in Chapter 4.

Fokker-Planck and Kolmogorov Equations

In a further generalization of the foregoing results, we assume that the probabilities p and q [defined as in (1.1.9)] are functions of the position of the particle. Then the probabilities associated with the point x are given as

$$p = \frac{1}{2}(a(x) + b(x)\delta), \qquad q = \frac{1}{2}(a(x) - b(x)\delta). \tag{1.1.30}$$

We assume that $0 < a(x) \leq 1$ and that $b(x)$ is chosen such that $0 \leq p, q \leq 1$. The probability $v(x,t)$ that the particle is at the point x at the time t now satisfies a difference equation,

$$v(x,t+\tau) = [1 - p(x) - q(x)]v(x,t) + p(x-\delta)v(x-\delta,t) + q(x+\delta)v(x+\delta,t). \tag{1.1.31}$$

On expanding all the terms in Taylor series around x and t as in (1.1.13), and assuming that as $\delta \to 0$ and $\tau \to 0$, $a(x)\delta^2/\tau \to D(x)$ and $b(x)\delta^2/\tau \to c(x)$, we obtain, in the limit,

$$\frac{\partial v(x,t)}{\partial t} = -\frac{\partial(c(x)v(x,t))}{\partial x} + \frac{1}{2}\frac{\partial^2(D(x)v(x,t))}{\partial x^2}. \tag{1.1.32}$$

This is known as the *Fokker-Planck equation* in the physical literature and as the *forward Kolmogorov equation* in the mathematical literature. It represents a generalization of the diffusion equation and describes the motion of a Brownian particle in an inhomogeneous medium, on the basis of our model. The initial condition for the probability density function $v(x,t)$ is again (1.1.18). However, the functions $c(x)$ and $D(x)$ can no longer be interpreted as in (1.1.24) and (1.1.25), since the calculations in Example 1.3 are not valid for the Fokker-Planck equation with variable c and D. Nevertheless, the coefficients $c(x)$ and $D(x)$ are known as the *drift* and *diffusion coefficients*, respectively. At an absorbing boundary we find as before that the density function must vanish. The boundary condition at a reflecting boundary point is considered in the exercises.

In the foregoing, the probability that the Brownian particle is at the point x at the time t was given in terms of its location at the preceding time period. This analysis gave rise to the difference, diffusion, and Fokker-Planck equations determined above. Now, however, we examine the probability that the particle is at the point x at the time t in terms of its location in the following time period. With p and q given as in (1.1.30), we find that the particle can only have reached the points $x-\delta$, x, and $x+\delta$ at the time $t+\tau$. Thus, v satisfies the difference equation

$$v(x,t) = [1 - p(x) - q(x)]v(x,t+\tau) + p(x)v(x+\delta,t+\tau) + q(x)v(x-\delta,t+\tau). \tag{1.1.33}$$

In contrast to (1.1.31), all the probabilities p and q are evaluated at x since they are all related to the motion of the particle from the point x at the time t. To determine the probability that the particle is at some point x on the real axis, we must know the probabilities for all points x at some later time. Thus, if $v(0,0) = 1$ and $v(x,0) = 0$ for $x \neq 0$, (1.1.33) gives the probability that the particle is at some point x at an earlier time $t < 0$. Consequently, (1.1.33) is a *backward difference equation*. Proceeding as in the discussion following (1.1.31), we obtain in the limit as $\delta \to 0$ and $\tau \to 0$ in (1.1.33),

$$\frac{\partial v(x,t)}{\partial t} = -c(x)\frac{\partial v(x,t)}{\partial x} - \frac{1}{2}D(x)\frac{\partial^2 v(x,t)}{\partial x^2}. \tag{1.1.34}$$

This is known as the *backward Kolmogorov equation*. If the condition (1.1.18) is assigned for the density function $v(x,t)$ at the time $t = 0$, it represents an end

condition for (1.1.34), as the equation is to be solved for $t < 0$. Although it is assumed that c and D are functions of x in (1.1.34), even if they are taken to be constants, the forward and backward Kolmogorov equation still differ in the sign of the second derivative term. In fact, (1.1.34) is the *adjoint equation* of (1.1.32), as follows from our discussion of adjoint operators in Section 3.6.

The forward and backward Kolmogorov equations determine the probability density at each point x in terms of the position of the particle at an earlier and at a later time, respectively. Even if the time direction is reversed in the backward equation, (i.e., we replace t by $-t$), the distinction as to how the solution is to be interpreted remains. However, we do not pursue these matters further. The form of the boundary conditions for the backward equation is considered in the exercises. Apart from its intrinsic interest, the backward equation also arises in the process of determining the Green's function for the forward equation, as will be seen in Chapter 7.

Properties of Partial Difference Equations and Related PDEs

We conclude this section by noting a number of consequences for the solutions of initial and boundary value problems for the diffusion equation (1.1.15) and its generalizations that follow from the difference equation formulation of the associated random walk problems. To begin, we observe that the difference equations (1.1.12), (1.1.26), (1.1.28), and (1.1.31) imply that the solutions of the initial and boundary value problems evolve in time. That is, the solution of any point at a given time t depends only on initial and boundary data given at earlier times. This characterizes the *causality property*; only past events influence future events. For the backward Kolmogorov equation, however, there is backward causality. That is, events in the future influence the past.

Second, we consider the initial value problem or the initial and boundary value problem with the boundary condition (1.1.27). The difference equation (1.1.12) represents $v(x, t+\tau)$ as a weighted average of v evaluated at two points at an earlier time t. (We have a weighted average since $p > 0$, $q > 0$, and $p + q = 1$.) Since v vanishes initially everywhere except at one point and v vanishes on the boundary, we conclude that $0 \leq v \leq 1$ everywhere and the maximum value of v (i.e., $v = 1$), as well as the minimum value of v (i.e., $v = 0$), is attained either on the boundary or on the initial line. This *maximum and minimum principle* carries over to the diffusion equation, as shown in Chapter 8. The diffusion process distributes the densities in the interior in a fairly uniform fashion so that the maximum and minimum values occur on the initial or boundary line. The difference equation (1.1.32), which yields the Fokker-Planck equation, does not express $v(x, t+\tau)$ as a weighted average of v evaluated at the earlier time t unless p and q are constants. Otherwise, since p and q are all evaluated at different points, the coefficients of v need not add up to 1. Consequently, we do not expect the general Fokker-Planck equation to satisfy a maximum or minimum principle. However, in the difference equation (1.1.33), $v(x, t)$ is given as a weighted average of v evaluated at the time $t + \tau$, since p and q are both evaluated at the same point. As a result there is a maximum principle for the backward Kolmogorov equation, as shown in Chapter 8.

Finally, requiring that δ^2/τ tends to a finite, nonzero limit as δ and τ tend to zero, implies that $\delta/\tau \to \infty$ in the same limit. This means that the speed of the particle in Brownian motion (which is given by the limit of δ/τ) is infinite. This fact is also implied by the solution (1.1.20) of the initial value problem (1.1.15) and (1.1.18). It shows that $v(x,t)$ is instantaneously nonzero for all x when $t > 0$, even though $v(x,t)$ vanishes for all $x \neq 0$ at $t = 0$. Thus there is a nonzero probability, however small it may be, that the particle is located in the neighborhood of any point as soon as t increases from zero.

Langevin Equation

The foregoing limitation (the infinite particle speed) of the theory of Brownian motion based on solution of the diffusion equation was noted by *Einstein*, who was the first to derive a diffusion equation to describe Brownian motion. He recognized that the diffusion equation yields a valid model only as t gets large. Since (with $c = 0$) (1.1.15) also represents the equation of heat conduction, in which v represents the temperature and D represents a thermal diffusion coefficient, the same difficulty with regard to the interpretation of the physical processes involved occurs for the heat equation.

A different analysis of the theory of Brownian motion was given by *Ornstein* and *Uhlenbeck* in an effort to overcome this shortcoming. It is based on the *Langevin equation* for the velocity $u(t) = x'(t)$ of the particle in Brownian motion. It has the form

$$u'(t) = -fu(t) + F(t). \tag{1.1.35}$$

This is just *Newton's law of motion for the particle.* The term $mu(t)$ is the momentum of the particle, and we have divided through by m in the equation. The positive constant f is a frictional resistance coefficient, while the term $F(t)$ represents the effects of random collisions with the other particles in the fluid or gas. They showed that (1.1.35) implies finite particle velocities, in particular for small times, and yields a description of the motion of the particle for large times based on the diffusion equation (1.1.15). Equation (1.1.35) and its generalizations that describe Brownian motion in inhomogeneous media, are all examples of *stochastic differential equations*, in that they contain random terms. The theory and application of such equations has been the subject of much research in recent years.

In the following section we show, on the basis of a modified random walk model, how to construct a limiting differential equation that yields a finite speed for the particle undergoing Brownian motion for small t and reduces to the preceding results as $t \to \infty$.

Exercises 1.1

1.1.1. Verify, by direct substitution, that the function $v(x,t)$ given in (1.1.20) is a solution of the diffusion equation (1.1.15).

1.1.2. Show that the function $v(x,t)$ given in (1.1.20) satisfies the following conditions when $t > 0$, (a) $\int_{-\infty}^{\infty} v(x,t)\,dt = 1$; (b) $\int_{-\infty}^{\infty} xv(x,t)\,dx = ct$; (c) $\int_{-\infty}^{\infty} x^2 v(x,t)\,dx = c^2t^2 + Dt$. Conclude thereby that $v(x,t)$ is a probability density function with mean ct and variance Dt. *Hint*: $\int_{-\infty}^{\infty} e^{-x^2}\,dx = \sqrt{\pi}$.

1.1.3. Again considering the function $v(x,t)$ of (1.1.20), show that the limit as $t \downarrow 0$ is zero when $x \neq 0$. Conclude from this and the result of Exercise 1.1.2(a) that $\lim_{t\downarrow 0} v(x,t) = \delta(x)$, so that the initial condition (1.1.18) is satisfied.

1.1.4. Show that if $l > 0$, the functions

$$v_{\pm}(x,t) = \frac{1}{\sqrt{2\pi Dt}}\left\{\exp\left(-\frac{x^2}{2Dt}\right) \mp \exp\left[-\frac{(x-2l)^2}{2Dt}\right]\right\}$$

satisfy the diffusion equation (1.1.15) (with $c = 0$), the initial condition (1.1.18), and, respectively, satisfy the boundary conditions (1.1.27) and (1.1.29) with $c = 0$. (The construction of these solutions may be based on the method of images that is discussed in Section 7.5.)

1.1.5. Derive the diffusion equation in two dimensions, $v_t = \frac{1}{2}D(v_{xx} + v_{yy})$, by constructing an appropriate difference equation in the manner leading to (1.1.12), but assuming that the particle is equally likely to move to the points $(x \pm \delta, y)$ and $(x, y \pm \delta)$ from the point (x,y). Use the Taylor series and define $D = \lim_{\tau\to 0,\delta\to 0}(\delta^2/2\tau)$ to obtain the diffusion equation from the difference equation.

1.1.6. Derive the Fokker-Planck equation in two dimensions $v_t = -(c_1 v)_x - (c_2 v)_y + \frac{1}{2}\left[(D_1 v)_{xx} + (D_2 v)_{yy}\right]$, by assuming that the probabilities for motion to the right or to the left and up or down are given as in (1.3.19), and constructing an appropriate difference equation. With D given as in Exercise 1.1.5, and the coefficients in the Fokker-Planck equation defined analogously to those in (1.1.32), obtain the differential equation from the difference equation as δ and τ tend to zero.

1.1.7. With the probabilities defined as in Exercise 1.1.6, derive a backward difference equation and the corresponding backward Kolmogorov equation in two dimensions.

1.1.8. Derive the diffusion equation in three dimensions, $v_t = \frac{1}{2}D(v_{xx} + v_{yy} + v_{zz})$, with $D = \lim_{\tau\to 0,\delta\to 0}(\delta^2/3\tau)$, by using assumptions similar to those given in Exercise 1.1.5.

1.1.9. Generalize the results of Exercise 1.1.6 and derive a difference equation and a Fokker-Planck equation for the three-dimensional case.

1.1.10. With the probabilities given as in Exercise 1.1.9, derive a backward difference equation and a backward Kolmogorov equation for the three-dimensional case.

1.1.11. With $x = k\delta$ and $t = n\tau$, where $k = 0, \pm 1, \pm 2, \ldots$ and $n = 0, 1, 2, \ldots$, $v(x,t) = v(k\delta, n\tau)$ gives the probability that the particle is at the point $k\delta$ at the time $n\tau$. (a) Express the difference equation (1.1.12) in terms of the function $v(k\delta, n\tau)$. (b) Given the initial conditions $v(0,0) = 1$ and $v(k\delta, 0) = 0$ for $k \neq 0$, solve the difference equation recursively for all k with $n = 1, 2, 3$, and 4. (c) Show that for each of the foregoing values of n we have $\sum_{k=-\infty}^{\infty} v(k\delta, n\tau) = 1$. (d) Show that the

solution of the initial value problem of part (b) is

$$v(k\delta, n\tau) = \begin{cases} \dfrac{n!}{\left(\frac{n+k}{2}\right)!\left(\frac{n-k}{2}\right)!} p^{(n+k)/2} q^{(n-k)/2}, & \dfrac{n \pm k}{2} = 0, 1, 2, \ldots, \\ 0, & \text{otherwise.} \end{cases}$$

Using this solution, verify the results of part (b).

1.1.12. With $a = 1$, D = constant, and $c(x) = -\omega x$ in (1.1.30) and (1.1.32), the Fokker-Planck equation describes Brownian motion with the particle subjected to an elastic *restoring* force. As $|x|$ increases, the probability that the particle moves farther away from the origin decreases. (a) Show that this *restoring* property follows from the definition of $p(x)$ and $q(x)$. (b) Using the methods of Example 1.3 and retaining the initial condition (1.1.18), determine the first three moments of the random variable x for this case. (c) Obtain the expectation and the variance using the results of part (b) and compare them with the expectation and variance given in the text for the Brownian motion of a *free* particle.

1.1.13. For the random walk problem modeled by (1.1.31), using the assumptions made in the text, show that at a reflecting boundary, (1.1.28) is replaced by $v(l, t+\tau) = p(l-\delta)v(l-\delta, t) + p(l)v(l, t) + [1 - p(l) - q(l)]v(l, t)$. Show that in the limit as δ and τ tend to zero, this yields $\frac{1}{2}(Dv)_x - cv = 0$ as the boundary condition for the Fokker-Planck equation at a reflecting boundary point $x = l$.

1.1.14. (a) Since the particle must remain fixed at an absorbing boundary point once it reaches that point, show that the boundary condition for the backward Kolmogorov equation (1.1.34) at an absorbing boundary point is $v = 0$. (b) Show that at a reflecting boundary point, the boundary condition associated with the backward difference equation (1.1.33) is $v(l, t) = pv(l, t+\tau) + [1 - p - q]v(l, t+\tau) + qv(l-\delta, t+\tau)$, with p and q both evaluated at $x = l$. Verify that as δ and τ tend to zero, this tends to $v_x = 0$. This is the boundary condition for the backward Kolmogorov equation (1.1.34) at a reflecting boundary.

1.1.15. If $x = -l$ and $x = l$ are reflecting boundary points, by integrating the Fokker-Planck equation (1.1.32) from $-l$ to $+l$, show that the integral of $v(x, t)$ over that interval remains constant. Thus, the probability that the particle remains in the interval is fixed for all time since no absorption takes place. This probability is initially equal to 1, and it remains at 1 for all time.

1.1.16. (a) Show that in the two- or three-dimensional case, we must have $v = 0$ at an absorbing boundary, for both the Fokker-Planck equation and the backward Kolmogorov equation. (b) If the particle is equally likely to move in all directions in the two- or three-dimensional case, show that the normal derivative of v at the boundary (i.e., the derivative of v in the direction of the exterior unit normal at the boundary of the region) must vanish at each reflecting boundary point. This boundary condition applies for both the Fokker-Planck and the backward Kolmogorov equations.

1.1.17. Show that the function $v(x,t;y) = \frac{1}{\sqrt{2\pi Dt}} \exp\left[-\frac{(x-y-ct)^2}{2Dt}\right]$ with constant c and D, is a solution of the diffusion equation (1.1.15) and the time-reversed backward Kolmogorov equation $v_t = cv_y + \frac{1}{2}Dv_{yy}$. [This equation results from (1.1.34) if we replace t by $-t$ and x by y.] For both equations the initial condition is $v(x,0;y) = \delta(x-y)$.

1.2 THE TELEGRAPHER'S EQUATION AND DIFFUSION

The random walk problem considered in Section 1.1 was shown to imply an infinite speed for a particle in Brownian motion in the continuum limit modeled by the diffusion equation (1.1.15). It may be reasonably argued that this is a consequence of the basic assumption that each step in the random walk is independent of the previous steps. Thus in the limit as the step length, as well as the time lapse between steps, tends to zero, the probability that the particle moves right or left tends to $\frac{1}{2}$. The particle is, therefore, equally likely to move to the right or to the left regardless of the direction in which it was moving previously. Consequently, the limiting path is totally irregular and the particle cannot be said to have a fixed finite velocity. The same is true for each of the models considered.

Correlated Random Walks and Their Limits

An assumption in the random walk problem that might be expected to yield a smoother path of motion for the particle in the limit (at least in the initial stages) is that there exists a *positive correlation* between two adjacent steps. This correlation is expected to increase to a maximum value of unity as the step length and the time between steps tend to zero. The correlation implies a tendency for the particle to continue moving in a given direction once it begins to move in that direction. If, initially, probabilities for motion to the right or left are established, the particle will maintain its tendency to move in a fixed direction for a certain time at a finite speed. After awhile, the inherent randomness of the process reduces the motion to that obtained in the preceding section. This assumption was introduced into the random walk problem by *R. Fürth* in a study of Brownian motion. He showed that it implies a finite velocity for the particle at small times and yields the results of Section 1.1 for large times.

The *correlated random walk* was considered independently by *G. I. Taylor* in a discussion of diffusion processes. It was reexamined by *S. Goldstein*, who formulated a difference equation characterizing the random walk and constructed its limiting partial differential equation. We shall use the results of Taylor and Goldstein in our discussion but retain the notation of Section 1.1. The correlated random walk is described in the following example.

Example 1.4. The Correlated Random Walk. The particle is assumed to start at $x = 0$ and move to the right or to the left with probability equal to $\frac{1}{2}$ initially. Subsequently, the particle has a constant probability of persistence in or reversal of direction at each step. At the ith step the random variable x_i assumes the values $+\delta$ or $-\delta$ and is readily seen to do so with probability equal to $\frac{1}{2}$. However, the steps are

no longer assumed to be independent, as in Section 1.1. X_n again gives the position of the particle after n steps [as in (1.1.1)], but we now have for the mean value of X_n,

$$\langle X_n \rangle = \left\langle \sum_{i=1}^{n} x_i \right\rangle = \sum_{i=1}^{n} \langle x_i \rangle = 0, \tag{1.2.1}$$

since the equal likelihood of a step to the right or left implies that $\langle x_i \rangle = 0$.

For the variance of X_n, we have $V(X_n) = \langle (X_n - \langle X_n \rangle)^2 \rangle = \langle X_n^2 \rangle$ on using (1.2.1), so that the variance equals the second moment of X_n. Also, we easily obtain $V(x_i) = \langle (x_i - \langle x_i \rangle)^2 \rangle = \langle x_i^2 \rangle = \delta^2$. The second moment $\langle X_n^2 \rangle$ is given as

$$\begin{aligned} \langle X_n^2 \rangle &= \left\langle \left(\sum_{i=1}^{n} x_i \right)^2 \right\rangle = \left\langle \sum_{i=1}^{n} x_i^2 + 2 \sum_{i,j=1, i<j}^{n} x_i x_j \right\rangle \\ &= \sum_{i=1}^{n} \langle x_i^2 \rangle + 2 \sum_{i,j=1, i<j}^{n} \langle x_i x_j \rangle = n\delta^2 + 2 \sum_{i,j=1, i<j}^{n} \langle x_i x_j \rangle, \end{aligned} \tag{1.2.2}$$

and we evaluate it in terms of the correlation coefficient between x_i and x_j.

The *correlation coefficient* $\rho(x_i, x_{i+k})$ between two random variables x_i and x_{i+k} is defined as

$$\rho(x_i, x_{i+k}) = \frac{\langle x_i x_{i+k} \rangle - \langle x_i \rangle \langle x_{i+k} \rangle}{\sqrt{V(x_i) V(x_{i+k})}}. \tag{1.2.3}$$

It vanishes if the x_i are independent random variables. We shall assume that partial correlations between two nonadjacent random variables x_i and x_j (i.e., $| j - i | > 1$) equal zero. Since the x_i are identically distributed random variables, the correlation coefficient between any two adjacent random variables is equal, so that

$$\rho(x_i, x_{i+1}) \equiv \rho. \tag{1.2.4}$$

Further, the correlation between x_i and x_{i+k} (with $k > 1$) occurs only through the intermediate random variables $x_{i+1}, x_{i+2}, \ldots, x_{i+k-1}$ since partial correlations are assumed to vanish. Thus $\rho(x_i, x_{i+k}) = \rho^k$. Since $V(x_i) = \delta^2$ for all i and $\langle x_i \rangle = 0$, we obtain from (1.2.4) $\langle x_i x_{i+k} \rangle = \delta^2 \rho^k$. Introducing this into (1.2.3) gives

$$\langle X_n^2 \rangle = \delta^2 \left[n + 2(n-1)\rho + 2(n-2)\rho^2 + \cdots + 2\rho^{n-1} \right]. \tag{1.2.5}$$

This series is easily summed in terms of the finite geometric series, and we have

$$\langle X_n^2 \rangle = \delta^2 \left[n + \frac{2n\rho}{1-\rho} - \frac{2\rho(1-\rho^n)}{(1-\rho)^2} \right]. \tag{1.2.6}$$

Proceeding as in Section 1.1, we set $n = t/\tau$ in (1.2.6) and consider the limit as $\delta \to 0$, $\tau \to 0$, and $n \to \infty$ with t fixed and $X_n \to x$ to obtain

$$\langle X_n^2 \rangle = \left(\frac{\delta}{\tau} \right)^2 \left[\frac{1+\rho}{1-\rho} t\tau + \frac{2\rho(1-\rho^{t/\tau})\tau^2}{(1-\rho)^2} \right]. \tag{1.2.7}$$

We assume that $\lim \delta/\tau = \gamma$, the finite velocity of the particle. For (1.2.7) to have a finite nonzero limit as $\tau \to 0$, we must have

$$\lim_{\tau \to 0} \frac{\tau}{1-\rho} = \frac{1}{2\lambda}, \tag{1.2.8}$$

where λ is a nonzero positive constant. [The factor 2 in (1.2.8) is introduced for convenience.] This is consistent with the fact that as $\tau \to 0$ the correlation coefficient ρ must tend to unity. It then follows that

$$\lim_{\tau \to 0} \rho^{t/\tau} = e^{-2\lambda t}, \tag{1.2.9}$$

and (1.2.7) tends to

$$\langle x^2 \rangle = \gamma^2 \left[\frac{t}{\lambda} - \frac{1}{2\lambda^2} \left(1 - e^{-2\lambda t}\right) \right]. \tag{1.2.10}$$

An expression of the form (1.2.10) for the variance of x was obtained by Ornstein and Uhlenbeck and Fürth in their (improved) theories of Brownian motion. It was also assumed in their theories (as we have done) that the mean displacement $\langle x \rangle$ of the particle equals zero. For large values of λt, (1.2.10) reduces to

$$\langle x^2 \rangle \approx \frac{\gamma^2}{\lambda} t, \tag{1.2.11}$$

which agrees with the result (1.1.21) for the mean square displacement of the particle in Brownian motion (in the case where $\langle x \rangle = 0$), if we set

$$D = \frac{\gamma^2}{\lambda}, \tag{1.2.12}$$

with D as the diffusion coefficient. For small values of λt, on expanding $e^{-2\lambda t}$ in a Taylor series we obtain $\langle x^2 \rangle \approx \gamma^2 t^2$. This shows that in the initial stages of the motion $\sqrt{\langle x^2 \rangle}/t \approx \gamma$, which means that the motion of the particle is essentially uniform with speed γ. Ornstein and Uhlenbeck, using Langevin's equation, obtained a density function for the random variable x that gives the displacement of the particle in Brownian motion. It reduces to the density function (1.1.20) for the normal distribution as $t \to \infty$, has zero mean, and its variance is given by (1.2.10). They also constructed a differential equation that their density function satisfies but were not able to show that their equation reduces to the diffusion equation (1.1.15) (with $c = 0$) as $t \to \infty$.

Partial Difference Equations for Correlated Random Walks and Their Limits

Following the approach of *S. Goldstein*, we now construct a partial difference equation and its limiting partial differential equation that characterize the correlated random walk described.

Let $\alpha(x,t)$ be the probability that a particle is at the point x at the time t and arrived there from the left, while $\beta(x,t)$ is the probability that a particle is at x at the time t and arrived there from the right. Thus $\alpha(x,t)$ and $\beta(x,t)$ characterize right- and left-moving particles, respectively. Also, let p be the probability that the particle persists in its direction after completing a step, whereas q is the probability that it reverses its direction after completing a step. The probabilities p and q are assumed not to vary from step to step, and we have $p+q=1$. Thus if a particle arrives at x from the left, p is the probability that it continues to the right in the next step, and q is the probability that it reverses its direction and goes to the left in the next step. (Note that p and q were defined differently in the preceding section.)

With steps of length δ occurring in time intervals of length τ, we immediately obtain the coupled system of difference equations for $\alpha(x,t)$ and $\beta(x,t)$,

$$\alpha(x,t+\tau)=p\alpha(x-\delta,t)+q\beta(x-\delta,t), \tag{1.2.13}$$

$$\beta(x,t+\tau)=p\beta(x+\delta,t)+q\alpha(x+\delta,t), \tag{1.2.14}$$

on using the preceding definitions of α,β,p, and q.

The foregoing assumption that as $\tau\to 0$ the correlation coefficient $\rho\to 1$ [see (1.2.8)], implies that as $\tau\to 0$, the probability p of persistence in direction should tend to unity, whereas the probability q of reversal should tend to zero. This means that we should have for small τ,

$$p=1-\hat{\lambda}\tau+O(\tau^2), \qquad q=\hat{\lambda}\tau+O(\tau^2), \tag{1.2.15}$$

where $\hat{\lambda}$ is the *rate of reversal* of direction. Now it can be shown that the correlation coefficient ρ is related to p and q as $\rho=p-q$. Since $p+q=1$, (1.2.15) yields $\rho=p-q=1-2q=1-2\hat{\lambda}\tau+O(\tau^2)$. Recalling (1.2.8), we conclude that $\hat{\lambda}=\lambda$ since

$$\lim_{\tau\to 0}\frac{\tau}{1-\rho}=\frac{1}{2\lambda}=\lim_{\tau\to 0}\frac{\tau}{2\hat{\lambda}\tau+O(\tau^2)}=\frac{1}{2\hat{\lambda}}. \tag{1.2.16}$$

It was observed by *Kac* that in view of (1.2.15) and the independence of the probabilities p and q for each step, the probability of reversal of direction in a given time span is determined by what is known as a *Poisson process*. This will be exploited below when an elementary property of the Poisson process is used.

Introducing Taylor expansions for α and β in (1.2.13)–(1.2.14), and using (1.2.15), yields

$$\alpha_t(x,t)=-(\delta/\tau)\,\alpha_x(x,t)-\lambda\alpha(x,t)+\lambda\beta(x,t)+O(\tau+\delta+\delta^2/\tau), \tag{1.2.17}$$

$$\beta_t(x,t)=(\delta/\tau)\,\beta_x(x,t)+\lambda\alpha(x,t)-\lambda\beta(x,t)+O(\tau+\delta+\delta^2/\tau). \tag{1.2.18}$$

Taking the limit as $\delta,\tau\to 0$, with $\delta/\tau\to\gamma$, yields a coupled system of PDEs:

$$\frac{\partial\alpha(x,t)}{\partial t}+\gamma\frac{\partial\alpha(x,t)}{\partial x}=-\lambda\alpha(x,t)+\lambda\beta(x,t), \tag{1.2.19}$$

$$\frac{\partial \beta(x,t)}{\partial t} - \gamma \frac{\partial \beta(x,t)}{\partial x} = \lambda \alpha(x,t) - \lambda \beta(x,t). \tag{1.2.20}$$

In the limit, $\alpha(x,t)$ and $\beta(x,t)$ are to be interpreted as probability density functions for right- and left-moving particles, respectively.

Since at the time $t = 0$ the particle is located at $x = 0$ and is equally likely to move to the right or to the left, the probabilities α and β vanish for $x \neq 0$ and equal $\frac{1}{2}$ at $x = 0$. In terms of the density functions $\alpha(x,t)$ and $\beta(x,t)$ this yields the initial conditions as follows from the discussion in Section 1.1:

$$\alpha(x,0) = \beta(x,0) = \frac{1}{2}\delta(x), \tag{1.2.21}$$

where $\delta(x)$ is the Dirac delta function.

The system (1.2.19)–(1.2.20) together with (1.2.21) constitutes an initial value problem for $\alpha(x,t)$ and $\beta(x,t)$ that can be solved directly. However, to compare our results with those of Section 1.1, we introduce the probability density function $v(x,t)$ associated with the point x and the time t without regard to the direction in which the particle is moving. It is given as

$$v(x,t) = \alpha(x,t) + \beta(x,t), \tag{1.2.22}$$

since the particle must have arrived at x from either the right or the left. Adding (1.2.20) to (1.2.19) gives

$$\frac{\partial(\alpha+\beta)}{\partial t} + \gamma \frac{\partial(\alpha-\beta)}{\partial x} = 0. \tag{1.2.23}$$

Subtracting (1.2.20) from (1.2.19) gives

$$\frac{\partial(\alpha-\beta)}{\partial t} + \gamma \frac{\partial(\alpha+\beta)}{\partial x} = -2\lambda(\alpha-\beta). \tag{1.2.24}$$

Differentiating (1.2.23) with respect to t and (1.2.24) with respect to x after multiplying by γ, and subtracting the second equation from the first, yields the following equation for $v(x,t) = \alpha(x,t) + \beta(x,t)$:

$$\frac{\partial^2 v(x,t)}{\partial t^2} - \gamma^2 \frac{\partial^2 v(x,t)}{\partial x^2} + 2\lambda \frac{\partial v(x,t)}{\partial t} = 0. \tag{1.2.25}$$

This partial differential equation is a special case of what is called the *telegrapher's equation*, an equation that governs the propagation of signals on telegraph lines. It can also be characterized as a *wave equation* with a damping effect due to the term $2\lambda\, \partial v/\partial t$. That is, if $\lambda = 0$, (1.2.25) reduces to the (one-dimensional) wave equation. The solution of (1.2.25) is presented in later chapters, however, the damping effect is considered below.

In contrast to the diffusion equation (1.1.15), which contains only one time derivative, the telegrapher's equation (1.2.25) has two time derivatives and, therefore, requires two initial conditions. From (1.2.21)–(1.2.22) we obtain, since $\alpha - \beta = 0$ at $t = 0$,

$$v(x,0) = \delta(x), \qquad v_t(x,0) = 0. \tag{1.2.26}$$

In the following example, we derive the moments of the continuous random variable x whose density function $v(x,t)$ satisfies (1.2.25)–(1.2.26).

Example 1.5. Expectation and Variance: The Continuous Case. To obtain the first three moments of the random variable x, (1.2.25) is multiplied by 1, x, and x^2, respectively, and the result is integrated from $-\infty$ and ∞ with respect to x. Integrating by parts, assuming that $v(x,t)$ vanishes sufficiently rapidly at infinity so that all contributions from the limits at infinity vanish, and using the initial conditions (1.2.26), we easily obtain the following equations.

$$\frac{d^2}{dt^2}\langle x^0\rangle + 2\lambda\frac{d}{dt}\langle x^0\rangle = 0, \qquad \langle x^0\rangle\Big|_{t=0} = 1, \qquad \frac{d}{dt}\langle x^0\rangle\Big|_{t=0} = 0. \quad (1.2.27)$$

$$\frac{d^2}{dt^2}\langle x\rangle + 2\lambda\frac{d}{dt}\langle x\rangle = 0, \qquad \langle x\rangle\Big|_{t=0} = 0, \qquad \frac{d}{dt}\langle x\rangle\Big|_{t=0} = 0. \quad (1.2.28)$$

$$\frac{d^2}{dt^2}\langle x^2\rangle + 2\lambda\frac{d}{dt}\langle x^2\rangle = 2\gamma^2, \quad \langle x^2\rangle\Big|_{t=0} = 0, \quad \frac{d}{dt}\langle x^2\rangle\Big|_{t=0} = 0. \quad (1.2.29)$$

The moments $\langle x^k\rangle$ are defined as in (1.1.22).

The solution of (1.2.27) is $\langle x^0\rangle = 1$, which shows that $v(x,t)$ is indeed a probability density function for all $t \geq 0$. The solution of (1.2.28) is clearly $\langle x\rangle = 0$, whereas that of (1.2.29) is identical with (1.2.10). These results confirm that $v(x,t)$ characterizes a random variable whose expectation and variance agree with those obtained above in the continuum limit of the correlated random walk.

Telegrapher's, Diffusion and Wave Equations

From (1.2.11)–(1.2.12) we see that as $t \to \infty$, the mean square displacement $\langle x^2\rangle$ of the particle tends to the form given in Section 1.1 for a particle undergoing Brownian motion. We may expect, therefore, that in some sense the telegrapher's equation (1.2.25) tends to the diffusion equation (1.1.15) (with $c = 0$) as $t \to \infty$. To establish a means of assessing the magnitude of the terms in (1.2.25) for large t, we introduce the change of scale $t = \sigma/\epsilon$, where $0 < \epsilon \ll 1$. Even for moderate values of σ, since ϵ is small, t takes on large values. We have $\partial/\partial t = \epsilon(\partial/\partial\sigma)$ and $\partial^2/\partial t^2 = \epsilon^2(\partial^2/\partial\sigma^2)$, so that (1.2.25) becomes

$$\epsilon^2\frac{\partial^2 v(x,\sigma)}{\partial\sigma^2} - \gamma^2\frac{\partial^2 v(x,\sigma)}{\partial x^2} + 2\lambda\epsilon\frac{\partial v(x,\sigma)}{\partial\sigma} = 0. \quad (1.2.30)$$

If we assume that v does not vary rapidly with respect to σ, we may discard the term $\epsilon^2(\partial^2 v/\partial\sigma^2)$ compared to $2\lambda\epsilon(\partial v/\partial\sigma)$ since $\epsilon \ll 1$ implies that $\epsilon^2 \ll \epsilon$. Thus (1.2.25) can be approximated for large t by the equation

$$\frac{\partial v(x,t)}{\partial t} = \frac{\gamma^2}{2\lambda}\frac{\partial^2 v(x,t)}{\partial x^2}, \quad (1.2.31)$$

where σ was replaced by t in (1.2.30) after discarding the second time derivative term. This equation is to be compared with the diffusion equation (1.1.15) where we must put $c = 0$. It has already been shown in (1.2.11)–(1.2.12) that with $D = \gamma^2/\lambda$, the mean square displacements obtained from the random walk models in this and the preceding section agree for large t. Consequently, if the parameters are identified in this manner, (1.2.32) is identical with (1.2.15). The validity of the diffusion equation is limited to large values of t in its use as a model for Brownian motion as we have seen. However, (1.2.25) may be taken to represent a valid model for Brownian motion for all $t \geq 0$. For small values of t, it yields the required finite particle velocities, and for large t, it reduces to the diffusion equation.

To examine the behavior of the solution of the initial value problem (1.2.25)–(1.2.26) for small t, we note that the speed of the particle is γ. Thus, since the particle is at $x = 0$ at the time $t = 0$ and it can travel to the right or the left, it can never reach the set of points x for which $|x| > \gamma t$. The density function $v(x, t)$ should, therefore, vanish for $|x| > \gamma t$, and this will be verified when the full solution of (1.2.25)–(1.2.26) is given in a later chapter.

The location of the particle on one of the lines $x = \pm\gamma t$, can only be the result of the particle never having reversed its direction from the time $t = 0$ on, when it started on one of the paths $x = \pm\gamma t$ in the (x, t)-plane. The probability of reversal of direction is characterized by a Poisson process as described by (1.2.15) with rate of reversal $\hat{\lambda} = \lambda$. As is well known for this process with rate λ, the probability of nonreversal (i.e., persistence) in direction is given by $e^{-\lambda t}$ at time t. Since the particle is equally likely to be on the line $x = \gamma t$ or $x = -\gamma t$, we conclude that the probability that the particle is on either line is

$$v(x, t)\Big|_{x=\pm\gamma t} = \frac{1}{2}e^{-\lambda t}. \tag{1.2.32}$$

For small values of t, $e^{-\lambda t} \approx 1$ so that most of the probability is concentrated on the lines $x = \pm\gamma t$. Therefore, we have essentially deterministic motion with speed γ along the lines $x = \pm\gamma t$. As t increases, $e^{-\lambda t}$ decays rapidly and the density $v(x, t)$ begins to become more concentrated in the interior region $| x | < \gamma t$. Eventually, the motion appears to become totally random and can be described by the random walk model and the limiting diffusion equation given in the preceding section.

In a further analysis of the effect of the randomness assumptions that led to the partial differential equations (1.2.15) and (1.2.25), it is of interest to consider the reduced equations obtained from (1.1.15) and (1.2.25) when D and λ, respectively, are equated to zero. Since D is a measure of the variance around the mean particle path, we expect that the reduced equation with $D = 0$ should yield deterministic motion with the mean speed c. Also, λ is the rate of direction reversal of the particle, so that if $\lambda = 0$, the particle should move with speed γ either along the path $x = \gamma t$ or $x = -\gamma t$. Putting $D = 0$ in (1.1.15) gives

$$\frac{\partial v(x, t)}{\partial t} + c\,\frac{\partial v(x, t)}{\partial x} = 0. \tag{1.2.33}$$

With the initial condition (1.1.18), the formal solution of (1.2.33) is

$$v(x,t) = \delta(x - ct), \tag{1.2.34}$$

where the Dirac delta function $\delta(x - ct)$ vanishes for $x - ct \neq 0$. If we formally differentiate (1.2.34), we find that it satisfies (1.2.33), and (1.2.34) equals $\delta(x)$ at $t = 0$. The density function $v(x,t)$ is concentrated on the path $x - ct = 0$, so that the particle moves (deterministically) along that path. Equivalently, the particle moves with the fixed velocity c. The case where $D \neq 0$ in (1.1.15) (i.e., the diffusion equation) may, therefore, be characterized as representing a random motion around a deterministic path for the particle given by $x = ct$ in the (x,t)-plane. The equation (1.2.33) is called a *wave equation* representing *unidirectional wave motion* with velocity c, for reasons that will become apparent when such equations are discussed in Chapter 2.

Putting $\lambda = 0$ in (1.2.25) yields the *wave equation*

$$\frac{\partial^2 v(x,t)}{\partial t^2} - \gamma^2 \frac{\partial^2 v(x,t)}{\partial x^2} = 0. \tag{1.2.35}$$

With the initial conditions (1.2.26), the formal solution of (1.2.35) is

$$v(x,t) = \frac{1}{2}\delta(x - ct) + \frac{1}{2}\delta(x + ct). \tag{1.2.36}$$

This result may be obtained from (1.2.33)–(1.2.34) by noting that with $\lambda = 0$, the system (1.2.19)– (1.2.20) reduces to two unidirectional wave equations,

$$\frac{\partial \alpha(x,t)}{\partial t} + \gamma \frac{\partial \alpha(x,t)}{\partial x} = 0, \quad \frac{\partial \beta(x,t)}{\partial t} - \gamma \frac{\partial \beta(x,t)}{\partial x} = 0, \tag{1.2.37}$$

with initial data $\alpha(x,0) = \beta(x,0) = \delta(x)/2$. Recalling that $v = \alpha + \beta$ and the result (1.2.34), immediately yields (1.2.36). The definition of the Dirac delta function implies, by way of (1.2.36), that the particle is restricted to move along the deterministic path $x = \gamma t$ or $x = -\gamma t$ with speed γ. The factor $\frac{1}{2}$ before each delta function is a consequence of the random choice of direction at the time $t = 0$ for the particle (i.e., either to the right or to the left). Once the particle chooses a direction it must continue to move in that direction, since $\lambda = 0$ implies that the probability of reversal of direction is zero. If $\lambda \neq 0$, we have seen in the foregoing that there is some initial directionality to the particle motion, but this disappears rapidly and the motion becomes completely random. The wave equation (1.2.35) permits two directions of motion (to the right and to the left) with speed γ.

Another important limiting form of the telegrapher's equation (1.2.25) occurs if we divide by 2λ and let $\lambda \to \infty$ and $\gamma \to \infty$ but require that γ^2/λ remain fixed and nonzero. Then, (1.2.25) reduces to (1.2.31). As the rate of reversal λ and the particle speed γ both tend to infinity, but with $\gamma^2/\lambda = D$ as in (1.2.12), the Brownian motion model of this section is effectively reduced to that of the preceding section. However, it is not correct to state that both models are completely equivalent under the assumptions of large reversal rates and speeds. For, as we have shown in (1.2.11),

the mean square displacement of the particle predicted by both models agrees only if λt is large. Thus, even for very large λ, if t is sufficiently small, the results of both models disagree. Further, if we replace (1.2.25) by (1.2.31), we lose a second order time derivative term, and (1.2.31) cannot accommodate both initial conditions (1.2.26). The neglect of the term v_{tt} because it is multiplied by a small parameter represents a *singular perturbation* of the equation. Such problems are studied in Section 9.3.

Position-Dependent Correlated Random Walks and Their Limits

As was done for the Brownian motion model considered in Section 1.1, we now introduce persistence and reversal probabilities that depend on the position of the particle. In a further generalization, we assume that these probabilities depend on the direction of motion of the particle and admit the possibility that the particle rests at each step. In the present correlated random walk model, we define the probabilities $\alpha(x,t)$ and $\beta(x,t)$ as before. However, p^+ and q^+ represent the probabilities of persistence and reversal in direction if the particle is moving to the right, and p^- and q^- represent the same probabilities for leftward motion. These probabilities are given as

$$p^{\pm}(x,t) = \sigma(x) - \lambda^{\pm}(x)\tau, \qquad q^{\pm}(x,t) = \lambda^{\pm}(x)\tau. \tag{1.2.38}$$

The probability that the particle rests at each step is given by $1 - p^{\pm} - q^{\pm} = 1 - \sigma$, and is taken to be independent of the direction in which the particle is moving. As $\tau \to 0$, both reversal probabilities tend to zero. This characterizes the correlated nature of the random walk. We again require that as δ and τ tend to zero, the ratio $\delta/\tau \to \gamma$, with γ as the finite particle velocity.

The coupled system of difference equations satisfied by $\alpha(x,t)$ and $\beta(x,t)$ is easily found to be

$$\alpha(x,t+\tau) = [1-p^+(x)-q^+(x)]\,\alpha(x,t)+p^+(x-\delta)\alpha(x-\delta,t)+q^-(x-\delta)\beta(x-\delta,t), \tag{1.2.39}$$

$$\beta(x,t+\tau) = [1-p^-(x)-q^-(x)]\,\beta(x,t)+p^-(x+\delta)\beta(x+\delta,t)+q^+(x+\delta)\alpha(x+\delta,t). \tag{1.2.40}$$

On expanding the terms in a Taylor series around x and t, and letting δ and τ tend to zero, we obtain

$$\alpha_t(x,t) + \gamma(\sigma(x)\alpha(x,t))_x + \lambda^+(x)\alpha(x,t) - \lambda^-(x)\beta(x,t) = 0, \tag{1.2.41}$$

$$\beta_t(x,t) - \gamma(\sigma(x)\beta(x,t))_x - \lambda^+(x)\alpha(x,t) + \lambda^-(x)\beta(x,t) = 0. \tag{1.2.42}$$

To obtain a single equation for the probability density function $v(x,t)$, defined as in (1.2.22), we assume that the sum of the reversal rates λ^+ and λ^- is a constant. More precisely, we set

$$\lambda^{\pm}(x) = \lambda \mp \sqrt{\lambda}\,\psi(x), \tag{1.2.43}$$

where λ is a constant and $\psi(x)$ is a given function. On manipulating (1.2.41)–(1.2.42), as was done for (1.2.19)–(1.2.20), we obtain

$$v_{tt}(x,t) - \gamma^2[\sigma(x)(\sigma(x)v(x,t))_x]_x + 2\sqrt{\lambda}\,\gamma\,(\sigma(x)\psi(x)v(x,t))_x + 2\lambda v_t(x,t) = 0. \tag{1.2.44}$$

This equation is the analog of the telegrapher's equation (1.2.25) for the case of Brownian motion in an inhomogeneous medium. The initial data for $v(x,t)$ at $t = 0$ are again given by (1.2.26).

If $\sigma = 1$ and $\psi = 0$, (1.2.44) reduces to (1.2.25). The difference $\lambda^- - \lambda^+$ in the reversal rates gives rise to the drift term (i.e., the first order x-derivative term) in (1.2.44). This term arises even if $\sigma = 1$, in which case there is a zero rest probability. To relate (1.2.44) to the Fokker-Planck equation (1.1.15), which was shown to characterize Brownian motion in an inhomogeneous medium on the basis of an uncorrelated random walk model, we divide by 2λ in (1.2.44) and let λ and γ tend to infinity while requiring that γ^2/λ have a nonzero limit. This yields

$$v_t(x,t) = -(c(x)v(x,t))_x + \frac{1}{2}\left[\sqrt{D(x)}\left(\sqrt{D(x)}\,v(x,t)\right)_x\right]_x, \tag{1.2.45}$$

where $\sqrt{D(x)} = \lim(\gamma/\sqrt{\lambda})\sigma(x)$ and $c(x) = \lim(\gamma/\sqrt{\lambda})\psi(x)\sigma(x)$. We observe that (1.2.45) does not have the same form as (1.1.15) unless the diffusion coefficient D is a constant.

The difference in the form of the diffusion equations (1.1.15) and (1.2.45) is well known in the theory of *stochastic differential equations*. The Fokker- Planck equation (1.1.15) results on solving *generalized Langevin equations* by using the *Ito stochastic calculus*. If the *Stratonovich stochastic calculus* is used, an equation of the form (1.2.45) results. Further, it has been shown that on solving certain generalized Langevin equations in which the random terms contain correlation effects, and then letting the correlation effects tend to zero, a diffusion equation of the type (1.2.45) results. This conclusion is consistent with our derivation of (1.2.45). It is also possible to construct backward difference equations, telegrapher's equations and their generalizations in the manner of the results in Section 1.1. These matters are considered in the exercises.

Boundary value problems based on the random walk model formulated in this section are considered in the exercises. A full discussion of initial and initial boundary value problems for the telegrapher's and related equations is given in later chapters. We conclude by noting that the difference equation formulation of the random walk problems in this section again indicates, as was the case in the preceding section, that the solution at the time t depends only on data for the problem given at earlier times. The fact that the particle speed γ is finite implies, as was indicated, that the density function $v(x,t)$ with the data (1.2.26) concentrated at the origin vanishes when $|\,x\,| > \gamma t$. This is true for the solutions of the telegrapher's and wave equations. For the equation (1.2.44) with the data (1.2.26), the region outside of which $v(x,t) = 0$ has a more complicated form. It will be determined later in the book. The existence of a finite (maximal) speed of propagation of disturbances is a fundamental property of hyperbolic partial differential equations, of which the wave and telegrapher's

equations are prototypes. This property is not shared by diffusion equations (which are equations of parabolic type), as shown in the preceding section. (A classification of equations into various types is given in Chapter 3.)

Exercises 1.2

1.2.1. Verify that (1.2.5) sums to (1.2.6).

1.2.2. Show that the limit in (1.2.9) is correct.

1.2.3. Obtain the results of Example 1.5 for the moments specified by solving the given initial value problems for the ordinary differential equations.

1.2.4. Consider the solution (1.1.20) of the initial value problem (1.1.18) for the diffusion equation (1.1.15). Using the approach of Exercise 1.1.3, show that the limit of this solution as $D \to 0$ is the delta function (1.2.34).

1.2.5. (a) Let $v = Ve^{-\lambda t}$ in the telegrapher's equation (1.2.25) and show that V satisfies the equation $V_{tt} - \gamma^2 V_{xx} - \lambda^2 V = 0$. (b) Put $z = \sqrt{\gamma^2 t^2 - x^2}$ and show that $V(x,t) = W(z)$ satisfies the ODE $W''(z) + (1/z)W'(z) - (\lambda^2/\gamma^2)W(z) = 0$. [Note that the region $z \geq 0$ is where the solution $v(x,t)$ was shown to be nonzero in the text.] (c) Let $W(z) = (1/\sqrt{z})w(z)$ and show that $w(z)$ satisfies the equation $w''(z) + [1/4z^2 - (\lambda^2/\gamma^2)]w(z) = 0$. (d) Observe that for $|x| < \gamma t$, as $\gamma t \to \infty$ we have $z = \sqrt{\gamma^2 t^2 - x^2} \approx \gamma t - (x^2/2\gamma t)$ and, consequently, $z \to \infty$ as $\gamma t \to \infty$. Thus we can approximate the equation for $w(z)$ when γt is large by $w''(z) - (\lambda^2/\gamma^2)w(z) = 0$. Using (1.2.12), show that as $\gamma t \to \infty$ we have approximately $v(x,t) \approx (a/\sqrt{\gamma t})e^{-(x^2/2Dt)}$, where a is an arbitrary constant. This has the form of the solution (1.1.20) of the diffusion equation with $c = 0$.

1.2.6. Put $x = k\delta$ and $t = n\tau$ as in Exercise 1.1.11 and set $\alpha(x,t) = \alpha(k\delta, n\tau)$ and $\beta(x,t) = \beta(k\delta, n\tau)$. Show that the difference equations (1.2.13)–(1.2.14) take the form $\alpha[k\delta,(n+1)\tau] = p\alpha[(k-1)\delta, n\tau)] + q\beta[(k-1)\delta, n\tau]$, $\beta[k\delta,(n+1)\tau] = p\beta[(k+1)\delta, n\tau)]+q\alpha[(k+1)\delta, n\tau]$, with the initial conditions $\alpha(0,0) = \beta(0,0) = \frac{1}{2}$, and $\alpha(k\delta, 0) = \beta(k\delta, 0) = 0$, for $k \neq 0$.

1.2.7. (a) Solve the difference equations of Exercise 1.2.6 recursively for all k and for $n = 1, 2, 3$, and 4. (b) Show that for the values of n above, $\sum_{k=-\infty}^{\infty} \alpha(k\delta, n\tau) = \sum_{k=-\infty}^{\infty} \beta(k\delta, n\tau) = \frac{1}{2}$.

1.2.8. (a) Solve the difference equations of Exercise 1.2.6 for $\alpha(\pm n\delta, n\tau)$ and $\beta(\pm n\delta, n\tau)$. Show that $\alpha(n\delta, n\tau) = \beta(-n\delta, n\tau) = \frac{1}{2}p^{n-1}$ and that $\alpha(-n\delta, n\tau) = \beta(n\delta, n\tau) = 0$ for $n \geq 1$. (b) With $x = \pm n\delta$ and $t = n\tau$ so that $x = \pm(\delta/\tau)t$, show that as $n \to \infty$, $\delta \to 0$, $\tau \to 0$, and $\delta/\tau \to \gamma$, we have $\alpha(\gamma t, t) \to \frac{1}{2}e^{-\lambda t}$, $\beta(-\gamma t, t) \to \frac{1}{2}e^{-\lambda t}$. These results are in agreement with (1.2.32) since $v = \alpha + \beta$.

1.2.9. With $v(k\delta, n\tau) = \alpha(k\delta, n\tau) + \beta(k\delta, n\tau)$, show that v satisfies the difference equation $v[k\delta,(n+1)\tau] = pv[(k-1)\delta, n\tau]+pv[(k+1)\delta, n\tau]-(p-q)v[k\delta,(n-1)\tau]$ with the initial conditions $v(0,0) = 1$, $v(k\delta, 0) = 0$, $k \neq 0$, and $v(-\delta, \tau) =$

$v(\delta,\tau) = \frac{1}{2}$, $v(k\delta,\tau) = 0$, $k \neq \pm 1$. [We observe that with $p = q = \frac{1}{2}$, the correlation coefficient $\rho = p - q$ vanishes and the difference equation reduces to (1.1.12). If $\rho = p - q \neq 0$, the probability $v(k\delta, n\tau)$ depends not only on the location of the particle at the preceding step but on the preceding two steps.]

1.2.10. (a) Solve the difference equation of Exercise 1.2.9 for $v(k\delta, n\tau)$ with $n = 2, 3$ and 4, and for all k. (b) Solve the equation for $v(\pm n\delta, n\tau)$ for $n \geq 2$. (c) Show that $v(k\delta, n\tau) = 0$ for $|k| > n$.

1.2.11. Let λ^+ and λ^- be constants and $\sigma = 1$ in (1.2.38) and in (1.2.41) and (1.2.42). Putting $v = \alpha + \beta$, show that $v(x,t)$ satisfies the telegrapher's equation $v_{tt} - \gamma^2 v_{xx} + (\lambda^+ + \lambda^-)v_t - \gamma(\lambda^+ - \lambda^-)v_x = 0$.

1.2.12. (a) Apply the method of Example 1.5 to show that as $t \to \infty$ the expectation $\langle x \rangle$ and variance $V(x)$ of the continuous random variable characterized by the equation in Exercise 1.2.11 are given as

$$\langle x \rangle = -\frac{\gamma(\lambda^+ - \lambda^-)}{(\lambda^+ + \lambda^-)}\, t, \qquad V(x) = \langle x^2 \rangle - \langle x \rangle^2 \approx \frac{8\gamma^2 \lambda^+ \lambda^-}{(\lambda^+ + \lambda^-)^3}\, t$$

(b) Compare the results of part (a) with those for the Brownian motion model of Section 1.1. Explain why the coefficient of t in the expression for $\langle x \rangle$ may be characterized as a *drift* velocity. (c) Show that $V(x) \approx \gamma^2 t^2$ for small t.

1.2.13. Generate the results of Exercise 1.2.12 for the expectation and variance by proceeding as follows. Assuming that x and t are large and of the same order of magnitude (say, $x = \hat{x}/\epsilon$ and $t = \hat{t}/\epsilon$, where $0 < \epsilon \ll 1$), we can neglect the second derivative terms in the equation for $v(x,t)$ given in Exercise 1.2.11 and obtain in a first approximation, $v_t - \frac{\gamma(\lambda^+ - \lambda^-)}{(\lambda^+ + \lambda^-)} v_x = 0$. Using this equation to obtain an expression for v_{xx}, show that in the next approximation we obtain the following diffusion equation for $v(x,t)$:

$$v_t - \frac{\gamma(\lambda^+ - \lambda^-)}{(\lambda^+ + \lambda^-)} v_x - \frac{4\gamma^2 \lambda^+ \lambda^-}{(\lambda^+ + \lambda^-)^3} v_{xx} = 0.$$

Demonstrate that this equation implies the results of Exercise 1.2.12(a).

1.2.14. With the persistence and reversal probabilities $p^\pm$ and $q^\pm$ defined as in (1.2.38), obtain the difference equations

$$\alpha(x,t) = [1 - p^+ - q^+]\alpha(x, t+\tau) + p^+ \alpha(x+\delta, t+\tau) + q^+ \beta(x-\delta, t+\tau),$$

$$\beta(x,t) = [1 - p^- - q^-]\beta(x, t+\tau) + q^- \alpha(x+\delta, t+\tau) + p^- \beta(x-\delta, t+\tau),$$

where the $p^\pm$ and $q^\pm$ are all evaluated at the point x. These equations determine the probabilities that right- and left-moving particles are located at the point x at the time t, in terms of their possible locations at the time $t + \tau$.

1.2.15. As δ and τ tend to zero, show that the system of difference equations derived in Exercise 1.2.14 yields the system of differential equations $\alpha_t + \gamma\sigma\alpha_x - \lambda^+\alpha + \lambda^+\beta = 0, \beta_t - \gamma\sigma\beta_x + \lambda^-\alpha - \lambda^-\beta = 0$. This system is the analog of the backward Kolmogorov equation for the model of Brownian motion considered in this section.

1.2.16. Put $v = \alpha + \beta$ and let λ^+ and λ^- be defined as in (1.2.43). Show that the system of Exercise 1.2.15 yields $v_{tt} - \gamma^2\sigma(\sigma v_x)_x - 2\sqrt{\lambda}\gamma\sigma\psi v_x - 2\lambda v_t = 0$. This is the backward form of the equation (1.2.44). It is to be solved for $t < 0$ in terms of the data (1.2.26) given at $t = 0$. It is also the adjoint equation of (1.2.44) (see Section 3.6). We note that if $\sigma = 1$ and $\psi = 0$, this is known as the backward telegrapher's equation.

1.2.17. Carry out the limit described in the discussion immediately preceding equation (1.2.45) for the equation of Exercise 1.2.16 and obtain $v_t = -cv_x - \frac{1}{2}\sqrt{D}\,(\sqrt{D}\,v_x)_x$. This is the backward form of the equation (1.2.45), and is again its adjoint.

1.2.18. Put $\sigma = 1$ in (1.2.38) and $\psi = -x$ in (1.2.43). Show that the limiting diffusion equation (1.2.45) has the form of the equation given in Exercise 1.1.12. Explain why the behavior of the reversal rates λ^+ and λ^- for $x < 0$ and $x > 0$ suggests that this problem is similar to that of Exercise 1.1.12, where the particle is subjected to an elastic restoring force.

1.2.19. Consider the motion of a particle in the semi-infinite interval $x < l$, and let $x = l$ be an absorbing boundary. (a) Show that for the discrete problem $\beta(l,t) = 0$ since the particle cannot reach that point from the right. For a particle that moves to the right, we have $\alpha(l, t+\tau) = p^+(l-\delta)\alpha(l-\delta, t)$. (b) As δ and τ tend to zero, if $\sigma(l) \neq 0$, conclude that $\alpha(l,t) = \beta(l,t) = 0$ are the required conditions at an absorbing boundary.

1.2.20. Let the particle move in the interval $x < l$, and assume that $x = l$ is a reflecting boundary. (a) Show that for the discrete problem we have $\beta(l, t+\tau) = p^+(l)\alpha(l,t) + [1 - p^-(l) - q^-(l)]\beta(l,t)$, $\alpha(l, t+\tau) = p^+(l-\delta)\alpha(l-\delta, t) + q^-(l-\delta)\beta(l-\delta, t) + [1 - p^+(l) - q^+(l)]\alpha(l,t)$. (b) In the limit as δ and τ tend to zero, obtain $\alpha(l,t) = \beta(l,t)$, $\alpha_t + \gamma(\sigma\alpha)_x = (\lambda^- - \lambda^+)\alpha$ as the boundary conditions at a reflecting boundary.

1.3 LAPLACE'S EQUATION AND GREEN'S FUNCTION

In the random walk problems of the preceding sections we were concerned with finding the probability that a particle located initially at the point $x = 0$ and moving in a random manner is located at a point x at a later time t. Consequently, we were dealing with time-dependent problems. In this section we consider three time-independent random walk problems in plane regions. The one- and three-dimensional versions of these problems are considered in the exercises.

Each of the problems we study involves a random walk in a bounded plane region with an absorbing boundary. The first problem examines the probability that a particle

starting at some point in the region reaches a specified point on the boundary and is absorbed there before it reaches and is absorbed at another boundary point. The second problem is essentially concerned with the probability that the particle reaches a fixed interior point before it is absorbed at the boundary. The third problem involves the determination of the expected or mean time it takes for a particle starting at some point in the interior until it is absorbed at the boundary. This is known as a mean first passage time problem. The number of steps required for the particle to reach the fixed boundary or interior point is not relevant. That is, the time it takes for the particle to reach that point is not considered or determined, so that these problems are *time independent.* For the mean first passage time problem, the possible times until absorption for each interior point are averaged out, so that there is no explicit time dependence. The first two problems may be thought to represent stationary or steady-state versions of appropriate modifications of the problems considered in Sections 1.1–1.2. In fact, the time independent form of the diffusion and telegrapher's equations (in two space dimensions) is Laplace's equation which is derived below. We remark that none of these problems is meaningful if there is a purely reflecting boundary. For in that case, the particle does not cease its motion on reaching the boundary. However, we can consider a mixed boundary condition, in which part of the boundary is absorbing and the rest of the boundary is reflecting.

Let A represent the bounded region under consideration and ∂A its (piecewise smooth) boundary curve. We enclose A and its boundary by a rectangle with sides $x = a, x = b, y = c,$ and $y = d$ where $a < b$ and $c < d$. With δ as the step length in the random walk, we assume the intervals $[a, b]$ and $[c, d]$ can be subdivided into the set of points $x_k = a+\delta k$ and $y_l = c+\delta l$, respectively, with $0 \leq k \leq n$, $0 \leq l \leq m$, $x_n = b$ and $y_m = d$. Each point that lies within A (i.e., does not lie on ∂A) is called an *interior point.* Each interior point (x_k, y_l) has four neighboring points in four perpendicular directions. If one of the neighboring points lies on ∂A or is exterior to A, we call it a *boundary point.* The points in the rectangle that are neither interior nor boundary points will not be considered.

Time-Independent Random Walks and Their Limits

In the first random walk problem, we ask for the probability that a particle starting at an interior point of the region A reaches the specified boundary point (x_i, y_j) before it reaches and is absorbed at any other boundary point. Let $v(x, y)$ be the probability that the particle starts at the interior point (x, y) and reaches the boundary point (x_i, y_j). We assume that the particle is equally likely to move to any of its four neighboring points from the interior point (x, y). Thus the probability that it moves to any of its four neighbors equals $\frac{1}{4}$. The probability that the particle reaches the boundary point (x_i, y_j) from the point (x, y) can be expressed in terms of the probability that it moves to any of its four neighboring points and reaches (x_i, y_j) from one of these points. Thus we obtain the partial difference equation

$$v(x, y) = \frac{1}{4}[v(x + \delta, y) + v(x - \delta, y) + v(x, y + \delta) + v(x, y - \delta)]. \qquad (1.3.1)$$

If (x, y) is a boundary point, we have

$$v(x,y) = \begin{cases} 1, & (x,y) = (x_i, y_j), \\ 0, & (x,y) \neq (x_i, y_j), \end{cases} \tag{1.3.2}$$

since when the particle is at a boundary point $(x, y) \neq (x_i, y_j)$ it is absorbed and cannot reach the point (x_i, y_j). If one of the neighboring points of (x, y) is a boundary point, (1.3.2) is to be used in (1.3.1).

In the limit as the step length $\delta \to 0$, the number of points in the subdivisions of $[a, b]$ and $[c, d]$ tend to infinity, and the boundary points defined actually lie on the boundary ∂A. Using Taylor's formula gives

$$\begin{cases} v(x \pm \delta, y) = v(x,y) \pm \delta\, v_x(x,y) + (\delta^2/2) v_{xx}(x,y) + O(\delta^3), \\ v(x, y \pm \delta) = v(x,y) \pm \delta\, v_y(x,y) + (\delta^2/2)\, v_{yy}(x,y) + O(\delta^3). \end{cases} \tag{1.3.3}$$

Inserting (1.3.3) into (1.3.1), dividing by δ^2, and letting $\delta \to 0$ yields

$$\frac{\partial^2 v(x,y)}{\partial x^2} + \frac{\partial^2 v(x,y)}{\partial y^2} = 0, \tag{1.3.4}$$

which is known as *Laplace's equation*. The function $v(x, y)$ is now interpreted as a probability density.

We assume that arc length s is defined on the boundary ∂A and that as $\delta \to 0$, the point (x_i, y_j) tends to the (boundary) point $(\hat{x}, \hat{y})$ on ∂A. The boundary conditions (1.3.2) are easily found to take the form

$$\int_{\partial A} v(x,y)\, ds = 1, \qquad v(x,y) = 0, \quad (x,y) \in \partial A, \quad (x,y) \neq (\hat{x}, \hat{y}). \tag{1.3.5}$$

For example, if $(\hat{x}, \hat{y})$ is a point in an open interval on the x-axis that comprises a portion of the boundary ∂A, we can set $v(x, y) = \delta(x - \hat{x})$ in that interval [where $\delta(x - \hat{x})$ is the Dirac delta function] and $v(x, y) = 0$ elsewhere on ∂A.

Green's Function

The second random walk problem we consider asks for the probability that a particle starting at an interior point (x, y) in the region A reaches a specified interior point (ξ, η) before it reaches a boundary point and is absorbed. The region A is subdivided as in the first problem, and interior and boundary points are defined as before with the step length again equal to δ. Since the problem is time-independent and the particle does not stop its motion once it first reaches (ξ, η), it is possible for the particle to pass through the point (ξ, η) more than once before it reaches and is absorbed at the boundary. Consequently, if the particle begins its motion at (ξ, η), it has unit probability of reaching (ξ, η) since it is there already. However, it can also move to one of its four neighboring points and reach (ξ, η) from there, if the neighbor is not a boundary point. Therefore, if we introduce a function $w(x, y)$ that characterizes

the prospects of a particle reaching (ξ, η) from the starting point (x, y), we cannot consider $w(x, y)$ to be a probability distribution since it may assume values exceeding unity. In particular, $w(\xi, \eta) \geq 1$, as we have seen.

The preceding random walk problem (as well as the first one of this section) was considered by *Courant, Friedrichs*, and *Lewy* in a classic early paper on difference methods for the partial differential equations of applied mathematics. They introduced the probabilities that a particle starting at (x, y) reaches (ξ, η) in $0, 1, 2, \ldots, n, \ldots$ steps and defined a function of (x, y) that equals the sum of all these probabilities. This function gives the expected number of steps that it takes for a particle starting at (x, y) to reach (ξ, η) before it is absorbed at the boundary. We shall take the function $w(x, y)$ introduced above to be defined in this manner. This function satisfies a difference equation, as we now show.

If the point (x, y) is a boundary point, we must have

$$w(x, y) = 0, \qquad (x, y) \quad \text{a boundary point,} \tag{1.3.6}$$

since the particle is absorbed at a boundary point and cannot reach (ξ, η) from there. If (x, y) is an interior point not equal to (ξ, η), we obtain the difference equation

$$w(x, y) = \frac{1}{4}[w(x + \delta, y) + w(x - \delta, y) + w(x, y + \delta) + w(x, y - \delta)] \tag{1.3.7}$$

on expressing the expectation $w(x, y)$ of reaching (ξ, η) from (x, y) in terms of the expectation of reaching (ξ, η) from each of the four neighboring points $(x \pm \delta, y)$ and $(x, y \pm \delta)$, on condition that the particle goes from each of these points to the point (x, y) in the next step. Since the particle is equally likely to move to each of its four neighboring points, the conditional probability equals $\frac{1}{4}$. If $(x, y) = (\xi, \eta)$, we have

$$w(\xi, \eta) = 1 + \frac{1}{4}[w(\xi + \delta, \eta) + w(\xi - \delta, \eta) + w(\xi, \eta + \delta) + w(\xi, \eta - \delta)], \tag{1.3.8}$$

since the particle has unit probability of reaching (ξ, η) considering that it is there to begin with, and it retains the possibility of reaching (ξ, η) from each of its four neighboring points as before.

For small δ, (1.3.7) and (1.3.8) take the form

$$\frac{\partial^2 w(x, y)}{\partial x^2} + \frac{\partial^2 w(x, y)}{\partial y^2} = \begin{cases} O(\delta^2), & (x, y) \neq (\xi, \eta), \\ -4/\delta^2 + O(\delta^2), & (x, y) = (\xi, \eta), \end{cases} \tag{1.3.9}$$

on using (1.3.3). This shows that $w(x, y)$ satisfies Laplace's equation at interior points $(x, y) \neq (\xi, \eta)$ as $\delta \to 0$. At the point (ξ, η), the right side of (1.3.9) blows up as $\delta \to 0$. Now as $\delta \to 0$, $w(x, y)$ is to be understood as a density function, so that the integral of $w(x, y)$ over some small neighborhood of (x, y) characterizes the property in which we are interested. If we consider a square with center at (ξ, η) and with side proportional to the step length δ, the area of the square multiplied by $4/\delta^2$ tends to a finite nonzero limit as $\delta \to 0$. Since the right side of (1.3.9) vanishes for $(x, y) \neq (\xi, \eta)$ and its integral over the aforementioned square has a

finite nonzero limit as $\delta \to 0$, we conclude that it must be proportional in the limit to the two-dimensional Dirac delta function.

The two-dimensional Dirac delta function with singular point (ξ, η) has the properties (see Section 7.2):

$$\iint_R \delta(x-\xi)\delta(y-\eta)\,dx\,dy = 1,\ \delta(x-\xi)\delta(y-\eta) = 0,\ (x,y) \neq (\xi,\eta), \quad (1.3.10)$$

where R any open region containing the point (ξ, η). As in Example 1.2, it follows that for continuous $f(x, y)$ we have the further property

$$\iint_R f(x,y)\delta(x-\xi)\delta(y-\eta)\,dx\,dy = f(\xi,\eta). \quad (1.3.11)$$

As the step length $\delta \to 0$, the boundary points of the discrete problem tend to points on ∂A, and (1.3.6) now states that $w(x, y)$ vanishes on ∂A. To analyze the properties of the solution of this boundary value problem, it is convenient to replace $w(x, y)$ by the *Green's function* $K(x, y; \xi, \eta)$ for this problem. The Green's function $K(x, y; \xi, \eta)$ for Laplace's equation with homogeneous boundary conditions is defined to be a solution of

$$\frac{\partial^2 K(x,y;\xi,\eta)}{\partial x^2} + \frac{\partial^2 K(x,y;\xi,\eta)}{\partial y^2} = -\delta(x-\xi)\delta(y-\eta), \qquad (x,y) \in A, \quad (1.3.12)$$

which satisfies the boundary condition

$$K(x,y;\xi,\eta) = 0, \qquad (x,y) \in \partial A. \quad (1.3.13)$$

As was shown above, the density function $w(x, y)$ differs from Green's function only by a constant factor. Thus both $w(x, y)$ and $K(x, y; \xi, \eta)$ are solutions of (special) inhomogeneous forms of Laplace's equation. (Green's functions are a useful tool for solving boundary value problems for partial differential equations and are discussed in Chapter 7.)

The two random walk problems considered so far in this section are not completely unrelated, since one problem asks for the probability that an interior point is reached and the other seeks the probability that a boundary point is reached. The following example shows that the two problems are indeed connected, by establishing a relationship between the density function $v(x, y)$ for the first random walk problem and the Green's function $K(x, y; \xi, \eta)$ for the second random walk problem.

Example 1.6. Green's Theorem. We formally apply Green's second theorem to the functions $v(x, y)$ and $K(x, y; \xi, \eta)$. Integrating over the region A and its boundary ∂A, we have

$$\iint_A \left(v\,\nabla^2 K - K\,\nabla^2 v\right) dx\,dy = \int_{\partial A} \left(v\,\frac{\partial K}{\partial n} - K\,\frac{\partial v}{\partial n}\right) ds, \quad (1.3.14)$$

where $\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2$ is the Laplacian operator, $\partial/\partial n$ is a derivative in the direction of the exterior normal to the boundary ∂A, and s is the arc length on ∂A. We have $\nabla^2 v(x,y) = 0$ and $\nabla^2 K(x,y;\xi,\eta) = -\delta(x-\xi)\delta(y-\eta)$, in view of (1.3.4) and (1.3.12). Further, (1.3.5) implies that on the boundary, $v(x,y)\partial K(x,y;\xi,\eta)/\partial n$ equals $v(x,y)\partial K(\hat{x},\hat{y};\xi,\eta)/\partial n$. Combined with the fact that $K(x,y;\xi,\eta)$ vanishes on ∂A, we obtain from (1.3.14)

$$v(\xi,\eta) = -\int_{\partial A} v\,\frac{\partial K}{\partial n}\,ds = -\frac{\partial K(\hat{x},\hat{y};\xi,\eta)}{\partial n}, \tag{1.3.15}$$

where (1.3.5) and (1.3.11) were used. This shows that if the Green's function K can be determined, the boundary value problem (1.3.4)–(1.3.5) can be solved since (ξ,η) is an arbitrary interior point in A.

Equation (1.3.15) has the following interpretation. The probability density $v(\xi,\eta)$ characterizes the probability that particles starting in a neighborhood of (ξ,η) reach the boundary point $(\hat{x},\hat{y})$. The normal derivative $\partial K(\hat{x},\hat{y};\xi,\eta)/\partial n$ is a measure of the flux of particle density associated with particles originating near the boundary point $(\hat{x},\hat{y})$ and reaching the interior point (ξ,η). Therefore, $-\partial K/\partial n$ is a measure of the flux in the reverse direction, and this should essentially equal $v(\xi,\eta)$.

Mean First Passage Times and Poisson's Equation

The third random walk problem we consider determines the mean or expected time it takes for a particle starting at an interior point (x,y) in the region A until it is absorbed at the boundary. This yields what is known as the *mean first passage time* for each point and we denote it by $u(x,y)$. Assuming that the particle takes steps of length δ at intervals of time τ and is equally likely to move to each of its four neighboring points from the point (x,y), we obtain the difference equation

$$u(x,y) = \tau + \frac{1}{4}[u(x+\delta,y) + u(x-\delta,y) + u(x,y+\delta) + u(x,y-\delta)]. \tag{1.3.16}$$

The expected time until absorption $u(x,y)$ is expressed in terms of the expected time until absorption for each of the four neighboring points, multiplied by the probability $\frac{1}{4}$ that the particle moves to each of these points. We must also add the time τ it takes for the particle to reach one of the neighboring points in a single step. If (x,y) is a boundary point, $u(x,y) = 0$ since the time until absorption at any boundary point is zero. In the limit as $\delta \to 0$ and $\tau \to 0$, while $\delta^2/2\tau \to D$, we obtain

$$\frac{1}{2}D(u_{xx}(x,y) + u_{yy}(x,y)) = -1 \tag{1.3.17}$$

on using (1.3.3). The inhomogeneous form of Laplace's equation, of which (1.3.17) is a special case, is known as *Poisson's equation*. The boundary condition for $u(x,y)$ is

$$u(x,y) = 0, \qquad (x,y) \in \partial A. \tag{1.3.18}$$

Position-Dependent Random Walks and Their Limits

Each of the three random walk problems discussed in this section can be generalized to permit position-dependent jump probabilities as was done in Sections 1.1 and 1.2. To do so, we assume that the probabilities that a particle located at the point (x, y) moves to the right or to the left are given by p_1 and q_1, respectively, while those for upward or downward motion are given by p_2 and q_2, respectively. These probabilities are defined as

$$p_i(x,y) = \frac{1}{4}[a_i(x,y) + b_i(x,y)\delta], \quad q_i(x,y) = \frac{1}{4}[a_i(x,y) - b_i(x,y)\delta], \quad i = 1,2. \tag{1.3.19}$$

The a_i and b_i are chosen such that $0 < p_1 + q_1 + p_2 + q_2 \leq 1$.

For the first random walk problem, the difference equation (1.3.1) is replaced by

$$v(x,y) = [1 - p_1 - q_1 - p_2 - q_2]v(x,y) + p_1 v(x+\delta, y) + q_1 v(x-\delta, y) + p_2 v(x, y+\delta) + q_2 v(x, y-\delta). \tag{1.3.20}$$

The coefficients of the v terms are all evaluated at (x, y), and the coefficient of $v(x, y)$ on the right of equation (1.3.20) represents the probability that the particle rests at the point (x, y). On expanding all the terms in (1.3.20) in Taylor series, as in (1.3.3), we easily obtain

$$a_1(x,y)\, v_{xx}(x,y) + a_2(x,y)\, v_{yy}(x,y) + 2b_1(x,y)\, v_x(x,y) + 2b_2(x,y)\, v_y(x,y) = 0 \tag{1.3.21}$$

on dividing by δ^2 and letting δ tend to zero. The boundary condition is again (1.3.5).

For the second random walk problem, the difference equation (1.3.7) is replaced by

$$w(x,y) = [1 - p_1 - q_1 - p_2 - q_2]w(x,y) + q_1 w(x+\delta, y) + p_1 w(x-\delta, y) + q_2 w(x, y+\delta) + p_2 w(x, y-\delta). \tag{1.3.22}$$

However, in view of our definition of $w(x, y)$, the probabilities p_i and q_i must all be evaluated at the same points as the expectations w that they multiply. The same is true for the modified form of (1.3.8). As a result, when all the terms are expanded for small δ, and we let δ tend to zero, after dividing by δ^2, we obtain for the Green's function $K(x, y; \xi, \eta)$ associated with this problem the equation

$$(a_1 K)_{xx} + (a_2 K)_{yy} - 2(b_1 K)_x - 2(b_2 K)_y = -\delta(x-\xi)\delta(y-\eta), \tag{1.3.23}$$

with the a_1, b_1, a_2, b_2 as functions of x, y. Again, K must vanish on the boundary. The differential operator acting on K in (1.3.23) is the adjoint of the operator acting on v in (1.3.21), as follows from our discussion in Section 3.6. This observation yields a relationship between v and K similar to that given in Example 1.6 and based on a generalization of Green's theorem. This is considered in the exercises. Similarly, the generalization of the third random walk problem considered in this section is examined in the exercises.

Properties of Partial Difference Equations and Related PDEs

We conclude this section with a number of comments regarding the properties of boundary value problems for Laplace's equation and the related equations derived in this section that are suggested by the difference equation treatment presented above.

The difference equation (1.3.1) together with the boundary condition (1.3.2), when applied to all interior points in the subdivision of A, yields a simultaneous system of equations for the functions $v(x_k, y_l)$ where (x_k, y_l) is a typical interior point. All the boundary data play a role in the solution for v at the interior points. The same is true for the other equations considered in this section. This contrasts with the situation encountered in the preceding sections. Although the solution was required for all $t > 0$, the solution for fixed t did not depend on data for later values of t.

Further, the difference equation (1.3.1) characterizes the value of $v(x, y)$ at the center of a diamond-shaped quadrilateral, as the mean or average of its values at the four vertices. With (x, y) as its center, the diamond has the points $(x \pm \delta, y)$ and $(x, y \pm \delta)$ as its vertices. This *mean value property* carries over to functions $v(x, y)$ that satisfy Laplace's equation. It is shown later that any solution $v(x, y)$ of (1.3.4) equals the average of its values on a circle with center at (x, y) as long as the circle is interior to the region A.

The foregoing mean value property for the difference equation implies that the maximum and minimum values of $v(x, y)$ must be assumed at boundary points of A. This follows since $v(x, y)$ is the average of its neighboring values and, therefore, cannot be greater than or less than those values. This *maximum* and *minimum principle* will be shown to be valid for certain boundary value problems for Laplace's equation. In fact, it follows from (1.3.1) that if the maximum or minimum of $v(x, y)$ is attained at an interior point of A, $v(x, y)$ must be constant throughout A. For if (x, y) is an interior maximum or minimum point, $v(x, y)$ cannot be represented as the average of its four neighboring values if any one of them is, respectively, smaller or larger than the value of v at (x, y). This property carries over to solutions of Laplace's equation, as will be shown, and is known as the *strong maximum and minimum principle.*

The representation (1.3.20) for $v(x, y)$ implies that it also satisfies a maximum and minimum principle. Consequently, we expect such a principle to be valid for solutions of (1.3.26), and this will be demonstrated. Further questions relating to the positivity of solutions of Poisson's equation (1.3.17) and the positivity of the Green's function determined from (1.3.12)–(1.3.13), are examined later in the book.

Exercises 1.3

1.3.1. (a) Show that in the one-dimensional case, if the particle is equally likely to move to the right or to the left, (1.3.1) is replaced by $v(x, y) = \frac{1}{2}[v(x-\delta)+v(x+\delta)]$. (b) Show that in the three-dimensional case, when the particle is equally likely to move to any of its six neighboring points, we obtain $v(x, y, z) = \frac{1}{6}[v(x + \delta, y, z) + v(x - \delta, y, z) + v(x, y - \delta, z) + v(x, y + \delta, z) + v(x, y, z - \delta) + v(x, y, z + \delta)]$.

1.3.2. Show that the limiting differential equations for Exercise 1.3.1(a) and (b) are (a) $v_{xx} = 0$ and (b) $v_{xx} + v_{xx} + v_{xx} = 0$, respectively.

1.3.3. Consider the interval $0 \leq x \leq 1$. Let $v(x)$ be the probability that the particle reaches the boundary point $x = 1$ and is absorbed before it reaches and is absorbed at the boundary $x = 0$. (a) Show that the appropriate boundary conditions in this case are $v(0) = 0$ and $v(1) = 1$. (b) Show that $v(x) = x$, which satisfies the boundary conditions of part (a), is a solution of both the difference equation of Exercise 1.3.1(a) and the (one-dimensional) differential equation of Exercise 1.3.2(a).

1.3.4. We define the Green's function $K(x;\xi)$ in the one-dimensional case, for the interval $0 \leq x \leq 1$, to be the solution of the boundary value problem $\partial^2 K(x;\xi)/\partial x^2 = -\delta(x-\xi), K(0;\xi) = K(1;\xi) = 0,\ 0 < \xi < 1$. (a) By integrating the equation for $K(x;\xi)$ over a small neighborhood of $x = \xi$, show that $[\partial K/\partial x]\big|_{x\uparrow\xi}^{x\downarrow\xi} = -1$ (i.e., the jump in $\partial K/\partial x$ at $x = \xi$ is -1). Assuming that $K(x;\xi)$ is continuous at $x = \xi$, define $K(x,\xi)$ as $K(x;\xi) = K_1(x;\xi),\ 0 \leq x \leq \xi$, and $K(x;\xi) = K_2(x;\xi),\ \xi \leq x \leq 1$. Show that we then obtain the equations $\partial^2 K_1(x;\xi)/\partial x^2 = 0,\ 0 < x < \xi,\ \partial^2 K_2(x;\xi)/\partial x^2 = 0, \xi \leq x \leq 1$. and the supplementary conditions $K_1(0;\xi) = K_2(l;\xi) = 0,\ K_1(\xi;\xi) = K_2(\xi;\xi),\ \partial K_2(\xi;\xi)/\partial x - \partial K_1(\xi;\xi)/\partial x = -1$. (b) Show that the solution of part (a) is $K_1(x;\xi) = (1-\xi)x,\ K_2(x;\xi) = (1-x)\xi$.

1.3.5. (a) Apply a one-dimensional form of Green's theorem (1.3.14) to show that with $v(x)$ and $K(x;\xi)$ defined as in the two preceding exercises we have $v(\xi) = -\partial K(1;\xi)/\partial x$. (b) Demonstrate that the result of part (a) is correct by using the explicit forms for the functions $v(x)$ and $K(x;\xi)$.

1.3.6. Solve the difference equation (1.3.1) in the square $0 \leq x \leq 3$ and $0 \leq y \leq 3$ at the points with coordinates $x = 1, 2$ and $y = 1, 2$, assuming that $\delta = 1$ and $(x_i, y_j) = (0, 1)$ in (1.3.2).

1.3.7. Show that $v(x, y) = 1$ is a solution of the boundary value problem for Laplace's equation in the square $0 \leq x \leq 1,\ 0 \leq y \leq 1$, with the boundary conditions $v(x, 0) = 1$ and $\partial v/\partial n = 0$ on the other three sides of the square. Noting that $\partial v/\partial n = 0$ corresponds to a reflecting boundary condition, interpret the solution.

1.3.8. (a) Show that in the one-dimensional case, if the particle is equally likely to move to the right or to the left, the mean first passage time $u(x)$ satisfies the difference equation $u(x, y) = \tau + \frac{1}{2}[u(x-\delta) + u(x+\delta)]$. (b) Show that in the three-dimensional case, if the particle is equally likely to move to any of its six neighboring points, the mean first passage time $u(x, y, z)$ satisfies the equation $u(x, y, z) = \tau + \frac{1}{6}[u(x+\delta, y, z) + u(x-\delta, y, z) + u(x, y-\delta, z) + u(x, y+\delta, z) + u(x, y, z-\delta) + u(x, y, z+\delta)]$.

1.3.9. Show that if $\lim \delta^2/\tau = D$ in part (a) and $\lim \delta^2/3\tau = D$ in part (b) of Exercise 1.3.8, the limiting differential equations for the one- and three-dimensional difference equations are, respectively, (a) $\frac{1}{2}Du''(x) = -1$ and (b) $\frac{1}{2}D[u_{xx} + u_{yy} + u_{zz}] = -1$.

1.3.10. Show that in the one-dimensional case, the equations (1.3.21) and (1.3.23) take the form $L[v] = a(x)v_{xx} + 2b(x)v_x = 0$, $L^*[K] = (aK)_{xx} - 2(bK)_x = -\delta(x - \xi)$.

1.3.11. Demonstrate that with the operators L and L^* defined as in Exercise 1.3.10, we have $vL^*[K] - KL[v] = [v(aK)_x - aKv_x - 2bKv]_x$.

1.3.12. Given the interval $0 \leq x \leq 1$ with $v(0) = 0$, $v(1) = 1$, and $K(0;\xi) = K(1;\xi) = 0$, integrate the expression given in Exercise 1.3.11 over the interval. Use the equations in Exercise 1.3.10 to conclude that $v(\xi) = -a(1)\partial K(1;\xi)/\partial x$. This is the analog of the result (1.3.15) and Exercise 1.3.5. A similar result is valid in the two- and three-dimensional cases.

1.3.13. Write down the equations that correspond to (1.3.21) and (1.3.23) in the three-dimensional case.

1.3.14. Show that the mean first passage time $u(x)$ for an inhomogeneous one-dimensional medium, with p and q given as in (1.1.30), satisfies the ODE $\frac{1}{2}D(x)u''(x) + c(x)u'(x) = -1$, where c and D are defined as in (1.1.32).

1.3.15. Consider the interval $0 \leq x \leq 1$ and the mean first passage time problem of Exercise 1.3.9(a). (a) Find the solution $u(x)$ if $u(0) = u(1) = 0$. (b) Determine $u(x)$ if $u(0) = u'(1) = 0$. In this case, the boundary at $x = 0$ is absorbing and that at $x = 1$ is reflecting. (c) If $u'(0) = u'(1) = 0$, in which case both boundary points are reflecting, show that the problem has no solution.

1.3.16. Use Green's theorem (1.3.14) with $v(x, y) = 1$ to show that the Green's function $K(x, y; \xi, \eta)$ that satisfies (1.3.12) and whose normal derivative vanishes on the boundary of A, does not exist. That is, in the case of a reflecting boundary, the Green's function problem has no solution.

1.3.17. Show that $v(x, y) = (x + iy)^n$, where $i = \sqrt{-1}$, is a solution of Laplace's equation (1.3.4) for all integers $n > 0$.

1.3.18. Consider $v(x, y) = (x + iy)^2$ and write it in the form $v(x, y) = f(x, y) + ig(x, y)$, where $f(x, y)$ and $g(x, y)$ are the real and imaginary parts of the function $v(x, y)$. Show that $f(x, y)$ and $g(x, y)$ have no relative maxima or minima in the (x, y)-plane and that both satisfy Laplace's equation.

1.3.19. Show that $v(x, y, z) = (x + iy\cos\theta + iz\sin\theta)^n$ is a solution of Laplace's equation in three dimensions (see Exercise 1.3.2) for all integers $n > 0$ and for $0 \leq \theta < 2\pi$.

1.3.20. Write $v(x, y, z) = (x + iy\cos\theta + iz\sin\theta)^2$ in the form $v(x, y, z) = F(x, y, z) + iG(x, y, z)$, where $F(x, y, z)$ and $G(x, y, z)$ are real-valued functions. Show that $F(x, y, z)$ and $G(x, y, z)$ are solutions of Laplace's equation and that they have no relative maxima and minima in (x, y, z)-space.

1.4 RANDOM WALKS AND FIRST ORDER PDEs

T he random walks considered in the preceding sections were all related, in some way, to the theory of Brownian motion. They gave rise to second order partial differential equations in the continuum limit. We now consider random walks that give rise to *first order partial differential equations* in the limit. (This possibility was indicated in Sections 1.1 and 1.2.) In this section our emphasis is on the relation between random walk difference equations and various difference approximations that are used in the numerical solution of first order partial differential equations by the *method of finite differences*, which is presented in Chapter 11. A variety of partial difference equations is generated and questions relating to their *stability properties* are addressed. Solution methods for first order partial differential equations are considered in Chapter 2.

Random Walks and Linear First Order PDEs: Constant Transition Probabilities

We begin by reconsidering the unrestricted random walk introduced at the beginning of Section 1.1. The particle can move to the right, to the left or remain stationary at each step, with steps of length δ occurring at time intervals of length τ. As before, the probability for a step to the right is p and that for a step to the left is q. Again, the mean displacement of the particle equals c, but now we require a zero variance, so that the particle appears to move deterministically. As shown in the discussion following (1.1.8), this requires that $p-q$ does not tend to zero in the continuum limit. As a result, D as given in (1.1.8) equals 0.

These results can be realized if we set

$$p = \frac{1}{2}(a+b), \qquad q = \frac{1}{2}(a-b), \tag{1.4.1}$$

with a and b restricted so that p and q represent probabilities. [This representation differs from (1.1.9), where b is replaced by $b\delta$.] Thus, with $r = 1/\tau$, we have $a\delta^2/\tau \to D$ as before, but $b\delta/\tau \to c$. Since we require a finite value for c, we conclude that $D = 0$. The difference equation for v is given by (1.1.16),

$$v(x,t+\tau) = (1-p-q)v(x,t) + pv(x-\delta,t) + qv(x+\delta,t), \tag{1.4.2}$$

and can be expressed as

$$v(x,t+\tau) = (1-a)\,v(x,t) + \left(\frac{a}{2}+\frac{b}{2}\right)v(x-\delta,t) + \left(\frac{a}{2}-\frac{b}{2}\right)v(x+\delta,t). \tag{1.4.3}$$

The Taylor expansion yields

$$v_t(x,t) = -b\,\frac{\delta}{\tau}\,v_x(x,t) + \frac{1}{2}a\,\frac{\delta^2}{\tau}\,v_{xx}(x,t) + O(\tau) + O\left(\frac{\delta^2}{\tau}\right). \tag{1.4.4}$$

In the limit as $\delta \to 0$ and $\tau \to 0$, we obtain the first order PDE

$$\frac{\partial v(x,t)}{\partial t} + c\,\frac{\partial v(x,t)}{\partial x} = 0. \tag{1.4.5}$$

This is known as the *unidirectional wave equation* and characterizes wave motion to the right if $c > 0$ and to the left if $c < 0$. (It is also referred to, occasionally, as the *advection equation.*)

We see that δ/τ must have a finite nonzero limit and we set $b = c\tau/\delta$. Then the difference equation (1.4.3) becomes

$$v(x,t+\tau) = (1-a)v(x,t) + \left(\frac{a}{2} + \frac{c\tau}{2\delta}\right) v(x-\delta,t) + \left(\frac{a}{2} - \frac{c\tau}{2\delta}\right) v(x+\delta,t). \tag{1.4.6}$$

The value of a plays no role in the limit. But since $p + q = a$, we conclude that $0 \le a \le 1$, as the sum of the probabilities cannot exceed 1. Furthermore, $0 \le p, q \le 1$ implies that $|c|\,\tau/\delta \le a$, and this places a restriction on the values of δ and τ in terms of a, for a fixed c (in the discrete case). (We note that no matter how a is specified we must have $|c|\,\tau/\delta \le 1$.)

We now introduce three choices for a and consider their relation to well-known finite difference schemes for the numerical solution of (1.4.5) discussed in Chapter 11. To begin, we assume that $c > 0$ and choose a so that $q = 0$. This means that the particle can only move to the right and this corresponds to what is known as the *explicit forward-backward scheme* for the solution of (1.4.5). If $q = 0$, we must set $a = c\tau/\delta$ and (1.4.6) becomes

$$v(x,t+\tau) = \left(1 - \frac{c\tau}{\delta}\right) v(x,t) + \frac{c\tau}{\delta}\, v(x-\delta,t). \tag{1.4.7}$$

For the difference equation to retain its probabilistic interpretation we must have $c\tau/\delta \le 1$. This agrees with the von Neumann stability condition for the explicit forward-backward scheme determined in Chapter 11.

Next we assume that $c < 0$ and choose a so that $p = 0$. This means that the particle can only move to the left, and this corresponds to what is known as the *explicit forward-forward scheme* for the solution of (1.4.5). If $p = 0$, we must set $a = -c\tau/\delta$. and (1.4.6) becomes

$$v(x,t+\tau) = \left(1 + \frac{c\tau}{\delta}\right) v(x,t) - \frac{c\tau}{\delta}\, v(x+\delta,t). \tag{1.4.8}$$

For the difference equation to retain its probabilistic interpretation, we must have $-c\tau/\delta \le 1$. This agrees with the von Neumann stability condition for the explicit forward-forward scheme determined in Chapter 11.

Finally, we choose $a = 1$. As a result, there can be nonzero probabilities for steps to the right and to the left. This corresponds to what is known as the *Lax-Friedrichs scheme* for the solution of (1.4.5). Then (1.4.6) becomes

$$v(x,t+\tau) = \left(\frac{1}{2} + \frac{c\tau}{2\delta}\right) v(x-\delta,t) + \left(\frac{1}{2} - \frac{c\tau}{2\delta}\right) v(x+\delta,t). \tag{1.4.9}$$

For the difference equation to retain its probabilistic interpretation we must have $|c|\,\tau/\delta \leq 1$. This agrees with the von Neumann stability condition for the Lax-Friedrichs scheme determined in Chapter 11.

If we assume that the particle is initially located at $x = 0$, the initial condition for (1.4.5) is $v(x,0) = \delta(x)$, and the solution is $v(x,t) = \delta(x - ct)$. Noting the behavior of the Dirac delta function $\delta(x)$, we find that $v(x,t)$ vanishes for $x - ct \neq 0$, so that the particle moves horizontally with the velocity c. We see that even for values of a that yield nonzero probabilities for motion to the left or to the right, deterministic motion with velocity c results.

It is possible to introduce an absorbing or reflecting boundary condition at some value of x. If the initial position of the particle is at $x = 0$, then if $c > 0$, the particle moves to the right and a boundary condition can be inserted at some point $x > 0$. If $c < 0$, there can be a boundary condition at a point $x < 0$. We do not consider these boundary conditions here.

Random Walks and Linear First Order PDEs: Variable Transition Probabilities

Next we consider random walks with position-dependent probabilities $p(x)$ and $q(x)$. Then $v(x,t)$ satisfies the difference equation

$$v(x,t+\tau) = [1 - p(x) - q(x)]\, v(x,t) + p(x-\delta)\, v(x-\delta,t) + q(x+\delta)\, v(x+\delta,t), \tag{1.4.10}$$

with $p(x) = [a(x) + b(x)]/2$, $q(x) = [a(x) - b(x)]/2$. If we put $c(x) = b(x)\delta/\tau$, the difference equation (1.4.10) becomes

$$v(x,t+\tau) = [1 - a(x)]\, v(x,t) + \left(\frac{a(x-\delta)}{2} + \frac{c(x-\delta)\tau}{2\delta}\right) v(x-\delta,t)$$

$$+ \left(\frac{a(x+\delta)}{2} - \frac{c(x+\delta)\tau}{2\delta}\right) v(x+\delta,t). \tag{1.4.11}$$

Using the Taylor expansion as before and going to the limit as $\delta \to 0$ and $\tau \to 0$, we obtain the first order partial differential equation

$$\frac{\partial v(x,t)}{\partial t} + \frac{\partial (c(x)v(x,t))}{\partial x} = 0. \tag{1.4.12}$$

We observe that (1.4.12) is not simply a generalization of (1.4.5) with c replaced by $c(x)$, but contains an additional term $c'(x)v(x,t)$.

If we modify the random walk problem in the manner that led to the difference equation (1.1.33), we obtain instead of (1.4.11),

$$v(x,t) = [1 - a(x)]\, v(x, t+\tau) + \left(\frac{a(x)}{2} + \frac{c(x)\tau}{2\delta}\right) v(x+\delta, t+\tau)$$

$$+ \left(\frac{a(x)}{2} - \frac{c(x)\tau}{2\delta}\right) v(x-\delta, t+\tau). \tag{1.4.13}$$

In the limit we obtain in place of (1.4.12),

$$\frac{\partial v(x,t)}{\partial t} + c(x)\frac{\partial v(x,t)}{\partial x} = 0. \tag{1.4.14}$$

Solution methods for (1.4.12) and (1.4.14) are presented in Section 2.2.

In correspondence with our discussion above, we consider three special choices of $a(x)$ that we apply to the difference equation (1.4.11). [The same approach works for (1.4.13).] To begin, we assume that $c(x) > 0$ and choose $a(x)$ so that $q(x) = 0$. This means that the particle can only move to the right and this corresponds to the *explicit forward-backward scheme* introduced above for the solution of (1.4.5). If $q(x) = 0$ we must set $a(x) = c(x)\tau/\delta$, and (1.4.11) becomes

$$v(x, t+\tau) = \left(1 - \frac{c(x)\tau}{\delta}\right) v(x,t) + \left(\frac{c(x)\tau}{\delta}\right) v(x-\delta, t). \tag{1.4.15}$$

For the difference equation to retain its probabilistic interpretation we must have $c(x)\tau/\delta \leq 1$. This generalizes the von Neumann stability condition for the explicit forward-backward scheme in the case of constant coefficients.

Next we assume that $c(x) < 0$ and choose $a(x)$ so that $p(x) = 0$. This means that the particle can only move to the left and this corresponds to the *explicit forward-forward scheme* introduced above for the solution of (1.4.5). If $p(x) = 0$ we must set $a = -c(x)\tau/\delta$, and (1.4.11) becomes

$$v(x, t+\tau) = \left(1 + \frac{c(x)\tau}{\delta}\right) v(x,t) - \left(\frac{c(x)\tau}{\delta}\right) v(x+\delta, t). \tag{1.4.16}$$

For the difference equation to retain its probabilistic interpretation we must have $-c(x)\tau/\delta \leq 1$. This generalizes the Neumann stability condition for the explicit forward-forward scheme in the case of constant coefficients.

Finally we put $a(x) = 1$. As a result, there can be nonzero probabilities for steps to the right and to the left. This corresponds to the *Lax-Friedrichs scheme* introduced above for the solution of (1.4.5). Then the difference equation (1.4.11) becomes

$$v(x, t+\tau) = \left(\frac{1}{2} + \frac{c(x)\tau}{2\delta}\right) v(x-\delta, t) + \left(\frac{1}{2} - \frac{c(x)\tau}{2\delta}\right) v(x+\delta, t). \tag{1.4.17}$$

For the difference equation to retain its probabilistic interpretation we must have $|c(x)|\,\tau/\delta \leq 1$. This generalizes the von Neumann stability condition for the Lax-Friedrichs scheme in the case of constant coefficients.

Random Walks and Nonlinear First Order PDEs

Each of the three random walks considered above can be generalized by permitting the probabilities p and q to depend on $v(x,t)$. For the *explicit forward-backward scheme* we put $p(x,t) = (f(v(x,t))\tau)/(v(x,t)\delta)$, $q(x,t) = 0$, where f is an arbitrary differentiable function. Since p is a probability, we must have $0 \leq [f(v(x,t))\tau]/[v(x,t)\delta] \leq 1$. [If $f(v(x,t)) = cv(x,t)$ the problem reduces to the constant coefficient case.] The difference equation (1.4.10) takes the form

$$v(x,t+\tau) = \left(1 - \frac{f(v(x,t))\tau}{v(x,t)\delta}\right) v(x,t) + \frac{f(v(x-\delta,t))\tau}{\delta}. \tag{1.4.18}$$

Using the Taylor expansion as before and going to the limit as $\delta \to 0$ and $\tau \to 0$, we obtain the first order quasilinear partial differential equation (in conservation form)

$$\frac{\partial v(x,t)}{\partial t} + \frac{\partial (f(v(x,t))}{\partial x} = 0. \tag{1.4.19}$$

For the *explicit forward-forward scheme* we put $p(x,t) = 0$, $q(x,t) = -[f(v(x,t))\tau]/[v(x,t)\delta]$, where f is an arbitrary differentiable function. Since q is a probability we must have $0 \leq -[f(v(x,t))\tau]/[v(x,t)\delta] \leq 1$. [If $f(v(x,t)) = cv(x,t)$ the problem reduces to the constant coefficient case.] The difference equation (1.4.10) takes the form

$$v(x,t+\tau) = \left(1 + \frac{f(v(x,t))\tau}{v(x,t)\delta}\right) v(x,t) - \frac{f(v(x+\delta,t))\tau}{\delta}. \tag{1.4.20}$$

Using the Taylor expansion as before and going to the limit as $\delta \to 0$ and $\tau \to 0$, we again obtain the first order quasilinear PDE (1.4.19).

In the *Lax-Friedrichs scheme* we put $p(x,t) = 1/2 + [f(v(x,t))\tau]/[2v(x,t)\delta]$, $q(x,t) = 1/2 - [f(v(x,t))\tau]/[2\,v(x,t)\delta]$, where f is an arbitrary differentiable function. We must have $0 \leq [|f(v(x,t))|\,\tau]/[v(x,t)\delta] \leq 1$, since p and q are probabilities. The difference equation (1.4.17) takes the form

$$v(x,t+\tau) = \left[\frac{1}{2} + \frac{f(v(x-\delta,t))\tau}{2v(x-\delta,t)\delta}\right] v(x-\delta,t) + \left[\frac{1}{2} - \frac{f(v(x+\delta,t))\tau}{2v(x+\delta,t)\delta}\right] v(x+\delta,t). \tag{1.4.21}$$

Using the Taylor expansion as before and going to the limit as $\delta \to 0$ and $\tau \to 0$ again yields the first order quasilinear PDE (1.4.19).

We can write (1.4.19) as $\partial v(x,t)/\partial t + f'(v(x,t))\partial v(x,t)/\partial x = 0$. Then, if we put $f'(z) = c(z)$, it takes the form

$$\frac{\partial v(x,t)}{\partial t} + c(v(x,t))\frac{\partial v(x,t)}{\partial x} = 0. \tag{1.4.22}$$

Solution methods for (1.4.22) are presented in Section 2.3. If $c(z) = c$, a constant, we obtain the unidirectional wave equation $\partial v(x,t)/\partial t + c\,\partial v(x,t)/\partial x = 0$.

Of particular interest is the case $c(z) = z$, which yields

$$\frac{\partial v(x,t)}{\partial t} + v(x,t)\,\frac{\partial v(x,t)}{\partial x} = 0. \tag{1.4.23}$$

This is known as the *inviscid Burgers' equation* and will be studied in some detail.

Exercises 1.4

1.4.1. Let the parameter a in (1.4.6) be chosen as $a = \theta|c|\tau/\delta$, where θ is to be specified. This is known as the *θ-scheme*. Assign values of θ that give rise to the forward-backward, forward-forward, and Lax-Friedrichs schemes considered in the text.

1.4.2. Determine a restriction on θ in Exercise 1.4.1 so that p and q (as given in (1.4.1)) retain their probabilistic interpretation.

1.4.3. Let $\theta = |c|\tau/\delta$ in Exercise 1.4.1. Determine that unless $\theta = 1$ no probabilistic interpretation can be given to the resulting scheme. It will be seen in Chapter 11 that this choice of θ results in the Lax-Wendroff difference scheme.

1.4.4. Consider the backward difference equation $v(x,t) = (1-p-q)v(x,t+\tau) + qv(x-\delta,t+\tau) + pv(x+\delta,t+\tau)$. Construct a Taylor expansion as in (1.4.4) and obtain (1.4.5) in the limit as $\delta \to 0$ and $\tau \to 0$.

1.4.5. Construct a θ-scheme, as in Exercise 1.4.1, for the backward difference equation of Exercise 1.4.4.

1.5 SIMULATION OF RANDOM WALKS USING MAPLE

The relationship between random walks and partial differential equations established in this chapter is not only of intrinsic interest, but can serve as a tool for the (approximate) numerical solution of partial differential equations. The difference equations can be solved recursively, with given initial and boundary values, to determine an approximate solution of the related partial differential equations for a range of values of x at a given time t. Alternatively, each random walk can be *simulated* by the use of a *random number generator*, and for given transition probabilities p and q it is possible to determine the location of a particle that starts at some initial point after it takes a finite number of steps of fixed length. We use *Maple* to implement these methods.

Maple contains many built-in procedures or codes that carry out processes such as summation, differentiation, integration, and the solution of difference, ordinary differential, and partial differential equations. Yet Maple's procedures cannot be applied directly to deal with the foregoing partial difference equations and their simulations.

As a result, a number of new Maple procedures have been written that solve the difference equations and simulate the random walks presented above. The simulation methods fall into the general category of *Monte Carlo methods* for the solution of the initial and boundary value problems for partial differential equations.

The procedures that solve the difference equations employ Maple's ability to carry out recursions. The simulations of the random walks make use of Maple's random number generator to determine the probability that the particle undertaking a random walk takes a step to the right, to the left, or rests at each increment in time. However, the location of a particle after a single simulated random walk does not yield a meaningful approximation to the solution of the related partial differential equation. A large number of random walks must be simulated and the results must be averaged to yield a useful result. Thereby, probabilities and probability densities are assigned to the points that can be reached in the random walks. Although more general finite difference methods for the solution of initial and boundary value problems for partial differential equations and Maple procedures for their implementation are presented in Chapter 11, specific procedures related to the difference equations considered in this chapter are introduced here for the purpose of comparing their solutions with exact results and those obtained via simulation.

Unrestricted Random Walks

We begin by considering *unrestricted random walks* with constant transition probabilities p and q. The related initial value problem for the *Fokker-Planck equation* (it reduces to the heat or diffusion equation if $c = 0$) is

$$\frac{\partial v(x,t)}{\partial t} + c\,\frac{\partial v(x,t)}{\partial x} = \frac{1}{2}D\,\frac{\partial^2 v(x,t)}{\partial x^2}, \qquad v(x,0) = \delta(x). \tag{1.5.1}$$

The solution of this problem, which is referred to as the *fundamental solution of the Fokker-Planck equation*, was given in (1.1.20) as

$$v(x,t) = \frac{1}{\sqrt{2\pi Dt}} \exp\left[-\frac{(x-ct)^2}{2Dt}\right]. \tag{1.5.2}$$

A number of Maple procedures have been constructed that deal with matters related to unrestricted random walks. The procedure $FundSolFP(x, t, D, c, options)$ evaluates the fundamental solution (1.5.2) of (1.5.1) at the point x and at the time t, with prescribed diffusion and drift coefficients D and c, respectively. The fifth argument of the procedure must be either the word *explicit* or the word *float*. It determines whether the result is given in explicit or floating-point form.

As an example, the solution at $x = 1$, $t = 2$, with $D = 1$ and $c = 4$ is given in explicit form as $FundSolFP(1, 2, 1, 4, explicit) = \exp(-49/4)/2\sqrt{\pi}$ and in floating point form as $FundSolFP(1, 2, 1, 4, float) = 0.1349856694 \times 10^{-5}$.

The procedure *NumRandomWalkConst*$(p, q, [a, b], n, \delta, \tau)$ solves unrestricted random walk problems with constant transition probabilities p and q, that start at $x = 0$. The integer n specifies the number of steps taken in the random walk and δ, τ are the step length and the time increment, respectively. The solution is found in the interval $[a\delta,\ b\delta]$, and a and b should be chosen so that it contains $[-n\delta, n\delta]$ as a subinterval. If an optional seventh argument is added (it can be any name or number), information about the problem is returned but no solution data are given.

Example 1.7. Unrestricted Random Walk: A General Case. The output of the procedure *NumRandomWalkConst*$(p, q, [-2, 2], 2, \delta, \tau, r)$ displays the following general information about the random walk. Fokker-Planck Equation: $u_t + (p - q)\delta/\tau\, u_x = (1/2)(p+q)\delta^2/\tau\, u_{xx}$, Diffusion Coefficient $= (p+q)\delta^2/\tau$, Drift Coefficient $= (p-q)\delta/\tau$, x step length $= \delta$, t step length $= \tau$, Number of steps $= 2$, Stability parameter $= (1/2)(p+q)$, $t = 2\tau$. The stability parameter must be greater than zero and less than or equal to $1/2$, since p and q are probabilities.

The output of the procedure *NumRandomWalkConst*$(p, q, [-2, 2], 2, \delta, \tau)$ displays probabilities and probability densities for the random walk at the time $t = 2\tau$ in tabular form:

$$\begin{bmatrix} t = 2\tau & & \\ x & \textit{Probability} & \textit{Density} \\ -2\delta & q^2 & q^2/2\delta \\ -\delta & -2(p+q-1)q & (p+q-1)q/\delta \\ 0 & 1-2p-2q+p^2+4pq+q^2 & (1-2p-2q+p^2+4pq+q^2)/2\delta \\ \delta & -2(p+q-1)p & (p+q-1)p/\delta \\ 2\delta & p^2 & p^2/2\delta \end{bmatrix}.$$

As indicated by the display, the probability density at a given point is determined by dividing the probability by 2δ, which equals twice the x-step length for this example. It is the density values that approximate the solution of related initial value problem for the related partial differential equation. The list of probabilities can be displayed separately by entering $[seq(RWalkC(p, q, i, 2), i = -2..2)]$, whose output is $[q^2,\ 2(1-p-q)q,\ (1-p-q)^2+2pq,\ 2(1-p-q)p,\ p^2]$, while the list of probability densities can be displayed by entering $[seq(RWalkDens[i, 2], i = -2..2)]$, whose output is $[q^2/2\delta,\ q(1-p-q)/\delta,\ ((1-p-q)^2+2pq)/2\delta,\ p(1-p-q)/\delta,\ p^2/2\delta]$. If a probability at a point equals 0, its value is displayed in the output array, but the density at the corresponding point is left blank. Also, zero density values are omitted in the output of $RWalkDens$.

Next we consider a special case of the above with $p = 1/2$, $q = 1/4$, and $\delta = \tau = 0.1$. The relevant procedure is *NumRandomWalkConst*$(1/2, 1/4, [-2, 2], 2, .1, .1)$, with the output

$t = 0.2$		
x	*Probability*	*Density*
-0.2	0.0625	0.3125
-0.1	0.1250	0.6250
0.0	0.3125	1.5625
0.1	0.2500	1.2500
0.2	0.2500	1.2500

The values displayed are consistent with those given for the general case.

The procedure $SimRandomWalkConst(p, q, [a, b], n, \delta, \tau, options)$ simulates unrestricted random walk problems with constant transition probabilities p and q that start at $x = 0$. At each step, Maple's *random number generator* is used to generate a number R that lies in the unit interval $[0, 1]$. If the number R satisfies the condition $0 \leq R < p$, the particle takes a step to the right. If the number R satisfies $p < R \leq p + q$, the particle takes a step to the left. But if $p + q < R \leq 1$, the particle remains stationary. The integer n determines the number of steps in the random walk and δ, τ are the step length and the time increment, respectively. Since the particle takes n steps, it cannot get beyond the point $-n\delta$ on the left and $n\delta$ on the right. The solution is found in the interval $[a\delta, b\delta]$, which should be chosen to include the interval $[-n\delta, n\delta]$. The output can be a plot of the random walk if the seventh argument is *plot*, or an integer N that may be positive, negative, or zero if the seventh argument is *data*. For example, if $N = 5$, the particle reaches the point 5δ in the random walk, whereas if $N = -3$, the point -3δ is reached. All points reached in the random walk can be accessed by entering $N(k)$, where k represents the kth step in the walk and the corresponding point is $N(k)\delta$.

If an eighth argument m a positive integer, is added in the procedure, m random walks are simulated by the procedure and probabilities and densities are assigned to points reached in the random walks, based on the average number of times that these points are reached at the end of each random walk. The results are displayed in a table. Densities for points that cannot be reached in the random walk are not displayed in the table.

The procedure *OutputArray* takes lists of outputs for the solution of a single problem obtained by different methods and displays them in a matrix array. For example, $OutputArray(x, t = 5, [0, 1, 2], [Method(A), Method(B)], [[.1, .2], [.3, .4], [.5, .6]])$ yields the output of two different methods at the prescribed values of x and at the given time t. The outputs of each of the methods must be given as lists as shown. The procedure is used in the examples below.

Example 1.8. The Simulation of an Unrestricted Random Walk. The random walk with $p = .6$, $q = .4$, $\delta = .1$, and $\tau = .1$, with 10 steps, can be simulated by invoking $SimRandomWalkConst(.6, .4, [-10, 10], 10, .1, .1, data)$. A possible list of outputs $[-2, 0, 4, -1, 3]$ is obtained on carrying out the procedure five times. This means that at the time $t = 1$ the particle reaches the points in the list $[-0.2, 0, 0.4, -0.1, 0.3]$. The random walk corresponds to the Fokker-Planck equation

$$\frac{\partial u(x,t)}{\partial t} + \frac{1}{5}\frac{\partial u(x,t)}{\partial x} = \frac{1}{20}\frac{\partial^2 u(x,t)}{\partial x^2}. \tag{1.5.3}$$

The procedure *NumRandomWalkConst*$(.6, .4, [-10, 10], 10, .1, .1)$ solves the random walk problem at the time $t = 1$. The output gives probabilities at the points $x = -1, -0.9, ..., 0, ..., 0.9, 1$. Then $[seq(RWalkDens[i, 10], i = 10..10)]$, causes the following list of probability densities to be exhibited: $[.0011, .00786, .0531, .2123, .5574, 1.003, 1.254, 1.075, .6047, .2016, .06057]$. Only nonzero entries are displayed. (The probabilities and densities are zero because the particle cannot reach those points in the random walks. This happens, in general, only if $p + q = 1$, as is the case here.) Densities at the points that are omitted can be determined via linear interpolation, but we do not carry this out.

$SimRandomWalkConst(.6, .4, [-10, 10], 10, .1, .1, data, 5000)$ carries out a simulation of 5000 of the foregoing random walks and determines a terminal point for each of the random walks at $t = 1$. The procedure $[seq(SRWalkD(.6, .4, i, 10), i = -10..10)]$ yields the list of averaged probability densities, $[.002, 0, .008, 0, .05, 0, .212, 0, .583, 0, 1., 0, 1.233, 0, 1.081, 0, .596, 0, .21, 0, .052]$. We observe that every other entry is zero, as found above. Each application of the procedure yields a different output (in general) because the numbers used in the procedure are found using a random number generator.

Figure 1.1 displays 500 simulated random walks as determined by invoking $SimRandomWalkConst(.6, .4, [-10, 10], 10, .1, .1, plot)$ 500 times. The horizontal line segments represent the movement of the particle in one step. The Maple procedure $plots[display]$ is used to generate the plot.

The procedure [*seq(FundSolFP*$(i/10, 1, .1, .2$,float),$i = -10..10)$] generates the list of numbers: $[.00094, .002974, .0085, .02197, .05140, .1089, .2085, .3614, .5668, .8043, 1.033, 1.20, 1.261, 1.2, 1.033, .8043, .5668, .3614, .2085, .1089, .0514]$ which represent the values of the exact (fundamental) solution of the initial value problem for the (related) Fokker-Planck equation (1.5.3) at the points $x = -1, -0.9, ..., 0, ..., 0.9, 1$ and at the time $t = 1.0$. In contrast to the random walk results, there are no zero values in this list.

Making use of the procedure *OutputArray* causes the foregoing results to be displayed in an array that makes it possible for each of the entries in the list to be compared directly. Only points where the random walks have nonzero values are displayed. There is fairly good agreement between the random walk results and those found from the exact solution. Improvement can be achieved by increasing the number of subdivisions of the given interval so that the x step length is decreased.

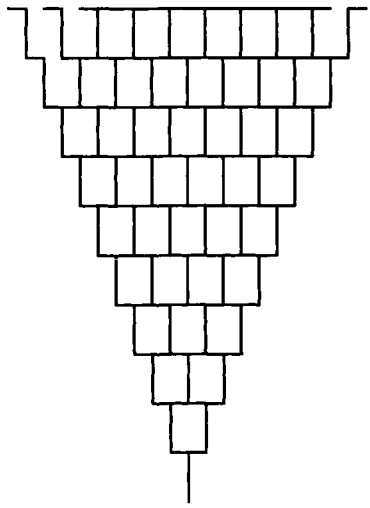

Figure 1.1 500 Simulated random walks.

Additionally, if the total number of simulations is increased, improved approximations to the probabilities are expected to be obtained:

$$
\left[\begin{array}{cccc}
t = 1.0 & & & \\
x & \mathit{FundSol} & \mathit{RandWalk} & \mathit{RWSim} \\
-1.0 & 0.0009419 & 0.001049 & 0.002000 \\
-0.8 & 0.008500 & 0.007864 & 0.008000 \\
-0.6 & 0.05142 & 0.05308 & 0.05000 \\
-0.4 & 0.2085 & 0.2123 & 0.2120 \\
-0.2 & 0.5669 & 0.5574 & 0.5830 \\
0.0 & 1.033 & 1.003 & 1.0 \\
0.2000 & 1.262 & 1.254 & 1.233 \\
0.4000 & 1.033 & 1.075 & 1.081 \\
0.6000 & 0.5669 & 0.6047 & 0.5960 \\
0.8000 & 0.2085 & 0.2016 & 0.2100 \\
1.0 & 0.05142 & 0.06047 & 0.05200
\end{array}\right] . \qquad (1.5.4)
$$

The solution curves generated by the list of points in the array and plotted using Maple's *plot* procedure are displayed in Figure 1.2.

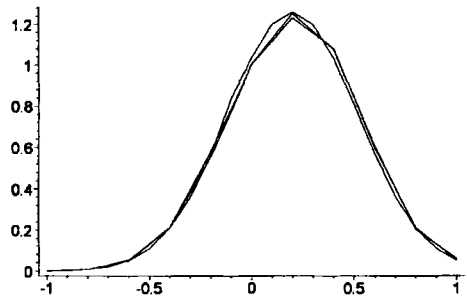

Figure 1.2 A plot of the three solution curves.

Restricted Random Walks

We consider restricted random walks in bounded intervals with constant transition probabilities p and q. (Random walks in semi-infinite intervals can also be considered by choosing the endpoint on the unbounded side to be sufficiently far from the boundary that the particle cannot reach it in the number of steps specified.) The boundaries can be absorbing or reflecting, as defined in Section 1.1. The related Fokker-Planck equation and initial condition for $v(x,t)$ are given as in (1.5.1). We require that $v(x,t) = 0$ at an absorbing boundary, and at a reflecting boundary, with $c = 0$ in the Fokker-Planck equation, we require that $v_x(x,t) = 0$. (The case with $c \neq 0$ is not considered here.)

In contrast to the situation for the unrestricted random walk considered above, there is no general, elementary, closed-form solution of the initial and boundary value problem for the related Fokker-Planck equation. The method of eigenfunction expansions, developed in Chapter 4, or the method of images of Chapter 7, must be used to solve the problem. These solutions involve infinite series, in general. We consider these results to represent exact solutions of the given problems and use truncated eigenfunction expansions to compare with results obtained via the random walk approach. (Although Maple procedures that generate solutions of partial differential equations in terms of eigenfunction expansions are presented in Chapter 4, we do not use them here. Finite expansions are constructed directly.)

Two Maple procedures that solve and simulate restricted random walk problems are presented. $NumRestRandomWalkConst(p, q, [a, b], n, lbdry, rbdry, absref, \delta, \tau)$ solves restricted random walk problems with constant transition probabilities p and q, that start at $x = 0$. The integer n specifies the number of steps taken in the random walk and δ, τ are the step length and the time increment, respectively. The arguments *lbdry* and *rbdry* are negative and positive integers l and r, respectively, that determine the location $l\delta$ and $r\delta$ of the left and right boundary, respectively. The solution is

found in the interval $[a\delta, b\delta]$ if the interval lies within that determined by the boundary points. Otherwise, the interval $[a\delta, b\delta]$ is cut off on the left or on the right by the boundary point. If the argument *absref* in the procedure is given as *absorbing*, the boundary conditions at both endpoints are absorbing, whereas if the argument *absref* is *reflecting*, the boundary conditions at both endpoints are reflecting. (We do not consider the case of mixed boundary conditions.) When an optional tenth argument is added (it can be any name or number), only information about the problem is returned, but no solution data are given.

$SimRestRandomWalkConst(p, q, [a, b], n, lbdry, rbdry, absref, \delta, \tau, option s)$ simulates restricted random walk problems with constant transition probabilities p and q, that start at $x = 0$. At each step, Maple's random number generator is used to generate a number R that lies in the unit interval $[0, 1]$. If the number R satisfies $0 \leq R < p$, the particle takes a step to the right. If the number R satisfies $p \leq R \leq p + q$, the particle takes a step to the left. But if $p + q < R \leq 1$, the particle remains stationary. The positive integer n determines the number of steps in the random walk and δ and τ are the step length and the time increment, respectively. In n steps the particle cannot go beyond the point $-n\delta$ on the left and $n\delta$ on the right. The solution is found in the interval $[a\delta, b\delta]$ if the interval lies within that determined by the boundary points. Otherwise, the interval $[a\delta, b\delta]$ is cut off on the left or on the right by the boundary point. When the argument *absref* is *absorbing*, the boundary conditions at both endpoints are absorbing, whereas if the argument *absref* is *reflecting*, the boundary conditions at both endpoints are reflecting. (We do not consider the case of mixed boundary conditions.) As was the case for the unrestricted random walk, the output can be a plot of the random walk if the 10th argument is *plot*, or an integer N that may be positive, negative, or zero if the tenth argument is *data*. All points reached in the random walk can be accessed by entering $N(k)$, where k represents the kth step in the walk, and the corresponding point is $N(k)\delta$. If an eleventh argument m is added in the procedure, where m is a positive integer, m random walks are simulated by the procedure. Then probabilities and densities are assigned to points reached in the random walks, determined by the average number of times that these points are reached at the end of each random walk.

Example 1.9. A Simulation of a Restricted Random Walk. The output of $NumRestRandomWalkConst(.5, .5, [-30, 10], 100, -30, 10, absorbing, .1, .01, p)$ gives general information about the restricted random walk that is considered and its related Fokker-Planck equation. It indicates that the random walk has $p = .5$, $q = .5$, $\delta = .1$, and $\tau = .01$. The particle takes 100 steps in the interval $-3 < x < 1$ starting at $x = 0$ with absorbing boundary conditions at $x = -3$ and $x = 0$, and the solution is found at $t = 1$. The random walk corresponds to the following problem for the diffusion equation

$$\frac{\partial u(x,t)}{\partial t} = \frac{1}{2}\frac{\partial^2 u(x,t)}{\partial x^2}, \qquad u(x,0) = \delta(x),\ u(-3,t) = 0,\ u(1,t) = 0, \quad (1.5.5)$$

in the region $-3 < x < 1,\ t > 0$.

A truncated form of the eigenfunction expansion solution (containing 50 terms) of the foregoing initial and boundary value problem is given as

$$u(x,t) \approx \frac{1}{2}\sum_{k=1}^{50} \sin\left(\frac{3\pi k}{4}\right) \exp\left(\frac{-\pi^2 k^2 t}{32}\right) \sin\left(\frac{\pi(x+3)k}{4}\right). \qquad (1.5.6)$$

Each term in the sum is a solution of the diffusion equation and vanishes at $x = -3$ and $x = 1$. The sum evaluated at $x = 0, 0.2, \ldots, 0.8, 1$, and $t = 1$ yields the values [.34495, .31209, .25735, .18350, .09551, 0.].

$NumRestRandomWalkConst(0.5, 0.5, [-30, 10], 100, -30, 10, absorbing, .1, .01)$ solves the random walk problem above at the time $t = 1$. The output of the procedure finds probabilities at the points $x = -3, -.29, \ldots, 0, \ldots, .9, 1$. On entering the procedure [*seq(RRWalkCD*$(.5, .5, i, 100), i = -30..10)$], a list of probability densities for each of the foregoing points is exhibited. We select the probability densities at the points $x = 0, 0.2, \ldots, 0.8, 1]$. This yields the list [0.34373, 0.31080, 0.25617, 0.18261, 0.09503, 0.]

The procedure $SimRestRandomWalkConst(.5, .5, [-30, 10], 100, -30, 10, absorbing, .1, .01, data, 10000)$ simulates 10,000 of the foregoing random walks and determines a terminal point for each of the random walks at $t = 1$. On using the procedure [*seq(SRRWalkD*$(.5, .5, i, 100), i = -30..10)$], a list of (averaged) probability densities is obtained, from which we exhibit the following list of probability densities at the points $x = 0, .2, \ldots, .8, 1$: [.34252, .33452, .27301, .18751, .08850, 0]. The procedure *OutputArray* displays the foregoing results in an array. We see that there is fairly good agreement between the random walk results and those found from the exact solution, but again there is room for much improvement:

$t = 1.0$			
x	*EigenExp*	*RandWalk*	*RWSim*
0.0	0.34495	0.34373	0.34252
0.2	0.31209	0.31080	0.33452
0.4	0.25735	0.25617	0.27301
0.6	0.18350	0.18261	0.18751
0.8	0.09551	0.09503	0.08850
1.0	0.0	0.0	0.0

The case of reflecting boundaries can be treated in a similar manner, but we do not consider it here.

Correlated Random Walks

The correlated (unrestricted) random walks considered in Section 1.2 were shown to correspond to the telegrapher's equation in the limit as the step length and the time between steps tends to zero. The related initial value problem for this partial differential equation was found to be [see (1.2.25)–(1.2.26)]

$$\frac{\partial^2 u(x,t)}{\partial t^2} - \gamma^2 \frac{\partial^2 u(x,t)}{\partial x^2} + 2\lambda \frac{\partial u(x,t)}{\partial t} = 0, \quad u(x,0) = \delta(x), \quad u_t(x,0) = 0. \tag{1.5.7}$$

As was the case for the unrestricted random walk and the related Fokker-Planck equation, the initial value problem (1.5.7) has an exact closed-form solution. This solution can be approximated by solving the initial value problem for the difference equations that correspond to the correlated random walk. The correlated random walk can be also be simulated and the results can also be used to generate numerical approximations to the solution of the foregoing initial value problem.

The exact solution of (1.5.7) is given as

$$u(x,t) = \frac{1}{2} e^{-\lambda t}(\delta(x-\gamma t) + \delta(x+\gamma t)) + \frac{1}{2} e^{-\lambda t}\left[\frac{\lambda}{\gamma} I_0(z) + \frac{\lambda t}{\sqrt{\gamma^2 t^2 - x^2}} I_1(z)\right] H(\gamma t - x) H(x + \gamma t), \tag{1.5.8}$$

where $I_0(z)$ and $I_1(z)$ are the modified Bessel functions of zero and first order, respectively, with $z = \lambda\sqrt{\gamma^2 t^2 - x^2}/\gamma$. (This result is derived in Chapter 7.) The function $H(z)$ represents the Heaviside function, which is defined as $H(z) = 0$, $z < 0$ and $H(z) = 1$, $z > 0$. The Maple procedure $SingSolTel(x, t, \gamma, \lambda, options)$ finds solutions of (1.5.7) at x and t, for the parameter values γ and λ. If the fifth argument is *explicit*, the solution is expressed in the form (1.5.8). If the fifth argument is *float*, the solution is expressed in floating-point form. For example, $SingSolTel(0.1, 2, 0.1, 3, float)$ gives the output 2.458859496.

Two Maple procedures that solve and simulate correlated random walk problems are given here. The Maple procedure *NumCorrRandomWalkConst*$(p, q, [a, b], n, \delta, \tau)$ solves correlated random walk problems with constant persistence and reversal probabilities p and q that start at $x = 0$. The integer n specifies the number of steps taken in the random walk and δ, τ are the step length and the time increment, respectively. The solution is found in the interval $[a\delta, b\delta]$. When an optional seventh argument is added, (it can be any name or number), only information about the problem is returned, but no solution data are given.

The procedure $SimCorrRandomWalkConst(p, q, [a, b], n, \delta, \tau, options)$ simulates correlated random walk problems with constant persistence and reversal probabilities p and q that start at $x = 0$. At each step, Maple's random number generator is used to generate a number R that lies in the unit interval $[0, 1]$. If the number R satisfies the condition $0 \leq R < p$, the particle takes a step that represents a persistence in its direction of motion. If it was moving to the right, it continues to move

in that direction in the next step. If it was moving to the left, it continues moving to the left. If the number R satisfies the condition $p < R \leq p + q$, the particle reverses its direction. If the number R satisfies the condition $p + q < R \leq 1$, the particle remains stationary. The integer n determines the number of steps in the random walk and δ, τ are the step length and the time increment, respectively. The particle cannot go beyond the point $-n\delta$ on the left and $n\delta$ on the right, in n steps. The solution is found in the interval $[a\delta, b\delta]$. As for the random walks considered above, the output can be a plot of the random walk if the seventh argument is *plot*, or an integer N that may be positive, negative, or zero if the seventh argument is *data*. All points reached in the random walk can be accessed by entering $N(k)$, where k represents the kth step in the walk, and the corresponding point is $N(k)\delta$. If an eighth argument m is added in the procedure, where m is a positive integer, m correlated random walks are simulated by the procedure and probabilities, and densities are assigned to point reached in the random walks, based on the average number of times that these points are reached at the end of each random walk.

It follows from our discussion in Section 1.2 that the probability of persistence p must be close to 1 for the random walk solution to agree closely with the solution of the telegrapher's equation. In the following example we simulate a correlated random walk whose persistence probability p equals 0.9. It will be seen, however, that p must be closer to 1 to obtain good agreement between the random walk problem and the telegrapher's equation.

Example 1.10. The Simulation of a Correlated Random Walk. The output of the procedure $NumCorrRandomWalkConst(.9, .1, [-10, 10], 10, .1, .1, p)$ displays general information about the correlated random walk that we consider and its related telegrapher's equation,

$$\frac{\partial^2 u(x,t)}{\partial t^2} - \frac{\partial^2 u(x,t)}{\partial x^2} + 2\,\frac{\partial u(x,t)}{\partial t} = 0. \tag{1.5.9}$$

The reversal rate is 1, the x and t step lengths are both 0.01, and the number of steps equals 10. The solution is found at $t = 1$. The persistence probability is $p = 0.9$ and the reversal probability is $q = 0.1$. At the initial step, the particle is equally likely to move to the right or to the left. Since $p \gg q$, the particle has a strong tendency to persist in its direction of motion once the direction is established initially. As p is fairly close to 1, we expect good agreement between the random walk result and the solution of the corresponding telegrapher's equation (1.5.9).

The Maple procedure $[e^{-1}/0.2,\ seq(SingSolTel(i/10, 1, 1, 1, float), i = -9.$ $.9),\ e^{-1}/0.2]$ determines the (exact) solution values at the points $x = -1, -.8, ..., 0,$ $..., .8, 1$ at the time $t = 1$. The values at the points $x = -1$ and $x = 1$ are determined from the expression (1.2.32).

$NumCorrRandomWalkConst(.9, .1, [-10, 10], 10, .1, .1)$ solves the difference equation for the correlated random walk and then $[seq(CWalkD(.9, .1, i, 10), i = -10..10)]$ yields the probability densities at the foregoing points. Finally, $SimCorr$ $RandomWalkConst(.9, .1, [-10, 10], 10, .1, .1, data, 5000)$ simulates 5000 correlated random walks, and $[seq(SCWalkD(.9, .1, i, 10), i = -10..10)]$ determines probability densities at the set of points that are listed above.

OutputArray displays the foregoing results in an array which shows that there is fairly good agreement between the random walk results and those found from the exact solution, but again there is room for much improvement. This could be achieved by choosing a value of p that is closer to 1. When doing so, however, it becomes necessary to adjust the other parameters in the random walk problem so that the same (related) telegrapher's equation results.

$t = 1.0$			
x	*SingSol*	*CRandWalk*	*CRandWalkSim*
-1.0	1.840	1.937	1.876
-0.8000	0.2970	0.3109	0.3110
-0.6000	0.3142	0.3326	0.3410
-0.4000	0.3266	0.3486	0.3450
-0.2000	0.3342	0.3585	0.3710
0.0	0.3368	0.3618	0.3580
0.2000	0.3342	0.3585	0.3550
0.4000	0.3266	0.3486	0.3600
0.6000	0.3142	0.3326	0.3170
0.8000	0.2970	0.3109	0.3400
1.0	1.840	1.937	1.928

It was shown in Section 1.2 that the solution of the telegrapher's equation tends to that of a related Fokker-Planck equation as t increases. In this case the related equation is the diffusion equation

$$\frac{\partial u(x,t)}{\partial t} = \frac{1}{2}\frac{\partial^2 u(x,t)}{\partial x^2}. \tag{1.5.10}$$

We compare the values of the solutions of the telegrapher's and diffusion equations at $x = 0$ and the times $t = 1, 3, 6, \ldots, 30$. The values for the telegrapher's equation are found from $[seq(SingSolTel(0, 3i, 1, 1, float), i = 1..10)]$ and are given as $[.2199, .1594, .1311, .1139, .1021, .09338, .08652, .08101, .07644, .07252]$. For the diffusion equation the values are found from $[seq(FundSolFP(0, 3i, 1, 0, float), i = 1..10)]$ as $[.2302, .1628, .1329, .1151, .103, .094, .08703, .0814, .0767, .07279]$. It is seen that the solution values agree more closely as t increases.

Time-Independent Random Walks

The random walks of Section 1.3 correspond to Laplace's and Poisson's equations in the limit as the step length tends to zero. We examine these random walks in rectangular regions in the (x, y)-plane. Maple procedures that find explicit solutions of the related difference equations and simulate the random walks are given here. The boundary value problems for the related partial differential equations do not have an explicit closed-form solution in general. We use the method of eigenfunction expansions (as presented in Chapters 4 and 7) to construct solutions of the problems. These are taken to represent exact solutions of these problems, and truncated forms of these expansions are used as approximate solutions that are compared with the results determined by solving the random walk problem.

The (steady-state) random walk problem in the rectangle $a \le x \le b$, $c \le y \le d$ is solved by the procedure $NumSteadyRandomWalk([i, j], x = a..b, n, y = c..d, m, numits, position, opt)$. The integers n and m are the number of x- and y-subintervals, respectively. The step sizes in the x- and y-directions are given as $a + (b - a)/n$ and $c + (d - c)/m$, respectively, and they must be equal. (However, we do not require that $b - a = d - c$.) The list $[i, j]$ in the first argument of the procedure specifies the coordinates of a boundary or interior point in the rectangle. The values of i and j are restricted to be positive integers. If $i = 0$, $i = n$, or $j = 0$, $j = m$, the point chosen is a boundary point, and the sixth argument of the procedure *position* must be the word *boundary*. (The numbers i and j determine the location of the boundary points in an obvious manner. The boundary points cannot be vertices of the rectangle, so that, for example, $[0, 0]$ and $[n, m]$ are excluded.) If the point $[i, j]$ corresponds to an interior point, the sixth argument of the procedure *position* must be the word *interior*. The system of difference equations is solved recursively, using iteration, and *numits* determines the number of iterations. (This method corresponds to the *Gauss-Seidel method* introduced in Chapter 11.) The eighth argument *opt* is optional. If it is omitted, a list of probabilities for all points in the subdivision of the rectangle is displayed. If there are eight arguments in the procedure, *opt* must be given as *plot*, *arrayp*, or *arrayd*. Then the output is a plot of a rectangle that exhibits probabilities at all points of the subdivision, an array that displays the probabilities, or an array that displays the probability densities, respectively.

$SimSteadyRandomWalk([i, j], x = a..b, n, y = c..d, m, [r, s], position, opt)$ simulates the random walk described above. The first six arguments of the procedure are defined as before and *position* reflects whether the point $[i, j]$ is a boundary or an interior point. The point specified by $[r, s]$ is the point in the interior of the rectangle where the particle starts its random walk. The random walk ends when the particle reaches the boundary. The eighth argument *opt* must be given as *plot* or *data*. The procedure gives the number of steps taken in the random walk, starting from the interior point $[r, s]$ until a boundary point is reached, if the last argument is *data*. Otherwise, the output is a plot that exhibits the steps taken by the particle during the random walk. When an optional ninth argument, a positive integer N, is added, the ratio of the number of times the boundary or interior point $[i, j]$ is reached in the N random walks (each of which starts at $[r, s]$) divided by N, is displayed.

Whereas $NumSteadyRandomWalk([i,j],\ x = a..b, n,\ y = c..d, m,\ numits,\ position, opt)$ determines probabilities and densities at all the grid points in the interior of the rectangle, $SimSteadyRandomWalk([i,j], x = a..b, n, y = c..d, m, [r,s], position, opt)$ only finds probabilities or densities associated with one interior point at a time. To determine probabilities and densities at all interior grid points, random walks must be simulated for each of these points.

The procedures *NumMeanFirstPassTime* and *SimMeanFirstPassTime* determine directly and via simulation, respectively, mean first passage times for steady random walks. (These problems are not strictly time-independent, but they are most closely related to steady random walks.) $NumMeanFirstPass\,Time(x = a..b, n, y = c..d, m, tstep, numits, opt)$ solves the mean first passage time problem in the rectangle $a \le x \le b,\ c \le y \le d$. The integers n and m are the number of x- and y-subintervals, respectively, in the problem. The step sizes in the x- and y-directions are given as $a + (b-a)/n$ and $c + (d-c)/m$, respectively, and they must be equal. (However, it is not required that $b - a = d - c$.) The number *tstep* represents the time step in the random walk. The system of difference equations is solved recursively, using iteration, and *numits* determines the number of iterations. The seventh argument *opt* is optional. If it is omitted, a list of mean times until absorption at the boundary, for all points in the grid of the rectangle is displayed. If there are seven arguments in the procedure, the seventh argument must be *plot* or *array*. Then the output is a plot of a rectangle that exhibits the mean first passage times or an array that displays these times at all points of the grid, respectively. If there are eight arguments in the procedure, the seventh and eighth arguments can be anything. The output of the procedure is the related Poisson equation together with some additional information.

$SimMeanFirstPassTime([i,j], x = a..b, n, y = c..d, m, tstep, opt)$ simulates the steps in a mean first passage time problem. The first argument $[i,j]$ specifies the point at which the random walk begins. The next five arguments are defined as above. The sixth argument, *opt*, must be the name *plot* or *data*. If the name is *plot*, the steps taken until the particle is absorbed are displayed and plotted. If the name is *data*, the number of steps taken until the particle is absorbed is displayed. If the name is *data* and an additional argument N is added, where N is required to be a positive integer, the mean time to absorption at the boundary for N random walks, each of which begins at $[i,j]$, is displayed.

The foregoing random walk problems are related to boundary value problems for Laplace's and Poisson's equations, as shown in Section 1.3. We consider specific random walk problems in the following example. The solutions of the related PDE problems are approximated by constructing finite sums of eigenfunction expansions. These partial sums are taken to represent exact solutions of the problems and are compared with the random walk results.

Example 1.11. A Simulation of a Steady Random Walk: Boundary Point. $NumSteadyRandomWalk([3,6], x = 0..1, 6, y = 0..1, 6, 50, boundary, arrayd)$ carries out the simulation of the random walk. The output of the procedure is given in two parts.

The main part is an array that displays the probability densities associated with the random walk. The first column in the array lists the y-values and the last row in the array lists the x-values displayed in the array (floating- point representations of fractions are used, e.g, $1/6 \approx 0.17$):

$$\begin{bmatrix} 1.0 & 0.0 & 0.0 & 0.0 & 5.999 & 0.0 & 0.0 & 0.0 \\ 0.83 & 0.0 & 0.2412 & 0.7217 & 2.089 & 0.7217 & 0.2412 & 0.0 \\ 0.67 & 0.0 & 0.2427 & 0.5567 & 0.9160 & 0.5567 & 0.2427 & 0.0 \\ 0.50 & 0.0 & 0.1731 & 0.3461 & 0.4614 & 0.3461 & 0.1731 & 0.0 \\ 0.33 & 0.0 & 0.1034 & 0.1932 & 0.2375 & 0.1932 & 0.1034 & 0.0 \\ 0.17 & 0.0 & 0.04725 & 0.08566 & 0.1022 & 0.08566 & 0.04725 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ y|x & 0.0 & 0.17 & 0.33 & 0.50 & 0.67 & 0.83 & 1.0 \end{bmatrix}.$$

Each boundary value, except that at the point (0.5, 1), is zero. The remaining entries represent probability densities at interior points.

Additionally, the first part of the output displays information about the random walk problem. The random walk takes place in the unit square $0 < x < 1$, $0 < y < 1$. The x and y step lengths both equal $1/6$ and the boundary point is given as $(1/2, 1)$. The number of iterations that will be carried out by the procedure in solving the difference equations is 50. The related partial differential equation is Laplace's equation $\nabla^2 u(x, y) = 0$, which is to be solved in the unit square with the (Dirichlet) boundary conditions $u(x, 0) = 0$, $u(x, 1) = \delta(x - 1/2)$, $u(0, y) = 0$, $u(1, y) = 0$. Again, $\delta(z)$ is the Dirac delta function.

The problem can be solved by constructing an eigenfunction expansion, of which a truncated form (with 25 terms) is given as

$$u(x, y) \approx \sum_{k=1}^{25} 4 \frac{\sin(\pi k/2) \sinh(k\pi y)[\cosh(\pi k) + \sinh(\pi k)] \sin(k\pi x)}{-1 + \cosh(2\pi k) + \sinh(2\pi k)}. \quad (1.5.11)$$

It is easily verified that each term in the sum is a solution of Laplace's equation and vanishes at $x = 0$, $y = 0$, and $x = 1$. The sum evaluated at the points with $x = [1/6, 2/6, 3/6, 4/6, 5/6]$ and $y = 2/6$ yields the list of values for the solution $u(x, y)$ as $[.1045, .1873, .2202, .1873, .1045]$.

Finally, we simulate 5000 random walks in the unit square (with the conditions given above), each of which originates at the foregoing list of points. The Maple procedure $[seq(SimSteadyRandomWalk([3, 6], x = 0..1, 6, y = 0..1, 6, [i, 2], boundary, data, 5000), i = 1..5)]$ simulates these random walks. Then $[seq(SWalkDSim([i, 2], i = 1..5)]$ finds the list of probability densities at these points as $[.09960, .2208, .2292, .1800, .1008]$.

Using the procedure *OutputArray*, we place the probability densities at $y = 0.33$ determined above in an array:

$$\begin{bmatrix} y = 0.33 & & & \\ x & EigenSol & RandWalk & RWSim \\ 0.17 & 0.1045 & 0.1034 & 0.09960 \\ 0.33 & 0.18735 & 0.1932 & 0.2208 \\ 0.50 & 0.2201 & 0.2375 & 0.2292 \\ 0.67 & 0.1873 & 0.1932 & 0.1800 \\ 0.83 & 0.1045 & 0.1034 & 0.1008 \end{bmatrix}.$$

There is fairly good agreement among the three results.

In the following example, we reconsider the random walk of Example 1.11 and replace the boundary point with an interior point. All three methods presented above will be considered and the results compared.

Example 1.12. A Simulated Steady Random Walk: Interior Point. We invoke $NumSteadyRandomWalk([3,3], x = 0..1, 6, y = 0..1, 6, 50, interior, arrayd)$. The first part of the output states that the random walk takes place in the unit square $0 < x < 1$, $0 < y < 1$ with the step length equal to $1/6$ and the interior point given as $(1/2, 1/2)$. The number of iterations carried out in solving the difference equations is 50. The remaining part of the output is an array that displays the probability densities in the form given in the preceding example:

$$\begin{bmatrix} 1.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.83 & 0.0 & 0.02885 & 0.05770 & 0.07692 & 0.05770 & 0.02885 & 0.0 \\ 0.67 & 0.0 & 0.05770 & 0.1250 & 0.1923 & 0.1250 & 0.05770 & 0.0 \\ 0.50 & 0.0 & 0.07692 & 0.1923 & 0.4422 & 0.1923 & 0.07692 & 0.0 \\ 0.33 & 0.0 & 0.05770 & 0.1250 & 0.1923 & 0.1250 & 0.05770 & 0.0 \\ 0.17 & 0.0 & 0.02885 & 0.05770 & 0.07692 & 0.05770 & 0.02885 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ y|x & 0.0 & 0.17 & 0.33 & 0.50 & 0.67 & 0.83 & 1.0 \end{bmatrix}.$$

The random walk problem is related to the following boundary value problem for Poisson's equation:

$$\frac{\partial^2 u(x,y)}{\partial x^2} + \frac{\partial^2 u(x,y)}{\partial y^2} = -\delta(x - 1/2)\,\delta(y - 1/2),\ 0 < x < 1,\ 0 < y < 1, \tag{1.5.12}$$

which is to be solved in the unit square with the (Dirichlet) boundary conditions $u(x,0) = 0$, $u(x,1) = 0$, $u(0,y) = 0$, $u(1,y) = 0$. Again, $\delta(z)$ is the Dirac delta

function, and we are effectively determining a Green's function for Dirichlet boundary conditions in the unit square. A truncated form (with 25 terms) of the eigenfunction expansion solution of the boundary value problem is given as

$$u(x,y) \approx \sum_{k=1}^{25} 2\,\frac{\sin(\pi k/2)\sinh(\pi k/2)\sinh(k\pi y)\sin(k\pi x)}{\pi k \sinh(\pi k)}$$

$$-\sum_{k=1}^{25} 2\,\frac{H(y-1/2)\sin(\pi k/2)\sinh(\pi k)\sinh[k\pi(-1+2y)/2]\sin(k\pi x)}{\pi k \sinh(\pi k)}, \tag{1.5.13}$$

where $H(z)$ is the Heaviside function. Each term in the sum is a solution of Laplace's equation and vanishes at $x = 0$, $y = 0$, $y = 1$, and $x = 1$. The sum evaluated at $x = [1/6, 2/6, 3/6, 4/6, 5/6]$ and $y = 1/6$ yields the list of values for the solution $u(x,y)$: [.03054, .05987, .07426, .05987, .03054].

Finally, we simulate 5000 random walks in the unit square (with the conditions given above) that originate at the foregoing list of points, using [*seq*(*SimSteadyRandomWalk*([3, 3], $x = 0..1, 6, y = 0..1, 6$, [i, 1],boundary,data,5000), $i = 1..5$)]. Then the procedure [*seq*(*SWalkDSim*([i, 1], $i = 1..5$)] yields the following list of probability densities at these points: [.03065, .05635, .0744, .0593, .0261].

OutputArray displays the probability densities at $y = 1/6$:

$$\begin{bmatrix} y = 1/6 & & & \\ x & EigenSol & RandWalk & RWSim \\ 1/6 & 0.03054 & 0.02885 & 0.03065 \\ 1/3 & 0.05987 & 0.05769 & 0.05635 \\ 1/2 & 0.07426 & 0.07692 & 0.07440 \\ 2/3 & 0.05987 & 0.05769 & 0.05930 \\ 5/6 & 0.03054 & 0.02885 & 0.02610 \end{bmatrix}.$$

We conclude our discussion with a consideration of a mean first passage time problem.

Example 1.13. A Mean First Passage Time Problem. We consider the boundary value problem for Poisson's equation,

$$\frac{\partial^2 u(x,y)}{\partial x^2} + \frac{\partial^2 u(x,y)}{\partial y^2} = -200, \qquad 0 < x < 0.6,\ 0 < y < 0.8, \tag{1.5.14}$$

which is to be solved subject to the (Dirichlet) boundary conditions $u(x,0) = 0$, $u(x,0.8) = 0$, $u(0,y) = 0$, $u(0.6,y) = 0$. A truncated (with 25 terms) eigenfunction expansion solution of the boundary value problem is given as

$$u(x,y) \approx \sum_{k=1}^{25} 144 \frac{\left(e^{\frac{5k\pi y}{3}} - e^{\frac{4\pi k}{3}} + e^{-\frac{\pi k(5y-4)}{3}} - 1\right)\left[(-1)^k - 1\right]\sin(\frac{5k\pi x}{3})}{k^3\pi^3\left[\cosh(\frac{4\pi k}{3}) + \sinh(\frac{4\pi k}{3}) + 1\right]}. \tag{1.5.15}$$

It is easily verified that each term in the sum vanishes at $x = 0$, $y = 0$, $y = 0.8$, and $x = 0.6$. The sum evaluated at the points with $x = [0.1, 0.2, 0.3, 0.4, 0.5]$ and $y = 0.4$ yields the values for $u(x,y)$: $[3.8716, 6.0486, 6.7481, 6.0486, 3.8716]$.

A related random walk problem is obtained by invoking the procedure $NumMean\ FirstPassTime(x = 0..0.6, 6, y = 0..0.8, 8, 1/2, 50, array, p)$. The output presents information about the random walk and states that the x step length = 0.1, the t step length = 1/2, with six x subintervals and eight y subintervals. If we drop the last argument in the procedure, we obtain the following array of mean first passage times for the random walk (the last row and first column in the array represent x- and y-values as before):

$$\begin{bmatrix}
0.80 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.70 & 0.0 & 2.043 & 3.048 & 3.354 & 3.048 & 2.043 & 0.0 \\
0.60 & 0.0 & 3.123 & 4.794 & 5.319 & 4.794 & 3.124 & 0.0 \\
0.50 & 0.0 & 3.657 & 5.686 & 6.335 & 5.686 & 3.657 & 0.0 \\
0.40 & 0.0 & 3.818 & 5.960 & 6.647 & 5.960 & 3.818 & 0.0 \\
0.30 & 0.0 & 3.657 & 5.686 & 6.335 & 5.686 & 3.657 & 0.0 \\
0.20 & 0.0 & 3.123 & 4.794 & 5.319 & 4.794 & 3.123 & 0.0 \\
0.10 & 0.0 & 2.043 & 3.048 & 3.354 & 3.048 & 2.043 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\
y|x & 0.0 & 0.10 & 0.20 & 0.30 & 0.40 & 0.50 & 0.60
\end{bmatrix}.$$

Finally, we simulate 5000 random walks in the rectangle $0 < x < 0.6$, $0 < y < 0.8$ (with the conditions given above), that originate at the foregoing list of points. The procedure $[seq(SimMeanFirstPassTime([i,4], x = 0..0.6, 6, y = 0..0.8, 8, 1/2, data, 5000), i = 1..5)]$ simulates these random walks. Then the procedure [*seq* (*FPTSim*([i,4],*i*=1..5)] finds the list of mean first passage times at these points as $[3.9275, 5.9127, 6.6331, 5.9489, 3.8412]$.

OutputArray places the mean first passage times at $y = 0.4$ determined above into an array:

$y = 0.4$			
x	*EigenSol*	*RandWalk*	*RWSim*
0.10	3.872	3.818	3.928
0.20	6.049	5.960	5.913
0.30	6.748	6.647	6.633
0.40	6.049	5.960	5.949
0.50	3.872	3.818	3.841

There is a fairly satisfactory agreement in the results.

Random Walks with Variable Transition Probabilities

For most of the random walk problems with constant transition probabilities that were treated above, related problems with variable transition probabilities were considered in Sections 1.1–1.3. Maple procedures have been created to deal with most of these problems and their simulations. Here we consider only one case, the restricted position-dependent random walk. The Maple procedures *NumRestRandomWalk* and *SimRestRandomWalk* solve and simulate such random walk problems. The related Fokker-Planck equations have variable coefficients, and in general, no closed-form solutions or full eigenfunction expansions are available for the associated initial and boundary value problems. Only random walks and their simulations are considered here.

The Maple procedure $NumRestRandomWalk(p(x), q(x), x, [a, b], n, lbdry, rbdry, absref, \delta, \tau)$ solves restricted random walk problems with variable transition probabilities $p(x)$ and $q(x)$ that start at $x = 0$. The integer n specifies the number of steps taken in the random walk, and δ, τ are the step length and the time increment, respectively. The arguments *lbdry* and *rbdry* are negative and positive integers l and r, respectively, that determine the location $l\delta$ and $r\delta$ of the left and right boundary, respectively. The solution is found in the interval $[a\delta, b\delta]$ if the interval lies within that determined by the boundary points. Otherwise, the interval $[a\delta, b\delta]$ is cut off on the left or on the right by the boundary point. If the argument *absref* in the procedure is given as *absorbing*, the boundary conditions at both endpoints are absorbing, whereas if the argument *absref* is *reflecting*, the boundary conditions at both endpoints are reflecting. When an optional eleventh argument is added (it can be any name or number), only information about the problem is returned but no solution data are given. (See Example 1.14.)

$SimRestRandomWalk(p(x), q(x), x, [a, b], n, lbdry, rbdry, absref, \delta, \tau, opts)$ simulates restricted random walk problems with variable transition probabilities $p(x)$ and $q(x)$, that start at $x = 0$. At each step, Maple's random number generator is used to generate a number R that lies in the unit interval $[0, 1]$. (Here x is the location of the particle following the previous step.) If the number R satisfies the condition $0 \leq R < p(x)$, the particle takes a step to the right. If $p(x) < R \leq p(x) + q(x)$,

the particle takes a step to the left. If $p(x) + q(x) < R \leq 1$, the particle remains stationary. The integer n determines the number of steps in the random walk and δ, τ are the step length and the time increment, respectively. The particle cannot go beyond the point $-n\delta$ on the left and $n\delta$ on the right, in n steps. The solution is found in the interval $[a\delta, b\delta]$ if it lies within that determined by the boundary points. Otherwise, $[a\delta, b\delta]$ is cut off on the left or on the right by the boundary point. If *absref* is given as *absorbing* or *reflecting*, both boundary conditions are absorbing or reflecting, respectively. The output is a plot of the random walk, or an integer N that may be positive, negative, or zero, if the eleventh argument is *plot* or *data*, respectively. All points reached in the kth step of the random walk can be accessed by entering $N(k)$. They are $N(k)\delta$. If a twelfth argument m (a positive integer) is added, m random walks are simulated.

Example 1.14. Restricted Position-Dependent Random Walk: Absorbing Boundary. $NumRestRandomWalk(6/10 - 1/10/(1+x^2), 4/10 + 1/10/(1+x^2), x, [-9, 9], 10, -10, 10, absorbing, 1/10, 1/10, p)$ has the following output; Fokker-Planck Equation: $\partial u/\partial t + (1/5)\partial/\partial x\left(\left\{1 - 1/\left(1 + x^2\right)\right\} u\right) = (1/20)\partial^2 u/\partial x^2$, with Diffusion Coefficient = $1/10$, Drift Coefficient = $1/5 - 1/[5(1 + x^2)]$, Left Boundary = -1, Right Boundary = 1, x step = $1/10$, t step = $1/10$, Steps = 10, x-intervals = 20, Stability parameter = $1/2$, $t = 1$. Omitting the last argument in the procedure and entering $[seq(RRWalkD(6/10 - 1/10/(1 + x^2), 4/10 + 1/10/(1 + x^2), x, i, 10), i = -9..9)]$ yields the following list of probability densities at the points $x = -0.8, -0.6, \ldots, 0, \ldots, 0.6, 0.8$ and at the time $t = 1$: [.0396, .2026, .5807, 1.037, 1.23, 1.013, .5898, .2368, .0595].

$SimRestRandomWalk(6/10 - 1/10/(1+x^2), 4/10+1/10/(1+x^2), x, [-10, 10], 10, -10, 10, absorbing, .1, .1, data, 10000)$ simulates 10,000 random walks under the foregoing conditions. $seq(SRRWalkD(6/10 - 1/10/(1+x^2), 4/10+1/10/(1+x^2), x, i, 10), i = -9..9)]$ gives the probability densities at $t = 1$ at the points listed above: [.043, .196, .558, 1.018, 1.214, 1.03, .623, .234, .0705]. *OutputArray* yields

$$\begin{bmatrix} t = 1.0 & & \\ x & RandWalk & RWSim \\ -0.80 & 0.0396 & 0.04300 \\ -0.60 & 0.2026 & 0.1960 \\ -0.40 & 0.5807 & 0.5580 \\ -0.20 & 1.037 & 1.018 \\ 0.0 & 1.230 & 1.214 \\ 0.20 & 1.013 & 1.030 \\ 0.40 & 0.5898 & 0.6230 \\ 0.60 & 0.2368 & 0.2340 \\ 0.80 & 0.0595 & 0.07050 \end{bmatrix}.$$

Exercises 1.5

1.5.1. Determine the fundamental solution of the Fokker-Planck equation (1.5.1), with $D = 2$ and $c = 10$, by using the procedure *FundSolFP*. (a) Use the Maple procedure *pdetest* to verify that the output satisfies the given PDE. (b) Plot the solution for $-2 \leq x \leq 2$, $0 \leq t \leq 2$ using *plot3d*. (c) Evaluate the fundamental solution at various values of x and t directly and by using the procedure *FundSolFP*. (d) Use int to integrate the fundamental solution with respect to x from $-\infty$ to ∞.

1.5.2. Use the procedure $NumRandWalkConst$ to verify all the results given in Example 1.7.

1.5.3. Refer to Example 1.8 and use the procedures $NumRandomWalkConst$ and $SimRandomWalkConst$ to obtain corresponding results. Since Maple's random number generator is used to determine the probabilities in the simulation procedure, different results are expected each time that procedure is invoked.

1.5.4. Generate a plot similar to that in Figure 1.1 using 500 simulations. Repeat the process using 1000 simulations.

1.5.5. Reproduce the table given in (1.5.4) (with possibly different values for the random walk simulation).

1.5.6. Reproduce the results of Example 1.9.

1.5.7. Reproduce the results of Example 1.10.

1.5.8. Reproduce the results of Example 1.11.

1.5.9. Reproduce the results of Example 1.12.

1.5.10. Reproduce the results of Example 1.13.

1.5.11. Reproduce the results of Example 1.14.

CHAPTER 2

FIRST ORDER PARTIAL DIFFERENTIAL EQUATIONS

2.1 INTRODUCTION

This chapter initiates our study of analytical methods for solving partial differential equations. We begin with scalar first order equations, that is, equations in which the highest derivative of the dependent variable is of first order. Our discussion is restricted to problems involving two independent variables and deals almost exclusively with the *method of characteristics* for solving these problems. The presentation is thereby simplified, but as shown in the exercises, this method is easily extended to handle equations with more independent variables. Other analytical solution methods for first order equations exist, but apart from the method of the *complete integral*, they are not discussed here. Numerical solution methods for the solution of first order PDEs based on random walk analogies were presented in Chapter 1 and finite difference methods for their solution are presented in Chapter 11.

First order equations can occur directly as models for physical processes or can arise as approximations to higher-order equations or systems of equations. Thus the *unidirectional linear wave equation* (1.2.33), studied in Example 2.2, approximates the diffusion equation of Section 1.1 when the diffusion coefficient is negligibly small. The *quasilinear wave equation* (2.3.9), or more generally (2.3.48), serves as an

Partial Differential Equations of Appplied Mathematics, Third Edition. By Erich Zauderer

approximation to Euler's equations of fluid dynamics (see Section 8.5) or represents a simple model for traffic flow [see also (1.4.22) and (1.4.23)]. The nonlinear *eiconal equation* (2.4.23) plays an important role in the theory of geometrical optics and represents an approximation to wave optics as shown in Section 10.1. In fact, linear, quasilinear, and nonlinear first order equations play significant roles in the approximation methods for higher order equations which are discussed in Chapters 9 and 10.

The formulation of mathematical models can lead naturally to systems of first order equations, as in the case of the correlated random walk of Section 1.2 [see (1.2.19)–(1.2.20)] or as in Euler's hydrodynamic equations of Section 8.5. Alternatively, as shown in Example 2.1, systems can be constructed by reduction of higher-order equations. Conversely, linear systems of equations with constant coefficients can always be reduced to single higher-order equations, as was shown in Section 1.2 for the special system (1.2.19)–(1.2.20) and as can be demonstrated in general. The methods of this section cannot be used directly for the solution of first order systems except in special cases, such as in the derivation of d'Alembert's solution of the wave equation in Example 2.4. (That is, if the system contains one or more first order equations that are uncoupled from the other equations, these can be solved separately and their solutions can then be introduced into the remaining equations. As in the case of Example 2.4, it may then be possible to solve the full system using the methods of this chapter.) In Chapter 3 we discuss first order systems of partial differential equations and their simplification. Although we do not present any general analytic solution methods in the following chapter and in this book for arbitrary first order systems, specific problems that involve linear and nonlinear systems are studied and various methods of solution are introduced. Additionally, numerical finite difference methods for the solution of first order PDEs and systems of PDEs are presented in Chapter 11.

Example 2.1. The Reduction of Higher-Order Equations to Systems. The *wave equation* (1.2.35) can be written in the factored form

$$v_{tt}(x,t) - \gamma^2 v_{xx}(x,t) = (\partial_t^2 - \gamma^2 \partial_x^2)v(x,t) = (\partial_t + \gamma \partial_x)(\partial_t - \gamma \partial_x)v(x,t) = 0, \tag{2.1.1}$$

with $\partial_t \equiv \partial/\partial t$ and $\partial_x \equiv \partial/\partial x$. We set $(\partial_t - \gamma \partial_x)v(x,t) = u(x,t)$ and find from (2.1.1) that $(\partial_t + \gamma \partial_x)u(x,t) = 0$, so that we obtain the system

$$\partial v(x,t)/\partial t - \gamma \partial v(x,t)/\partial x = u(x,t), \quad \partial u(x,t)/\partial t + \gamma \partial u(x,t)/\partial x = 0. \tag{2.1.2}$$

Although this system is coupled, the second equation can be solved independent of the first. Its solution $u(x,t)$ can then be introduced into the first equation and $v(x,t)$ can be determined. This is done in Example 2.4.

The transformation $w(x,t) = e^{\lambda t}v(x,t)$ converts the *damped wave equation* (1.2.25) into

$$\frac{\partial^2 w(x,t)}{\partial t^2} - \gamma^2 \frac{\partial^2 w(x,t)}{\partial x^2} - \lambda^2 w(x,t) = 0. \tag{2.1.3}$$

The factorization method can then be applied to yield

$$\frac{\partial w(x,t)}{\partial t} - \gamma \frac{\partial w(x,t)}{\partial x} = u(x,t), \quad \frac{\partial u(x,t)}{\partial t} + \gamma \frac{\partial u(x,t)}{\partial x} = \lambda^2 w(x,t). \quad (2.1.4)$$

Now both equations must be solved simultaneously, and the methods developed in this chapter do not lead to a straightforward solution of (2.1.4). [A direct factorization of equation (1.2.25) is presented in the exercises.]

The factorization technique applied to *Laplace's equation* (1.3.4) yields

$$\frac{\partial^2 v(x,y)}{\partial x^2} + \frac{\partial^2 v(x,y)}{\partial y^2} = (\partial_x + i\partial_y)(\partial_x - i\partial_y)v(x,y) = 0. \quad (2.1.5)$$

where $i = \sqrt{-1}$, $\partial_x \equiv \partial/\partial x$, and $\partial_y \equiv \partial/\partial y$. Proceeding as for the wave equation leads to a first order coupled system with real and imaginary coefficients. As we generally require real solutions of Laplace's equations, this procedure is unsatisfactory. A system of equations more customarily associated with Laplace's equation is obtained by putting $\partial v(x,y)/\partial x = u(x,y)$ and $\partial v(x,y)/\partial y = w(x,y)$. Assuming that the mixed partial derivatives are equal, we easily obtain, on using (2.1.5),

$$\partial w(x,y)/\partial x = \partial u(x,y)/\partial y, \quad \partial w(x,y)/\partial y = -\partial u(x,y)/\partial x. \quad (2.1.6)$$

These are the *Cauchy-Riemann equations*, which are of fundamental importance in the theory of complex variables. They must be solved simultaneously.

For the *diffusion equation* (1.1.15), no simple factorization is possible. However, if we put $u(x,t) = \partial v(x,t)/\partial x$, we obtain the system

$$\partial v(x,t)/\partial x = u(x,t), \quad \partial v(x,t)/\partial t - (D/2)\,\partial u(x,t)/\partial x = -cu(x,t). \quad (2.1.7)$$

Again, this system must be solved simultaneously.

This chapter emphasizes methods that yield explicit solutions of PDEs. The reduction of higher-order PDEs to (real) first order systems is of importance in their numerical solution via the finite difference methods of Chapter 11.

Exercises 2.1

2.1.1. Show that if we set $u = v_t$ and $w = v_x$ in the wave equation (2.1.1), we obtain the system $\{u_x = w_t,\ u_t = \gamma^2 w_x\}$. Note that this system is coupled, and in contrast to (2.1.2), neither of the equations can be solved independent of the other. Thus the manner in which an equation of higher order is represented as a system can play a significant role.

2.1.2. Express the telegrapher's equation (1.2.25) in the form $(\partial_t + \gamma\partial_x)(\partial_t - \gamma\partial_x + 2\lambda)v - 2\lambda\gamma v_x = 0$, and obtain the coupled system $\{v_t - \gamma v_x = u - 2\lambda v,\ u_t + \gamma u_x - 2\lambda\gamma v_x = 0\}$.

2.1.3. Consider the PDE $v_{tt} - \gamma^2 v_{xx} + a v_x + b v_t + c v = 0$ with constant coefficients. Use factorization to obtain $\mathbf{u}_t(x,t) + A\mathbf{u}_x(x,t) + C\mathbf{u}(x,t) = \mathbf{0}$, where $\mathbf{u}(x,t) = \begin{bmatrix} v(x,t) \\ u(x,t) \end{bmatrix}$, $A = \begin{bmatrix} -\gamma & 0 \\ a - \gamma b & \gamma \end{bmatrix}$, $C = \begin{bmatrix} b & -1 \\ c & 0 \end{bmatrix}$. Show that if $a = \gamma b$ and $c = 0$, the differential equation for u can be solved independent of that for v.

2.1.4. Show that the functions $u(x,y)$ and $w(x,y)$ in the Cauchy-Riemann equations (2.1.6) both satisfy Laplace's equation.

2.2 LINEAR FIRST ORDER PARTIAL DIFFERENTIAL EQUATIONS

Method of Characteristics

The most general first order linear partial differential equation has the form

$$a(x,t)\frac{\partial v(x,t)}{\partial x} + b(x,t)\frac{\partial v(x,t)}{\partial t} = c(x,t)v(x,t) + d(x,t), \tag{2.2.1}$$

where a, b, c, and d are given functions. Unless stated otherwise, these functions are assumed to be continuously differentiable. At each point (x,t) where the vector $[a(x,t), b(x,t)]$ is defined and nonzero, the left side of (2.2.1) (essentially) represents a directional derivative of $v(x,t)$ in the direction of $[a(x,t), b(x,t)]$. The equations

$$\frac{dx}{ds} = a(x,t), \qquad \frac{dt}{ds} = b(x,t) \tag{2.2.2}$$

determine a family of curves $x = x(s)$, $t = t(s)$. (A particular curve in the family is given by specifying a point through which the curve must pass.) The tangent vector to the curve $x = x(s)$, $t = t(s)$ is $[x'(s), t'(s)]$. The system (2.2.2) states that the tangent vector coincides with the direction of the vector $[a(x,t), b(x,t)]$ at each point where $[a(x,t), b(x,t)]$ is defined and nonzero. Therefore, the derivative of $v(x,t)$ along these curves becomes

$$\frac{dv}{ds} = \frac{dv[x(s),t(s)]}{ds} = \frac{\partial v}{\partial x}\frac{dx}{ds} + \frac{\partial v}{\partial t}\frac{dt}{ds} = a\frac{\partial v}{\partial x} + b\frac{\partial v}{\partial t} = cv + d \tag{2.2.3}$$

on using the chain rule and (2.2.1) and (2.2.2).

The family of curves $x = x(s)$, $t = t(s)$, $v = v(s)$ determined by solving the system of ODEs (2.2.2)–(2.2.3), are called the *characteristic curves* of the PDE (2.2.1). Since the equations (2.2.2) can be solved independently of (2.2.3), the curves in the (x,t)-plane determined from (2.2.1) are occasionally also referred to as characteristic curves or *characteristic base curves*. The approach we develop to solve (2.2.1) by using the solutions of (2.2.2)–(2.2.3) is called the *method of characteristics*. As shown, it is based on the geometric interpretation of the partial differential equation (2.2.1).

The existence and uniqueness theory for ordinary differential equations, assuming certain smoothness conditions on the functions a, b, c, and d, guarantees that exactly

one solution curve $(x(s), t(s), v(s))$ of (2.2.2)–(2.2.3) (i.e., a characteristic curve) passes through a given point (x_0, t_0, v_0) in (x, t, v)-space. As a rule, we are not interested in determining a general solution of the partial differential equation (2.2.1) but rather a specific solution $v = v(x, t)$ [i.e., a surface in (x, t, v)-space] that passes through or contains a given curve C. This problem is known as the *initial value problem* for (2.2.1).

The initial value problem for (2.2.1) is solved by the *method of characteristics* as follows. We represent the initial curve C parametrically as

$$x = x(\tau), \qquad t = t(\tau), \qquad v = v(\tau) \tag{2.2.4}$$

for a given range of values of the parameter τ. The curve may be of finite or infinite extent and must have a continuous tangent vector at each point. Every value of τ fixes a point on C through which a unique characteristic curve passes. The family of characteristic curves determined by the points of C may be parameterized as

$$x = x(s, \tau), \qquad t = t(s, \tau), \qquad v = v(s, \tau), \tag{2.2.5}$$

where $s = 0$ corresponds to the initial curve C. That is, we have $x(0, \tau) = x(\tau)$, $t(0, \tau) = t(\tau)$, and $v(0, \tau) = v(\tau)$.

The equations (2.2.5), in general, yield a parametric representation of a surface in (x, t, v)-space that contains the initial curve C. Assuming that the equations $x = x(s, \tau)$, $t = t(s, \tau)$ can be inverted to give s and τ as (smooth) functions of x and t [which is the case if the Jacobian determinant $\Delta(s, \tau) = x_s t_\tau - t_s x_\tau \neq 0$ at C], these functions can be introduced into the equation $v = v(s, \tau)$. The resulting function $v = V(x, t)$ satisfies (2.2.1) in a neighborhood of the curve C in view of (2.2.3), the initial condition (2.2.4) {i.e., $V(x(\tau), t(\tau)) = v(\tau)$}, and is the unique solution of the given initial value problem. It is referred to as the *integral surface* for the problem. The smoothness requirements placed on the functions a, b, c, and d in equation (2.2.3) imply that $V(x, t)$ is continuously differentiable near the curve C. Figure 2.1 shows how the integral surface is constructed in terms of the initial and characteristic curves.

If the foregoing method does not result in an integral surface, the initial value problem may not have a solution at all, or it may have infinitely many solutions. The latter situation arises if the initial curve C is itself a characteristic curve, in what is known as a *characteristic initial value problem.* Even if an integral surface can be constructed, it may be discontinuous or may fail to be differentiable along some curve.

Examples

The following examples illustrate the use of the method of characteristics in solving initial value problems for linear first order equations and show how problems with nonsmooth or nonunique solutions can arise and how to deal with them. (The role of characteristics in relation to the nonexistence or nonuniqueness of solutions of initial value problems is reexamined from a different point of view in Section 3.2.)

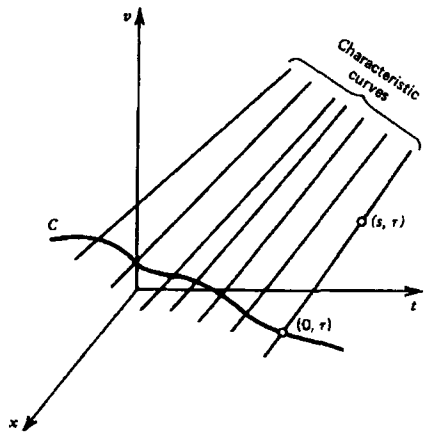

Figure 2.1 Construction of the integral surface.

We observe that although the method of characteristics reduces the problem of solving first order PDEs to that of solving a system of ODEs, the exact solution of this (generally) nonlinear system is most often not an elementary task. Numerical methods for the solution of the initial value problem for this system may have to be used.

Example 2.2. Unidirectional Linear Wave Motion. In Chapter 1

$$\partial v(x,t)/\partial t + c\,\partial v(x,t)/\partial x = 0, \tag{2.2.6}$$

with a constant coefficient c, was said to represent *unidirectional wave motion* [see (1.2.33)]. (This equation is also known as the *advection equation.*) We examine the initial value problem for (2.2.6) with the initial condition at $t = 0$,

$$v(x,0) = F(x), \tag{2.2.7}$$

where $F(x)$ is a given function.

Using the *method of characteristics*, we parameterize the initial curve C as

$$x = \tau, \qquad t = 0, \qquad v = F(\tau). \tag{2.2.8}$$

The characteristic equations [i.e., (2.2.2)–(2.2.3)] become

$$\frac{dx}{ds} = c, \qquad \frac{dt}{ds} = 1, \qquad \frac{dv}{ds} = 0. \tag{2.2.9}$$

Solving (2.2.9) subject to (2.2.8), with $s = 0$ as the initial curve, gives

$$x(s,\tau) = cs + \tau, \qquad t(s,\tau) = s, \qquad v(s,\tau) = F(\tau). \tag{2.2.10}$$

Using the first two equations to solve for s and τ as functions of x and t yields

$$s = t, \quad \tau = x - ct. \tag{2.2.11}$$

Substitution of this result in the equation for v in (2.2.10) gives

$$v(s, \tau) = F(\tau(x, t)) = F(x - ct). \tag{2.2.12}$$

Clearly, if $F(x)$ is differentiable, the solution $v(x, t) = F(x - ct)$ satisfies (2.2.6) as well as the initial condition (2.2.7). [The formal solution $v(x, t) = \delta(x - ct)$ of (1.2.33) given in Chapter 1 follows from (2.2.12).]

Let the initial function $v(x, 0) = F(x)$ represent a *wave form*. Then the solution $v(x, t) = F(x - ct)$ shows that a point x for which $x - ct$ = constant as t increases, will always occupy the same position on the wave form. If $c > 0$, that point x moves to the right with the speed $dx/dt = c$. As x is a typical point, we see that the entire initial wave form $F(x)$ moves to the right with speed c without change of shape. If $c < 0$, the direction of motion is reversed. Therefore, (2.2.6) characterizes *unidirectional wave motion* with velocity c.

The *characteristic base curves* for (2.2.6) are $x - ct = \tau$ in view of (2.2.11). For each value of τ, (2.2.10) shows that v is constant on these curves. This implies that the initial data are transmitted along these curves, in the sense that whatever the value of v is at $t = 0$ at some point x, it is retained at all points x that lie on the characteristic curve through that initial point. As a result, the initial data are transmitted at the characteristic velocity $dx/dt = c$ (to the right or to the left) as t increases from zero.

The expression $v = F(x - ct)$ obtained in (2.2.12) represents a *general solution* of (2.2.6) since $F(z)$ is an arbitrary function. It can be used, as an alternative to the method of characteristics, to solve initial value problems for (2.2.6). For example, if we specify that $v = f(t)$ on the curve $x = h(t)$, we obtain $v(h(t), t) = f(t) = F(h(t) - ct)$. With $z = h(t) - ct$, there exists a differentiable inverse function $t = g(z)$ if $z'(t) = h'(t) - c \neq 0$, that is, if the initial base curve $x = h(t)$ is not tangent to a characteristic base curve at any point. Then the unique solution of the problem is $v(x, t) = f(g(x - ct))$.

If $h'(t) = c$ along the entire curve, we find that $x - ct$ is a constant along that curve, so that the initial base curve is a *characteristic*. On denoting the constant by a (i.e., $x - ct = a$), we have $v = f(t) = F(a)$ on the initial curve. Then the problem has no solution unless $f(t) = A$ = constant. If so, this is a *characteristic initial value problem* for (2.2.6), since it follows from (2.2.9) that the initial curve in (x, t, v)-space is characteristic. The solution to this problem is not unique. Any differentiable function $F(z)$ with $F(a) = A$ yields a solution $v = F(x - ct)$ of the problem.

If $h'(t) = c$ at a finite number of points, a unique continuous inverse function $t = g(z)$, as given above, exists. However, $g(z)$ will not be differentiable at these points, so that $v = f(g(x - ct))$ is a continuous, nondifferentiable solution along the characteristic base curves through these points.

Example 2.3. Unidirectional Linear Wave Motion: Specific Initial Data. Consider the unidirectional linear wave equation (2.2.6) with the initial value $v = f(t)$ on the parabola $x = h(t) = t^2$. Since $h'(c/2) = c$, we see that the parabola is tangent to the characteristic base curve $x - ct = -c^2/4$ at the point $(x,t) = (c^2/4, c/2)$. Using the general solution $v = F(x - ct)$, we find that $F(t^2 - ct) = f(t)$, on applying the initial condition. With $z = t^2 - ct$, the inverse function $t = g(z)$ is defined (piecewise) as $t = g_-(z) = c/2 - (z + c^2/4)^{1/2}$, $t \leq c/2$, and $t = g_+(z) = c/2 + (z + c^2/4)^{1/2}$, $t > c/2$. It is continuous for $z \geq -c^2/4$, but it is not differentiable at $z = -c^2/4$.

Thus, on the characteristic base curves that intersect $x = t^2$ with $t \leq c/2$, we have the solution $v = v_-(x,t) = f(g_-(x - ct))$, while those curves that intersect the parabola with $t > c/2$ yield $v = v_+(x,t) = f(g_+(x - ct))$. However, in the region $x - ct \geq -c^2/4$, where the solution to the initial value problem is defined, each characteristic intersects the parabola twice except at the point $(x,t) = (c^2/4, c/2)$, where the characteristic is tangent to the parabola. Since $v = f(t)$ is assigned arbitrarily and $v_-(x,t)$ and $v_+(x,t)$ are constant on each characteristic, we see that the solution is double valued for $x - ct > -c^2/4$. Since any neighborhood of the point of tangency on the parabola contains points where the solution is double valued, the domain of definition of the integral surface cannot be restricted to yield a nonsingular solution. Therefore, the initial value problem has no solution.

On reformulating the problem and specifying that $v = f(t)$ on $x = t^2$ for $t \leq c/2$, the unique solution is $v = f(g_-(x - ct))$. It is continuous for $x - ct \geq -c^2/4$, but it is not differentiable along the line $x - ct = -c^2/4$. As such, it must be interpreted as a solution in a *generalized sense*, in a manner to be defined below. However, if we require that $t \leq t_0 < c/2$, the solution has the same functional form as before but is now differentiable wherever it is defined. Similarly, we can consider an initial value problem on $x = t^2$ with $t \geq c/2$ and obtain a unique solution.

With regard to the initial value problem for the general first order equation (2.2.1), if the characteristic base curves, determined from (2.2.2), intersect the initial base curve more than once, the problem generally has no solution. The solution $v(x,t)$ is completely determined along a characteristic base curve by its value at a point on the initial base curve. Thus, if the projection of this characteristic on the (x,t)-plane (i.e., the characteristic base curve) intersects the initial base curve more than once, the solution will in general be multivalued. Each intersection point carries an initial value for v and thereby determines a characteristic curve that differs from those characteristics determined at other intersection points, except in the unlikely case that the data at all these points are compatible.

Of particular interest is the *initial and boundary value problem* for (2.2.1), with the initial condition $v(x, t_0) = f(x)$ for $x > x_0$ and the boundary condition $v(x_0, t) = g(t)$ for $t > t_0$. The problem is to be solved in the quarter-plane $x > x_0$, $t > t_0$. If each characteristic base curve intersects both the initial and the boundary line of the quarter-plane, the problem has no solution in general. However, if each characteristic

base curve intersects the initial and the boundary line of the quarter-plane only once, a solution can be found in general. [A similar result is obtained if the initial condition is replaced by $v(x, t_0) = f(x)$ for $x < x_0$, so that the quarter-plane is $x < x_0$, $t > t_0$.] Thus, for (2.2.6) with $c < 0$, the initial and boundary value problem in the quarter-plane $x > x_0$, $t > t_0$ has no solution in general.

Example 2.4. d'Alembert's Solution of the Wave Equation. We now derive the solution of the initial value problem for the wave equation (2.1.1), which was first obtained by *d'Alembert*. The wave equation was encountered in Section 1.2 and it describes many physical phenomena, as we shall see. It is often introduced as an equation describing the transverse displacement of a tightly stretched string, and is discussed from this point of view in Example 4.9. The constant γ^2 then equals the tension T divided by the density ρ of the string, and $v(x, t)$ is the vertical displacement of a point x on the string at the time t. Appropriate initial conditions for the wave equation are

$$v(x, 0) = f(x), \qquad \partial v(x, 0)/\partial t = g(x). \tag{2.2.13}$$

For the vibrating string, these represent its initial displacement and velocity. It is then assumed that the string is sufficiently long that disturbances arising at the ends of the string do not affect the vibration generated by the initial data, within the time span in which the motion is observed.

We solve the initial value problem (2.1.1) and (2.2.13) by using the system (2.1.2). The initial value of $v(x, t)$ in the system is given as in (2.2.13), and the initial condition for $u(x, t)$ is obtained from the first equation of (2.1.2) as $u(x, 0) = g(x) - \gamma f'(x)$. Using (2.2.12) we immediately obtain $u(x, t) = g(x - \gamma t) - \gamma f'(x - \gamma t)$, since $F(x) = g(x) - \gamma f'(x)$. Substituting into (2.1.2) gives

$$\frac{\partial v(x, t)}{\partial t} - \gamma \frac{\partial v(x, t)}{\partial x} = g(x - \gamma t) - \gamma f'(x - \gamma t), \tag{2.2.14}$$

with the initial condition $v(x, 0) = f(x)$.

From (2.2.2)–(2.2.3) we obtain the characteristic equations for (2.2.14) as

$$\frac{dx}{ds} = -\gamma, \qquad \frac{dt}{ds} = 1, \qquad \frac{dv}{ds} = g - \gamma f', \tag{2.2.15}$$

with the initial curve $v(x, 0) = f(x)$ parameterized as $x = \tau$, $t = 0$, $v = f(\tau)$. The solutions of the problem above yield the system of curves

$$x = -\gamma s + \tau, \qquad t = s, \qquad v = \int_0^s (g - \gamma f')\, d\sigma + f(\tau). \tag{2.2.16}$$

Now both g and f' are functions of $x - \gamma t$, and the first two equations in (2.2.16) imply that $x - \gamma t = -2\gamma s + \tau$. Thus $v(s, \tau)$ can be expressed as $v(s, \tau) = \int_0^s [g(-2\gamma\sigma + \tau) - \gamma f'(-2\gamma\sigma + \tau)]\, d\sigma + f(\tau)$. The change of variables $\lambda = -2\gamma\sigma + \tau$ yields $v(s, \tau) = -(1/2\gamma) \int_\tau^{-2\gamma s + \tau} g(\lambda)\, d\lambda + (1/2) f(-2\gamma s + \tau) + (1/2) f(\tau)$. Again the

first two equations in (2.2.16) imply that $\tau = x + \gamma t$ and $-2\gamma s + \tau = x - \gamma t$, so that in terms of x and t, the solution is

$$v = V(x,t) = \frac{1}{2}[f(x+\gamma t) + f(x-\gamma t)] + \frac{1}{2\gamma}\int_{x-\gamma t}^{x+\gamma t} g(\lambda)\, d\lambda. \qquad (2.2.17)$$

This is *d'Alembert's solution* of the initial value problem for the wave equation. It can be verified directly that it satisfies (2.1.1) and (2.2.13).

Let $G(z) = \int^z g(\lambda)\, d\lambda$, that is, G is the antiderivative of g. Then (2.2.17) can be expressed as

$$v(x,t) = \left[\frac{1}{2}f(x+\gamma t) + \frac{1}{2\gamma}G(x+\gamma t)\right] + \left[\frac{1}{2}f(x-\gamma t) - \frac{1}{2\gamma}G(x-\gamma t)\right]. \qquad (2.2.18)$$

Since $\gamma > 0$, the first and second bracketed terms represent waves traveling to the left and right, respectively, with speed γ. Each wave travels without change of shape. However, the presence of these wave forms is not always apparent in the solution, because of the interference caused by the interaction of these traveling waves when they are superposed or summed as in (2.2.18).

Generalized Solutions

The family of characteristics (2.2.5) yields a smooth solution $v = V(x,t)$ of the initial value problem (2.2.1) and (2.2.4) only if the data (2.2.4) and the functions a, b, c, and d in (2.2.1) are smooth, and the Jacobian determinant $\Delta(s,\tau)$ is not zero. If the initial value $v = v(\tau)$ in (2.2.4) is singular at $\tau = \tau_0$, this singularity is transmitted along the characteristic (2.2.5) that passes through that point on the initial curve (2.2.4). Then the integral surface is singular along that characteristic, and depending on the singularity of $v(\tau)$, it may not have first derivatives or it may be discontinuous there.

For example, if the first derivative of $F(x)$ in (2.2.7) has a jump discontinuity at $x = 0$, the solution $v = F(x - ct)$ of (2.2.6) has discontinuous x- and t-derivatives along the characteristic (base) curve $x = ct$. If $F(x)$ has a jump discontinuity at $x = 0$, then $v = F(x - ct)$ has a jump discontinuity along $x = ct$. In each case, $v = F(x-ct)$ is not a solution of (2.2.6) in the classical sense, as it is not differentiable along $x = ct$. We must treat it as a *generalized solution* and now show how this is to be done in the general case.

Let $v(x,t)$ be a continuously differentiable solution of (2.2.1) except along the arbitrary curve $x = x(s)$, $t = t(s)$, where it is continuous but its first derivatives have (finite) jump discontinuities. We evaluate (2.2.1) at two points (x,t) on either side of the curve and consider the difference of these two equations. In the limit as these points approach a common point on the curve, we obtain

$$a[v_x] + b[v_t] = 0, \qquad (2.2.19)$$

where the brackets represent the jumps in the first derivatives across the curve. All other terms in (2.2.1) are assumed to be continuous across the curve. A continuously

differentiable solution of (2.2.1) everywhere except along a curve across which the jumps in the first derivatives satisfy (2.2.19) is said to be a *generalized solution* of (2.2.1).

In fact, the curve $x = x(s)$, $t = t(s)$ across which the jumps in the first derivatives satisfy (2.2.19) cannot be arbitrary but must be a *characteristic base curve*. Since $[v] = 0$ by assumption, if we differentiate this equation along the curve (using the chain rule), we find that $d[v]/ds$ is given as

$$x'(s)[v_x] + t'(s)[v_t] = 0. \tag{2.2.20}$$

Now (2.2.19)–(2.2.20) yields a simultaneous homogeneous linear system for the jumps $[v_x]$ and $[v_t]$. As we assume that they are not both zero, we can effectively conclude that $x'(s) = a$ and $t'(s) = b$. In view of (2.2.2), the curve must be a characteristic, and (2.2.19) states that $d[v]/ds = 0$ along the characteristic. It is shown in the exercises that the jumps $[v_x]$ and $[v_t]$ satisfy ordinary differential equations of first order along the characteristics. From these equations we conclude that unless the derivatives undergo jumps in the data for the problem, the solution cannot have any jumps in the derivatives.

If $v(x,t)$ is a continuously differentiable solution of (2.2.1) except along a curve $x = x(s)$, $t = t(s)$ where v has a jump discontinuity, we must generalize the concept of a solution of (2.2.1) even further. To do so we write (2.2.1) in *conservation form*,

$$(av)_x + (bv)_t = (c + a_x + b_t)v + d, \tag{2.2.21}$$

and integrate (2.2.21) over an arbitrary closed and bounded region R in the (x,t)-plane with a piecewise smooth boundary curve S. On applying Green's theorem in the plane, we obtain

$$\int_S av\,dt - bv\,dx = \iint_R [(c + a_x + b_t)v + d]\,dx\,dt, \tag{2.2.22}$$

with integration over S taken in the positive direction. This *integral relation* is equivalent to the differential equation (2.2.1) in any region where $v(x,t)$ is continuously differentiable, as follows from Green's theorem and the arbitrariness of the region R (see Exercise 8.1.9). However, the integral relation remains valid even if $v(x,t)$ has a jump discontinuity across some curve.

If we apply the integral relation (2.2.22) to a region R that surrounds an arbitrary portion of the curve $x = x(s)$, $t = t(s)$ across which $v(x,t)$ has a jump discontinuity (see Figure 6.3), and let the boundary curve S collapse onto the discontinuity curve, we obtain

$$(a\,t'(s) - b\,x'(s))[v] = 0, \tag{2.2.23}$$

since the limiting line integrals are taken in opposite directions and $v(x,t)$ has different limits on both sides of the curve. (Again, $[v]$ is the jump in v across the curve.) Although the limit implies only that the integral of (2.2.23) over a portion of the discontinuity curve vanishes, the arbitrariness of that portion implies that the integrand must itself be zero (see Exercise 8.1.9).

Since $[v] \neq 0$ in (2.2.23) by assumption, we must have $at'(s) - bx'(s) = 0$. But the vector $[t'(s), -x'(s)]$ is normal to the curve $x = x(s)$, $t = t(s)$, so we again conclude that the discontinuity curve must be a characteristic base curve along which $x'(s) = a$, $t'(s) = b$. Thus, if $v(x,t)$ has (together with its first derivatives) a jump discontinuity across a characteristic base curve (2.2.2) but is a continuously differentiable solution of (2.2.1) elsewhere, it is a *generalized solution* of (2.2.1).

As examples, we consider

$$v_1(x,t) = \begin{cases} x - ct, & x < ct, \\ 1, & x > ct, \end{cases} \qquad v_2(x,t) = \begin{cases} x - ct, & x < 0, \\ 1, & x > 0. \end{cases} \tag{2.2.24}$$

$v_1(x,t)$ is a generalized solution of (2.2.6) but $v_2(x,t)$ is not, since $x = ct$ is a characteristic base curve of (2.2.6) whereas $x = 0$ is not.

If we apply the discussion that precedes equation (2.2.19) to the present case, where both v and its first derivatives have jump discontinuities across the characteristic base curve (2.2.2), we obtain

$$a[v_x] + b[v_t] = c[v], \tag{2.2.25}$$

since v also has a jump discontinuity. As we have shown, (2.2.25) is equivalent to the ODE $d[v]/ds = c[v]$ along the discontinuity curve. It not only determines the variation of the jump along that curve, but also shows that there can be no discontinuity in a generalized solution of (2.2.1) unless it arises in the data for the equation. The case when the functions a, b, c, and d have jump discontinuities also requires the introduction of generalized solutions, and these are considered in the exercises.

Example 2.5. Equations with Singular Points. In our discussion of (2.2.1) it has been assumed that the coefficients a and b do not both vanish at some point (x,t). If they are both zero at $x = x_0$, $t = t_0$, (x_0, t_0) is a critical point for the (autonomous) system (2.2.2), $x'(s) = a(x,t)$, $t'(s) = b(x,t)$, from which the characteristic base curves are determined. Either two or more solution curves of this system intersect at the critical point or they spiral around or encircle that point. In each case, difficulties arise in the formulation of initial value problems for (2.2.1) in the neighborhood of the critical point, or singularities occur at that point in solutions with initial data given outside its neighborhood. Consequently, we refer to that point as a *singular point* for the equation (2.2.1).

We consider the initial value problem

$$x\frac{\partial v(x,t)}{\partial x} + t\frac{\partial v(x,t)}{\partial t} = cv(x,t), \qquad v(x,1) = f(x), \tag{2.2.26}$$

where c is a constant in the PDE. The PDE has a singular point at $(x,t) = (0,0)$, and we examine its effect on the solution of the initial value problem.

The initial curve C for the problem can be parameterized as

$$x = \tau, \qquad t = 1, \qquad v = f(\tau), \tag{2.2.27}$$

and the characteristic equations are

$$\frac{dx}{ds} = x, \qquad \frac{dt}{ds} = t, \qquad \frac{dv}{ds} = cv. \tag{2.2.28}$$

The solutions of (2.2.28) that satisfy (2.2.27) at $s = 0$ are

$$x(s,\tau) = \tau e^s, \qquad t(s,\tau) = e^s, \qquad v(s,\tau) = f(\tau)e^{cs}. \tag{2.2.29}$$

Solving for s and τ in terms of x and t, gives $s = \log t$, $\tau = x/t$, which is valid for $t > 0$. Inserting this into the equation for v, we obtain as the solution of the initial value problem (2.2.26)

$$v(x,t) = f\left(\frac{x}{t}\right) t^c. \tag{2.2.30}$$

The critical point $(x,t) = (0,0)$ for the system $x'(s) = x$, $t'(s) = t$ is called a *proper node* in the theory of ODEs. The characteristic base curves are the pencil of straight lines $x = \lambda t$ $(-\infty < \lambda < \infty)$, together with $t = 0$, all of which intersect at the origin. Thus, if $c = 0$ in (2.2.26) so that (2.2.30) reduces to $v = f(x/t)$, the solution is not defined for arbitrary $f(x)$, when $t = 0$ and $x \neq 0$, and is infinitely multivalued at the origin when $x = 0$. Furthermore, if $f(x) = 1$, (2.2.30) is not defined on the x-axis if $c < 0$; is continuous but not differentiable there if $c = 1/3$; but is a valid solution for all x and t if $c = 1$.

Next, we consider the initial value problem

$$t\frac{\partial v(x,t)}{\partial x} - x\frac{\partial v(x,t)}{\partial t} = 0, \qquad v(x,0) = f(x),\ x > 0. \tag{2.2.31}$$

The PDE has a singular point at $(x,t) = (0,0)$. The characteristic equations for (2.2.31) are

$$x'(s) = t, \qquad t'(s) = -x, \qquad v'(s) = 0. \tag{2.2.32}$$

If we parameterize the initial curve C as $x = \tau$, $t = 0$, $v = f(\tau)$, we obtain the solution of the system (2.2.32) that satisfies these initial conditions at $s = 0$, in the form $x(s,\tau) = \tau\cos s$, $t(s,\tau) = -\tau\sin s$, $v(s,\tau) = f(\tau)$. On solving for τ in terms of x and t, we find that $\tau = (x^2 + t^2)^{1/2}$, so that

$$v(x,t) = f\left((x^2 + t^2)^{1/2}\right). \tag{2.2.33}$$

The solution $v(x,t)$ is completely determined for $x^2 + t^2 > 0$ in terms of the data given on the positive x-axis. It is not differentiable at the origin.

The critical point $(0,0)$ for the system $x'(s) = t$, $t'(s) = -x$, is called a *center*. The characteristic base curves are concentric circles whose center is at the origin. They are fully determined by specifying a point that they pass through on the positive x-axis. The solution $v(x,t)$ is constant on each circle.

Characteristic Initial Value Problems

If the *Jacobian determinant* $\Delta(s,\tau) = x_s t_\tau - t_s x_\tau$ associated with the first two equations in (2.2.5) vanishes on the initial (base) curve $x = x(\tau)$, $t = t(\tau)$, the tangent vector to the curve has the direction of the vector $[a, b]$ at each point. Therefore, it must be a *characteristic base curve.* We change the variable on the initial base curve from τ to s, with s chosen so that the components $x'(s)$ and $t'(s)$ of the tangent vector satisfy (2.2.2). Then we have a *characteristic initial value problem* if the initial value v, expressed in terms of the parameter s, satisfies equation (2.2.3). If this equation is not satisfied, the initial value problem has either no solution or a nonsmooth solution.

As an example, we consider (2.2.26) with the initial condition $v = f(t)$ on $x = t$ for $t > 0$. On representing the initial curve parametrically as $x = \tau$, $t = \tau$ and $v = f(\tau)$, we find that the solution of the characteristic equations is $x = \tau e^s$, $t = \tau e^s$, and $v = f(\tau)e^{cs}$. The Jacobian determinant $\Delta(s,\tau)$ vanishes not only at $s = 0$ but identically. [As shown in the exercises, this is always the case if the initial base curve is characteristic and the coefficients a and b in (2.2.1) are continuously differentiable.] To determine what condition $f(t)$ must satisfy for this to be a characteristic initial value problem, we set $\tau = \tau(s)$ and find that the first two characteristic equations in (2.2.28) are satisfied if $\tau = e^s$ so that $x = t = e^s$. Then $v = f(e^s)$ and the third equation in (2.2.36) requires that $e^s f'(e^s) = cf(e^s)$. The general solution of this equation is $f(e^s) = \beta e^{cs}$, where β is an arbitrary constant. In terms of x and t this requires that on $x = t$ we must have $v = f(t) = \beta t^c$ to obtain a characteristic initial value problem. If $v = f(t)$ does not have the required functional form, the problem has no solution.

If $a \neq 0$ in (2.2.1), we can divide through by a in (2.2.1) and replace the characteristic equations (2.2.2) and (2.2.3) by

$$\frac{dv}{dx} = \frac{c}{a}v + \frac{d}{a} \quad \text{on} \quad \frac{dt}{dx} = \frac{b}{a}. \tag{2.2.34}$$

Similarly, if $b \neq 0$ in (2.2.1), we obtain the characteristic equations

$$\frac{dv}{dt} = \frac{c}{b}v + \frac{d}{b} \quad \text{on} \quad \frac{dx}{dt} = \frac{a}{b}. \tag{2.2.35}$$

Consider the initial value problem for (2.2.1) with $v = f(x)$ on $t = h(x)$. If $t = h(x)$ is a solution of $t'(x) = b/a$, we have a characteristic initial value problem, if $v = f(x)$ satisfies $f'(x) = (c/a)f(x) + d/a$. If we have $v = f(t)$ on $x = g(t)$, and $x = g(t)$ satisfies the equation $x'(t) = a/b$, we must have $f'(t) = (c/b)f(t) + d/b$ for this to be a characteristic initial value problem. Both characteristic initial value problems have nonunique solutions, as is shown most easily by constructing a general solution of (2.2.1) for each case.

Let $\phi(x,t) = \lambda$ represent a family of characteristic base curves for (2.2.1). In terms of (2.2.34), we assume that they can be expressed as $t = h(x,\lambda)$ and for each λ we solve the linear ordinary differential equation for v. If we express the general solution as $v = F(x,\lambda)$, the general solution of (2.2.1) is $v = F(x,\phi(x,t))$. In connection with (2.2.35), we express the characteristics as $x = g(t,\lambda)$ and for each λ we solve

for v. With the general solution given as $v = G(t, \lambda)$, the general solution of (2.2.1) is $v = G(t, \phi(x,t))$. [We note that $\phi(x,t)$ satisfies the equation $a\,\phi_x + b\,\phi_t = 0$.] Each of the foregoing general solutions of the ODEs contains an arbitrary constant that can vary with λ. Consequently, the corresponding general solutions of (2.2.1) depend on an arbitrary function, each of which is constant on each characteristic base curve. Therefore, a characteristic initial value problem has infinitely many solutions, since all that is required of the arbitrary functions is that their value at $\phi(x,t) = \lambda_0$, the initial characteristic (along which they are constant), be chosen so that the initial condition is satisfied.

Example 2.6. Characteristic Initial Value Problems. The PDE (2.2.26) can be expressed in either of the forms (2.2.34) or (2.2.35), say if x and t are both positive. We obtain either

$$\frac{dv}{dx} = \frac{c}{x}\,v \qquad \text{on} \qquad \frac{dt}{dx} = \frac{t}{x}, \tag{2.2.36}$$

or

$$\frac{dv}{dt} = \frac{c}{t}\,v \qquad \text{on} \qquad \frac{dx}{dt} = \frac{x}{t}. \tag{2.2.37}$$

The family of characteristic base curves is $\phi(x,t) = x/t = \lambda$. This yields $t = h(x,\lambda) = x/\lambda$ or $x = g(t,\lambda) = \lambda t$ as the explicit representations of these curves. (We note that $t = 0$ is also a characteristic curve.) Along $t = x/\lambda$, we find that $v = F(x,\lambda) = \beta(\lambda)x^c$, while on $x = \lambda t$ we obtain $v = G(t,\lambda) = \beta(\lambda)t^c$. Thus, the general solution of (2.2.26) has the form $v(x,t) = \beta(x/t)x^c$ or $v(x,t) = \beta(x/t)t^c$ and they both contain an arbitrary function $\beta(x/t)$. [The arbitrary function $\beta(x/t)$ can be replaced by the arbitrary function $\gamma(t/x)$. Then, at $t = 0$ the solution must be of the form $\gamma(0)x^c$.]

For example, if we specify that $v = 5x^c$ on the line $t = x/3$, we have a characteristic initial value problem. The (nonunique) solution is $v(x,t) = \beta(x/t)x^c$, with $\beta(3) = 5$ but $\beta(x/t)$ otherwise arbitrary. However, if $v = \sin x$ on $t = x/3$, the initial value problem has no solution, since v does not satisfy the ordinary differential equation in (2.2.36).

An interesting characteristic initial value problem occurs for the equation

$$\frac{\partial v(x,t)}{\partial x} + \sqrt{t}\,\frac{\partial v(x,t)}{\partial t} = v(x,t), \tag{2.2.38}$$

where we assume that $t \geq 0$. The characteristic equations can be given in the form (2.2.34) as

$$\frac{dv}{dx} = v \qquad \text{on} \qquad \frac{dt}{dx} = \sqrt{t}. \tag{2.2.39}$$

To study the characteristic base curves that originate on the x-axis, we must solve the nonlinear equation $t'(x) = \sqrt{t(x)}$ with $t(\lambda) = 0$. The solution of the initial value problem for each value of the parameter λ is not unique. Two solutions are $t = 0$ and $t = \frac{1}{4}(x-\lambda)^2$, and there are infinitely many others.

We see that $t = \frac{1}{4}(x-\lambda)^2$ is a family of characteristics for (2.2.38), and each curve in this family is tangent to the x-axis at its point of intersection. Thus, the

line $t = 0$ is an envelope of characteristics (see the Appendix) in addition to being a characteristic, as we have shown. In view of the discussion that precedes this example, the characteristic family is given as $\phi(x,t) = x - 2\sqrt{t} = \lambda$, and we easily find from (2.2.39) that the general solution of (2.2.38) is

$$v(x,t) = \beta(x - 2\sqrt{t})e^x \tag{2.2.40}$$

for arbitrary $\beta(\lambda)$. On the characteristic $t = 0$, (2.2.39) yields $v = \beta e^x$.

Noting the form that v must have on the characteristic $t = 0$, we expect that the initial value problem for (2.2.38) with $v(x,0) = f(x)$ does not have a solution unless $f(x) = \beta\, e^x$ with constant β. The solution is then given as $v(x,t) = \beta e^x$. We may, however, use the general solution (2.2.40) to obtain, for arbitrary $f(x)$, the solution $v(x,t) = f(x - 2\sqrt{t})e^{2\sqrt{t}}$, which is continuous at $t = 0$ but not differentiable there. It must be understood as a generalized solution. If $f(x) = \beta e^x$, this solution agrees with that obtained above and is differentiable at $t = 0$. In either case the solution is determined uniquely.

The foregoing results appear to be at variance with our general conclusions concerning characteristic initial value problems. If $v(x,0) = f(x)$ has the required exponential form, the initial value problem is characteristic and yet the solution was found to be unique. If $f(x)$ does not have the required form, the initial value problem should have no solution and yet we found a generalized solution. These inconsistencies can be traced to the fact that the coefficient $\sqrt{t}$ of v in (2.2.38) is not differentiable at $t = 0$. The line $t = 0$ is a singular solution of $t'(x) = \sqrt{t(x)}$, and it cannot be obtained from the family of characteristics determined above for any specific choice of λ. Therefore, the general solution of (2.2.38) given above does not yield nonuniqueness here. Since each of these characteristics intersects $t = 0$, the general solution can be used to solve an initial value problem given on the x-axis. Each characteristic intersects the x-axis only once, so the solution is unique. Because the intersection is tangential, the solution is not differentiable at $t = 0$.

Exercises 2.2

2.2.1. Solve the initial value problem for the damped unidirectional wave equation $v_t(x,t) + cv_x(x,t) + \lambda v(x,t) = 0$, $v(x,0) = F(x)$, where $\lambda > 0$ and $F(x)$ is prescribed.

2.2.2. (a) Solve the initial value problem for the inhomogeneous equation $v_t(x,t) + cv_x(x,t) = f(x,t)$, $v(x,0) = F(x)$, with prescribed $f(x,t)$ and $F(x)$. (b) Solve the problem when $f(x,t) = xt$ and $F(x) = \sin(x)$.

2.2.3. Discuss the solution of the wave equation (2.1.1) for the following cases: (a) $v(x,0) = f(x) = x$, $v_t(x,0) = g(x) = 0$. (b) $f(x) = 0$, $g(x) = x$. (c) $f(x) = \sin(x)$, $g(x) = -\gamma\cos(x)$. (d) $f(x) = \sin(x)$ $g(x) = \gamma\cos(x)$. Based on

the results obtained, observe that solutions of the wave equation may or may not have the form of traveling waves.

2.2.4. Consider the inhomogeneous wave equation $v_{tt}(x,t) - \gamma^2 v_{xx}(x,t) = F(x,t)$. (a) Apply the method of Example 2.1 to reduce it to the system $v_t(x,t) - \gamma v_x(x,t) = u(x,t)$, $u_t(x,t) + \gamma u_x(x,t) = F(x,t)$. (b) Solve the initial value problem for the inhomogeneous wave equation with the initial data $v(x,0) = f(x)$, $v_t(x,0) = g(x)$.

2.2.5. Solve the *signaling* problem for (2.2.6) in the region $x > 0$ with the boundary condition $v(0,t) = G(t)$ for $-\infty < t < +\infty$.

2.2.6. Show that the initial value problem $v_t(x,t) + v_x(x,t) = 0$, $v(x,t) = x$ on $x^2 + t^2 = 1$, has no solution. However, if the initial data are given only over the semicircle that lies in the half-plane $x + t \leq 0$, the solution exists but is not differentiable along the characteristic base curves that issue from the two endpoints of the semicircle.

2.2.7. Solve the initial value problem $v_t(x,t) + e^x v_x(x,t) = 0$, $v(x,0) = x$.

2.2.8. (a) Show that the initial and boundary value problem for (2.2.6) with $c > 0$ can be solved in the quarter-plane $x > 0$, $t > 0$ with the data $v(x,0) = F(x)$ and $v(0,t) = G(t)$, if $F(x)$ and $G(t)$ are arbitrary smooth functions that satisfy the conditions $G(0) = F(0)$ and $G'(0) = -cF'(0)$. (b) Show that for $c < 0$, the problem can be solved only if the data $F(x)$ and $G(t)$ satisfy a compatibility condition. In this case, the characteristic base curves intersect both the x- and the t-axes.

2.2.9. If $v_x(x,t)$ has a jump across the characteristic base curve $x = x(s)$, $t = t(s)$ but the generalized solution v of (2.2.1) is continuous there, show that $\frac{d}{ds}[v_x] = (c - a_x + (a/b)\, b_x)\, [v_x]$. If v_x has a jump at a point on the initial curve, this determines the variation of $[v_x]$ along the characteristic through that point. Once $[v_x]$ is determined, $[v_t]$ can be obtained from (2.2.19).

2.2.10. Let $f(\tau)$ in (2.2.27) be given as $f(\tau) = 0$ for $\tau < 0$ and $f(\tau) = \tau$ for $\tau > 0$. Obtain the generalized solution of this problem and use it to determine the jump in v_x. Show that this jump satisfies the equation found in Exercise 2.2.9.

2.2.11. If the function $d(x,t)$ in (2.2.1) has a jump discontinuity across a (noncharacteristic) curve but the coefficients $a(x,t)$, $b(x,t)$, and $c(x,t)$ are smooth functions, obtain an equation for the jumps in $v_x(x,t)$ and $v_t(x,t)$ across that curve. [Assume that $v(x,t)$ is continuous there.] Show that if $d(x,t)$ is discontinuous across a characteristic base curve, the solution $v(x,t)$ must also have a jump along that curve.

2.2.12. Solve the initial value problem of Exercise 2.2.2 if (a) $f(x,t) = H(x)$ and (b) $f(x,t) = H(x - ct)$. In both cases, set $F(x) = 0$. [$H(z)$ is the Heaviside function defined as $H(z) = 0$ for $z < 0$ and $H(z) = 1$ for $z > 0$.] Verify the conclusions of Exercise 2.2.11.

2.2.13. Let $f(x) = H(x)$ in (2.2.31), where $H(x)$ is the Heaviside function (see Exercise 2.2.12). Find the generalized solution of this initial value problem for (2.2.26) and verify that (2.2.25) is satisfied for this equation. Also, show that the jump in

v satisfies the ordinary differential equation given in the discussion in the text that follows equation (2.2.25).

2.2.14. Let the coefficients $a(x,t)$ and $b(x,t)$ in (2.2.1) have jump discontinuities across a curve $x = x(s)$, $t = t(s)$. Assume that $v(x,t)$ also has a jump across that curve and use the integral relation (2.2.22) to show that the values of $a(x,t)$, $b(x,t)$, and $v(x,t)$ on opposite sides of the curve (with unit normal vector $\mathbf{n}$) are related by $[a_1, b_1] \cdot \mathbf{n} v_1 = [a_2, b_2] \cdot \mathbf{n} v_2$. (The subscripts denote values on opposite sides of the curve.)

2.2.15. Consider the initial value problem $v_t(x,t) + c(x)v_x(x,t) = 0$, $c(x) = c_1$, $x < 0$, $c(x) = c_2$, $x > 0$, $v(x,0) = f(x)$, where c_1 and c_2 are positive constants. Use the jump condition obtained in Exercise 2.2.14 to solve this problem. *Hint*: First solve for $v(x,t)$ in the region $x < 0$ and $t > 0$. Then use the jump condition as well as the initial condition to solve for v in $x > 0$ and $t > 0$. Note that $v(x,t)$ has a jump across $x = c_2 t$ unless $f(0) = 0$.

2.2.16. Show that the initial value problem $(t - x)v_t(x,t) - (t + x)v_x(x,t) = 0$, $v(x,0) = f(x)$, $x > 0$, has no solution if $f(x)$ is an arbitrary function. *Hint*: The singular point at the origin gives rise to characteristics that spiral around the origin.

2.2.17. Show that the equation $a(x)v_x(x,t) + b(t)v_t(x,t) = 0$ has the general solution $v(x,t) = F[A(x) - B(t)]$, where $A'(x) = 1/a(x)$ and $B'(t) = 1/b(t)$.

2.2.18. Show that the equation $a(t)v_x(x,t) + b(x)v_t(x,t) = 0$ has the general solution $v(x,t) = F[B(x) - A(t)]$, where $B'(x) = b(x)$ and $A'(t) = a(t)$.

2.2.19. Given the initial value problem $tv_t(x,t) + xv_t(x,t) = cv(x,t)$, $v(x,x) = f(x)$, $x > 0$, where c is a constant, determine a condition on the function $f(x)$ so that this becomes a characteristic initial value problem. Supposing that $f(x)$ satisfies that condition, obtain a (nonunique) solution of the problem.

2.2.20. Show that the initial value problem $u_t(x,t) + u_x(x,t) = x$, $u(x,x) = 1$ has no solution. Observe that the initial curve $t = x$ is a characteristic base curve and explain why this is not a characteristic initial value problem.

2.2.21. (a) Show that $u_t(x,t) + cu_x(x,t) = F(x,t)$, $u(x,x/c) = f(x)$ is a characteristic initial value problem if $f'(x) = (1/c)F[x, x/c]$. (b) Verify that

$$u(x,t) = \frac{1}{c}\int_{x-ct}^{x} F\left[r, \frac{r - x + ct}{c}\right] dr + g(x - ct)$$

is a solution of the characteristic initial value problem of part (a) if $g(0) = f(0)$ but $g(z)$ is an otherwise arbitrary function. (c) Let $F(x,t) = \cos(x + t)$ in part (a). Determine the appropriate choice of $f(x)$ and obtain the general solution of the problem.

2.2.22. Show that the Jacobian $\Delta(s,\tau)$ satisfies the equation $\partial\Delta(s,\tau)/\partial s = (a_x + b_t)\Delta(s,\tau)$. Thus, if $\Delta(s,\tau)$ vanishes at $s = 0$ and a_x and b_t are continuous, $\Delta(s,\tau)$ vanishes identically. *Hint*: Use equations (2.2.2).

2.2.23. Determine $\Delta(s,r)$ for equation (2.2.38) with initial data on $t = 0$. Note that the Jacobian vanishes at $s = 0$ but not for $s > 0$. Show that $\Delta(s,\tau)$ satisfies the equation of Exercise 2.2.22 and explain why the result of that exercise is not contradicted.

2.2.24. Obtain the general solution of the equation $v_t(x,t) + t^{1/3}v_x(x,t) = v(x,t)$. Show that $t = 0$ is a characteristic for this equation but that the initial value problem with $v(x,0) = f(x)$ has a unique but nondifferentiable solution (at $t = 0$) for arbitrary $f(x)$. Show that the characteristics are all tangent to the x-axis.

2.2.25. Generalize the method of characteristics to problems in three dimensions. Given the linear equation $av_x + bv_y + cv_z = dv + e$, where a, b, c, d, e, and v are functions of (x,y,z), obtain the characteristic equations $dx/ds = a$, $dy/ds = b$, $dz/ds = c$, $dv/ds = dv + e$. The initial value problem for this case specifies an initial hypersurface given parametrically in (x,y,z,v)-space, as $x = x(\lambda,\tau)$, $y = y(\lambda,\tau)$, $z = z(\lambda,\tau)$, $v = v(\lambda,\tau)$, with λ and τ as the parameters. The family of characteristic curves is now given as $x = x(s,\lambda,\tau)$, $y = y(s,\lambda,\tau)$, $z = z(s,\lambda,\tau)$, $v = v(s,\lambda,\tau)$, with s as the running parameter along a curve and (λ,τ) as a two-parameter family specifying the individual curves. If the equations for (x,y,z) can be inverted to yield (s,λ,τ) as smooth functions of (x,y,z), the function $v = v(x,y,z)$ obtained thereby is a solution of the initial value problem. Discuss situations in which the method of characteristics may not give a solution or yields a nonunique solution.

2.2.26. Using the method of characteristics, solve the initial value problem $v_x(x,y,z) + v_y(x,y,z) + v_z(x,y,z) = 0$, $v(x,y,0) = f(x,y)$, where $f(x,y)$ is prescribed. (a) Parametrize the initial data and let $x = \lambda$, $y = \tau$, $z = 0$, and $v = f(\lambda,\tau)$ and set up the characteristic equations. (b) Show that the solution of these equations is $x = s + \lambda$, $y = s + \tau$, $z = s$, $v = f(\lambda,\tau)$. (c) Conclude that the solution of the initial value problem is $v(x,y,z) = f(x - z, y - z)$.

2.2.27. Extend the method of characteristics to linear equations in n dimensions. Given the equation $\sum_{i=1}^{n} a_i(x_1,\ldots,x_n)\partial v/\partial x_i = b(x_1,\ldots,x_n)v + c(x_1,\ldots,x_n)$, where $v = v(x_1,\ldots,x_n)$, obtain the system of ODEs $dx_i/ds = a_i$, $(i = 1,\ldots,n)$, $dv/ds = bv + c$. The initial values for this problem are given as $x_i = x_i(\tau_1,\ldots,\tau_{n-1})$ and $v = v(\tau_1,\ldots,\tau_{n-1})$ at $s = 0$, with $\tau_1,\ldots,\tau_{n-1}$ as parameters. The characteristic curves are $x_i = x_i(s,\tau_1,\ldots,\tau_{n-1})$ and $v = v(s,\tau_1,\ldots,\tau_{n-1})$, where $i = 1,2,...,n$. If $s,\tau_1,\ldots,\tau_{n-1}$ can be solved for in terms of $x_1,\ldots,x_n$, then $v = v(x_1,\ldots,x_n)$ is a solution of the initial value problem.

2.2.28. Use the method of characteristics to solve the problem $\sum_{i=1}^{n} v_{x_i}(x_1,\ldots,x_n) = 0$, $v(x_1,\ldots,x_{n-1},0) = f(x_1,\ldots,x_{n-1})$ where $f(x_1,\ldots,x_{n-1})$ is given. (a) Show that the solution of the characteristic equations can be given as $x_i = s + \tau_i$ $(i = 1,\ldots,n)$, $x_n = s$, $v = f(\tau_1,\ldots,\tau_{n-1})$. (b) Conclude that $v(x_1,\ldots,x_n) = f(x_1 - x_n, x_2 - x_n,\ldots,x_{n-1} - x_n)$ is the solution of the problem.

2.3 QUASILINEAR FIRST ORDER PARTIAL DIFFERENTIAL EQUATIONS

A first order partial differential equation of the form

$$a(x,t,u)\,\frac{\partial u}{\partial x} + b(x,t,u)\,\frac{\partial u}{\partial t} = c(x,t,u) \tag{2.3.1}$$

is said to be *quasilinear*, since it is linear in the derivative terms but may contain nonlinear expressions of the form $u\,\partial u/\partial t$ or u^2. Such equations occur in a variety of nonlinear wave propagation problems and in other contexts, but we emphasize the wave propagation aspect in the examples. Proceeding as in Section 2.2, we interpret (2.3.1) geometrically for the purpose of constructing a solution via the *method of characteristics.*

Method of Characteristics

We assume that a solution $u = u(x,t)$ of (2.3.1) can be found and examine its properties as implied by (2.3.1). The solution $u = u(x,t)$ is called an *integral surface*, and we express it in implicit form as $F(x,t,u) = u(x,t) - u = 0$. The gradient vector $\nabla F = [\partial u/\partial x, \partial u/\partial t, -1]$ is normal to the integral surface $F(x,t,u) = 0$. On transposing the term c in (2.3.1), we can express the resulting equation as a scalar or dot product,

$$a\,\frac{\partial u}{\partial x} + b\,\frac{\partial u}{\partial t} - c = [a,b,c]\cdot\left[\frac{\partial u}{\partial x}\,,\frac{\partial u}{\partial t}\,,-1\right] = 0. \tag{2.3.2}$$

The vanishing of the dot product of the vector $[a,b,c]$ with the gradient vector ∇F, implies that these vectors are orthogonal. Accordingly, the vector $[a,b,c]$ lies in the tangent plane of the integral surface $u = u(x,t)$ at each point in the (x,t,u)-space where ∇F is defined and nonzero. At each point (x,t,u), the vector $[a,b,c]$ determines a direction that is called the *characteristic direction.* As a result, the vector $[a,b,c]$ determines a *characteristic direction field* in (x,t,u)-space, and we can construct a family of curves that have the characteristic direction at each point. If the parametric form of these curves is $x = x(s)$, $t = t(s)$, and $u = u(s)$, we must have

$$\frac{dx}{ds} = a(x,t,u), \qquad \frac{dt}{ds} = b(x,t,u), \qquad \frac{du}{ds} = c(x,t,u), \tag{2.3.3}$$

since $[dx/ds,\ dt/ds,\ du/ds]$ is the tangent vector along the curves. The characteristic equations (2.3.3) differ from those in the linear case, since the equations for x and t are not, in general, uncoupled from the equation for u. The solutions of (2.3.3) are called the *characteristic curves* of the quasilinear equation (2.3.1).

Assuming that a, b, and c are sufficiently smooth and do not all vanish at the same point, the theory of ordinary differential equations guarantees that a unique

characteristic curve passes through each point (x_0, t_0, u_0). The *initial value problem* for (2.3.1) requires that $u(x,t)$ be specified on a given curve in the (x,t)-plane. This determines a curve C in (x,t,u)-space referred to as the *initial curve*. To solve this initial value problem, we pass a characteristic curve through each point of the initial curve C. If these curves generate a surface, this *integral surface* is the solution of the initial value problem. We now state and outline the proof of a theorem that gives conditions under which a unique solution of the initial value problem for (2.3.1) can be obtained.

Let a, b, and c in (2.3.1) *have continuous partial derivatives in all three variables. Suppose that the initial curve C, given parametrically as $x = x(\tau)$, $t = t(\tau)$, and $u = u(\tau)$, has a continuous tangent vector and that*

$$\Delta(\tau) = \frac{dt}{d\tau}\, a[x(\tau), t(\tau), u(\tau)] - \frac{dx}{d\tau}\, b[x(\tau), t(\tau), u(\tau)] \neq 0 \tag{2.3.4}$$

on C. Then there exists one and only one solution $u = u(x,t)$, defined in some neighborhood of the initial curve C, that satisfies (2.3.1) *and the initial condition $u[x(\tau), t(\tau))] = u(\tau)$.*

The proof of this theorem proceeds along the following lines. The characteristic system (2.3.3) with initial conditions at $s = 0$ given as $x = x(\tau)$, $t = t(\tau)$, $u = u(\tau)$ has a unique solution of the form

$$x = x(s,\tau), \qquad t = t(s,\tau), \qquad u = u(s,\tau), \tag{2.3.5}$$

with continuous derivatives in s and τ, and with

$$x(0,\tau) = x(\tau), \qquad t(0,\tau) = t(\tau), \qquad u(0,\tau) = u(\tau). \tag{2.3.6}$$

This follows from the existence and uniqueness theory for ordinary differential equations. The Jacobian of the transformation $x = x(s,\tau)$, $t = t(s,\tau)$ at $s = 0$ is

$$\Delta(s,\tau)|_{s=0} = \left| \begin{array}{cc} \frac{\partial x}{\partial s} & \frac{\partial x}{\partial \tau} \\ \frac{\partial t}{\partial s} & \frac{\partial t}{\partial \tau} \end{array} \right|_{s=0} = \frac{dt}{d\tau}\, a - \frac{dx}{d\tau}\, b\Big|_{s=0} = \Delta(\tau), \tag{2.3.7}$$

and it is nonzero in view of the assumption (2.3.4). By the continuity assumption, the Jacobian determinant does not vanish in a neighborhood of the initial curve. Therefore, the implicit function theorem guarantees that we can solve for s and τ as functions of x and t near the initial curve.

Then

$$u(s,\tau) = u(s(x,t), \tau(x,t)) = U(x,t) \tag{2.3.8}$$

is a solution of (2.3.1). This is readily seen on substituting $u(s(x,t), \tau(x,t))$ into (2.3.1), using the chain rule and the characteristic equations (2.3.3). The *uniqueness* of the solution follows since any two integral surfaces that contain the same initial curve must coincide along all the characteristic curves passing through the initial

curve. This is a consequence of the uniqueness theorem for the initial value problem for (2.3.3). This completes our proof.

The condition (2.3.4) essentially means that the initial curve C is not a characteristic or is not the envelope of characteristic curves. If the initial curve is characteristic, so that $\Delta(\tau) = 0$ and, in addition, it is a solution of the characteristic equations (2.3.3), solutions exist but they are not unique. This is called a *characteristic initial value problem.* If the initial curve is the envelope of characteristic curves (see the Appendix), so that $\Delta(\tau) = 0$ along the curve but the curve does not satisfy (2.3.3), the integral surface may not be differentiable along the initial curve. If the initial curve is noncharacteristic but $\Delta(\tau) = 0$ at a discrete set of points, special problems that originate at these points occur with the solution. Examples of each of these types are considered in the exercises.

Wave Motion and Breaking

We now consider an important example of a quasilinear first order PDE, the simplest case of *unidirectional nonlinear wave motion.* It is often referred to as the *inviscid Burgers' equation* in view of its connection with the (viscous) *Burgers' equation*, which is of importance in fluid mechanics (see Section 10.3). We concentrate on the initial value problem for this equation and consider various features of *nonlinear wave motion*, that distinguish it from linear wave motion. (In other contexts, it is referred to as a *convection equation*, and when a second x-derivative term is added, as a *convection-diffusion equation.*)

Example 2.7. Unidirectional Nonlinear Wave Motion. The simplest quasilinear equation characterizing one-directional nonlinear wave motion is

$$\frac{\partial u(x,t)}{\partial t} + u(x,t)\,\frac{\partial u(x,t)}{\partial x} = 0, \tag{2.3.9}$$

the inviscid Burgers' equation. We introduce the initial value $u(x,0) = f(x)$ for a given smooth function $f(x)$.

To solve the initial value problem for (2.3.9), we parameterize the initial curve as

$$x = \tau, \qquad t = 0, \qquad u = f(\tau). \tag{2.3.10}$$

We find that $\Delta(\tau) = -1$, so that $\Delta(\tau) \neq 0$ along the entire initial curve. The characteristic equations are

$$\frac{dx}{ds} = u, \qquad \frac{dt}{ds} = 1, \qquad \frac{du}{ds} = 0, \tag{2.3.11}$$

with initial conditions at $s = 0$ given by (2.3.10). Denoting the solutions by

$x(s,\tau)$, $t(s,\tau)$, and $u(s,\tau)$, we see from $du/ds = 0$ that $u(s,\tau)$ is constant along the characteristics. Thus,

$$u(s,\tau) = u(0,\tau) = f(\tau), \tag{2.3.12}$$

on using (2.3.10). Inserting (2.3.12) into $dx/ds = u$, we immediately obtain

$$x(s,\tau) = \tau + sf(\tau), \qquad t(s,\tau) = s. \tag{2.3.13}$$

If the conditions of the theorem are satisfied, this system can be inverted near $s = 0$ to give $s = t$ and $\tau = \tau(x,t)$. Then (2.3.12) yields the solution

$$u = f(\tau(x,t)) = U(x,t). \tag{2.3.14}$$

Since $s = t$ and $\tau = x - sf(\tau) = x - tu$ in view of (2.3.13), the *implicit form* of the solution is

$$u = f(x - tu). \tag{2.3.15}$$

We consider the solution $u = u(x,t)$ of (2.3.9) to represent a *wave*. The *wave form* at the time t is given by the curve $u = u(x,t)$ in the (x,u)-plane with t as a parameter and $u = f(x)$ as the initial wave form. The (numerical) value of $u(x,t)$ represents the height of the wave at the point x and the time t. To determine the motion of the wave, we must find the velocity dx/dt of each point on the wave. The first two characteristic equations in (2.3.11) imply that $dx/dt = u$. Thus, the greater the amplitude $|u(x,t)|$ of the wave, the greater the speed of the corresponding point x on the wave. [This contrasts with the situation for linear wave motion determined by equation (2.2.6), where $dx/dt = c$ = constant, so that each point and, consequently, the entire wave form moves with a single speed.] If $u(x,t) > 0$, the point x moves to the right; if $u(x,t) = 0$, it remains fixed; whereas if $u(x,t) < 0$, it moves to the left. Thus if $u(x,t)$ takes on both positive and negative values, individual points maintain a fixed direction of motion, but different portions of the wave form can move at different speeds either to the right or to the left. Therefore, the wave motion need not be totally unidirectional as is the case for linear wave motion. Assuming that $u(x,t) > 0$, points x where $u(x,t)$ has larger values and the wave is higher, move more rapidly to the right than points x where $u(x,t)$ has smaller values and the wave is lower. If, initially, there are higher portions of the wave form located to the left or the rear of lower portions, the higher points may eventually overtake (and pass) the lower points. The wave is said to *break* the first time this happens. At the time t when the wave breaks, the function $u(x,t)$ becomes multivalued and is no longer a valid solution of (2.3.9). Similar difficulties can occur if $u(x,t)$ is not restricted to be positive. The geometric aspects of the breaking process are demonstrated in Example 2.8.

In many processes described by the quasilinear wave equation (2.3.9), the function $u(x,t)$ represents a physical quantity, such as density, which is intrinsically expected to be single valued. Thus when the wave $u(x,t)$ breaks and becomes multivalued, the equation (2.3.9) that describes the physical process is no longer an acceptable model for the process. In general, this means that certain higher derivative terms that were neglected in the derivation of the quasilinear equation become significant and must be retained. It will be shown in Section 10.3, in a discussion of the viscous *Burgers' equation*, how the addition of a second derivative term to the inviscid Burgers' equation (2.3.9) can yield solutions valid for all time. A similar situation was encountered in our discussion of the diffusion equation (1.1.15). With Dirac delta function initial data, the solution of the diffusion equation which is given in (1.1.20) is a smooth function for all x and for $t > 0$. However, if the diffusion coefficient D is equated to zero and the first order unidirectional wave equation (1.2.33) results, the solution (1.2.34) remains sharply singular for all $t > 0$. This indicates that the inclusion of higher derivative terms tends to smooth out the solution.

An alternative method for dealing with the breaking phenomenon is to introduce a discontinuous solution of (2.3.9) known as a *shock wave* that extends the validity of the solution beyond the breaking time. Such solutions of (2.3.9) must be interpreted in a generalized sense, as was done for discontinuous solutions of the linear equation (2.2.1). This approach has the advantage that a complete analysis of the breaking wave solution can be carried out on the basis of the first order equation (2.3.9) without having to analyze a higher order nonlinear equation. However, as will be seen, the correct determination of the (shock wave) discontinuity solution is not as straightforward as it was in the linear case. In general, an appeal to the physical origin of the problem is needed to make the correct determination.

In any case, even if the wave breaks at the time t, the solution (2.3.15) of the initial value problem remains valid until that time. Thus it is important to determine the time when the wave $u(x,t)$ first begins to break. In addition, to determine the shock wave it is necessary to know not only the time but also the point(s) x at which the wave breaks. We consider two methods for doing so.

First, we use implicit differentiation to determine the slope of the wave form $u = u(x,t) = f(x - tu)$ at the time t. The slope $\partial u/\partial x$ is readily found to be

$$\frac{\partial u}{\partial x} = \frac{f'(x - tu)}{1 + tf'(x - tu)}. \tag{2.3.16}$$

From (2.3.13) we see that the characteristic base curves for (2.3.9) are the family of straight lines $x - tf(\tau) = \tau$, with τ as a parameter. When the denominator on the right side of (2.3.16) [i.e., $1 + tf'(x - tu) = 1 + tf'(\tau)$] first vanishes, the slope $\partial u/\partial x$ becomes infinite and the wave begins to break. Since the denominator vanishes when

$$t = -\frac{1}{f'(\tau)}, \tag{2.3.17}$$

the *breaking time* is determined from the value of τ at which t has its smallest non-negative value.

The second method for determining the breaking time considers the characteristic curves (2.3.12)–(2.3.13) of the equation (2.3.9). We have shown that $u = f(\tau)$ on the characteristic base curves

$$x - tf(\tau) = \tau, \tag{2.3.18}$$

where τ is a parameter. If for two or more values of τ the straight lines (2.3.18) intersect, u will in general be multivalued at the intersection point, for $u(x,t)$ must equal $f(\tau)$ at that point, and $f(\tau)$ may have different values on each of the lines (2.3.18) that intersect at the point. To find the possible intersection points, we must determine the envelope of the family of straight lines (2.3.18). As shown in the Appendix, if $F(x,t,\tau) = 0$ is a one-parameter family of curves, the envelope of this family is obtained by eliminating the parameter τ from the system $\{F(x,t,\tau) = 0,\ F_\tau(x,t,\tau) = 0\}$. In our case, $F = x - tf(\tau) - \tau = 0$ and $\partial F/\partial\tau = -tf'(\tau) - 1 = 0$. The equation $\partial F/\partial\tau = 0$ shows that the times at which the characteristic curves touch the envelope are given by $t = -1/f'(\tau)$, so that the initial breaking time is in agreement with that determined from the first method.

Once the value of τ that yields the breaking time t is determined, the point x at which the wave breaks is found from (2.3.18). There may be more than one value of the parameter τ that yields the same breaking time t. In general, each of the values of τ yields a different breaking point x. It can also happen that different parts of the wave break at different times. While the classical (differentiable) solution is not valid after the first breaking time, the later breaking times play a role in the construction of shock waves.

We note that if $f'(\tau) > 0$ for all τ, the right side of (2.3.17) must always be negative, so that the wave never breaks. Geometrically, this means that the initial wave form $u = f(x)$ is a monotonically increasing function of x. Since the larger the values of u, the faster the corresponding points on the wave move, no point x on the wave can overtake any other point and no breaking occurs. However, if $f'(\tau) < 0$ for all τ, the initial wave form is a monotonically decreasing function. Thus points in the rear portion of the wave move faster than points in the front portion and eventually overtake them. When this happens, the wave breaks. If $f'(\tau)$ takes on both positive and negative values, wave breaking generally occurs.

The foregoing results concerning the breaking of the wave $u = f(x - tu)$ can be interpreted from a different point of view. From (2.3.17)–(2.3.18) we conclude that the breaking time is determined from the point on the curve $x = \tau - f(\tau)/f'(\tau)$, $t = -1/f'(\tau)$, $u = f(\tau)$ where t has its smallest value. As is easily verified, this curve has the property that $\Delta(\tau)$ [as defined in (2.3.4)] vanishes on it. Thus if this curve were chosen as an initial curve for (2.3.9), the theorem stated for the general case would not guarantee that this initial value problem has a unique (smooth) solution. In fact, (2.3.16) implies that the solution would not be differentiable along the initial curve. We infer from the above that even when the initial data guarantee a unique solution for the given quasilinear equation, this solution breaks down along a curve where the condition (2.3.4) [i.e., $\Delta(\tau) \neq 0$] is violated.

The breaking curve determined above is not a characteristic curve for (2.3.9) since u is not constant on it. For the sake of completeness, we show that if the initial curve

for (2.3.9) is characteristic, the initial value problem has infinitely many solutions. Let $u = c$ = constant on the line $x = ct$. [This curve is a characteristic for (2.3.11) since with $t = s$, $x = cs$, the characteristic equations (2.3.11) are satisfied.] Then $u = f(x - tu)$ is a solution of the characteristic initial value problem as long as $f(0) = c$ but f is an otherwise arbitrary but smooth function.

Unidirectional Nonlinear Wave Motion: An Example

The following example continues our discussion of the unidirectional wave equation. Three specific initial wave forms are selected and the resulting wave motions are examined in some detail. Differing features of unidirectional nonlinear wave motion are thereby brought out. Difficulties that arise as a result of the breakdown of solutions are exhibited. How these problems can be resolved will be shown in the sequel.

Example 2.8. Unidirectional Nonlinear Wave Motion: Specific Initial Data. In this example three special choices of initial values for (2.3.9) (i.e., $\partial u/\partial t + u\,\partial u/\partial x = 0$) are given and the corresponding solutions analyzed.

1. The initial value is

$$u(x, 0) = f(x) = -x. \tag{2.3.19}$$

Since $f'(x) = -1$, we expect (in view of the foregoing discussion) that the wave will break and the solution $u(x, t)$ will become multivalued. In fact [since $f(\tau) = -\tau$], (2.3.17) gives $t = 1$ as the breaking time.

Now (2.3.15) gives the implicit form of the solution as

$$u = f(x - tu) = -(x - tu), \tag{2.3.20}$$

from which it follows that

$$u(x, t) = \frac{x}{t - 1}. \tag{2.3.21}$$

We easily verify that (2.3.21) satisfies (2.3.9) and the initial condition (2.3.19), as well as the fact that it blows up at the time $t = 1$.

The motion of the wave $u(x, t) = x/(t - 1)$ is indicated in Figure 2.2. As t increases, the wave form $u(x, 0) = -x$ executes a clockwise rotation around the origin in the (x, u)-plane. Since $u = 0$ at $x = 0$, that point (i.e., $x = 0$) is stationary. Also, $|u|$ increases linearly with $|x|$, and points x farther away from the origin have a linearly increasing velocity, yielding the effect indicated in the figure. At $t = 1$ the wave form $u(x, t)$ coincides with the u-axis and becomes infinitely multivalued.

The breakdown of the solution and its multivaluedness at $t = 1$ may also be determined by considering the characteristic base curves, which for this initial value problem are given by $x = (1 - t)\tau$, with τ as a parameter. When $t = 1$, these characteristic lines all intersect at the point $(x, t) = (0, 1)$. Since each of the lines

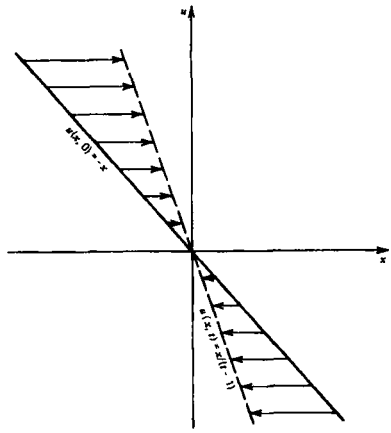

Figure 2.2 The wave motion.

carries a different value of τ, and $u(x,t)$ is constant on each of these lines, it must become infinitely multivalued at $t = 1$. This is indicated in Figure 2.3.

If the initial condition (2.3.19) is replaced by $f(x) = x$, we have $f(\tau) = \tau$ and $f'(\tau) = 1$. Then, according to (2.3.17), the wave never breaks. This fact is borne out by the solution $u(x,t) = x/(t+1)$ of this problem, which is defined for all $t > 0$.

More generally, if $u(x,0) = f(x) = \alpha + \beta x$ (with constant α and β), we determine from (2.3.15) that the solution is $u(x,t) = (\beta x + \alpha)/(\beta t + 1)$. Thus if $\beta < 0$, the wave breaks when $t = -1/\beta$, but if $\beta > 0$, the wave never breaks. The characteristic lines are given as $x = \alpha t + (\beta t + 1)\tau$, and at the breaking time $t = -1/\beta$, we have $x = -\alpha/\beta$. At this point $u(x,t)$ is stationary and the wave form rotates around this point. At the breaking time the wave form $u(x,t)$ is vertical and is infinitely multivalued.

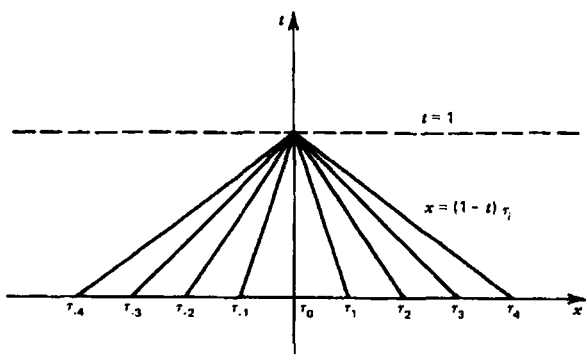

Figure 2.3 The characteristic curves.

2. The initial condition is

$$u(x,0) = f(x) = 1 - x^2. \tag{2.3.22}$$

Using (2.3.16) gives the implicit form of the solution as $u = f(x - tu) = 1 - (x - tu)^2$. Solving this quadratic equation for u gives two branches, $u(x,t) = x/t - [1 \pm \sqrt{1 + 4t(t - x)}]/2t^2$. To satisfy the initial condition (2.3.22) we must choose the minus sign and obtain

$$u(x,t) = \frac{x}{t} - \frac{1 - \sqrt{1 + 4t(t - x)}}{2t^2}. \tag{2.3.23}$$

Differentiating (2.3.23) with respect to x, we find that the slope $\partial u/\partial x$ becomes infinite when the radical in (2.3.23) vanishes, that is, when $x = t + 1/4t$. As long as $x < t + 1/4t$, the radical in (2.3.23) is real valued and the solution $u(x,t)$ is well defined. An easy calculation shows that $x < t + 1/4t$ for all $x < 1$ since $t > 0$. At $x = 1$, the radical vanishes when $t = \frac{1}{2}$, and as $x \to \infty$, the time t at which it vanishes decreases to zero. Thus if the initial condition is given over the infinite interval $-\infty < x < \infty$, the wave begins to break immediately at $x = +\infty$.

The characteristic base curves for this problem are $x - t(1 - \tau^2) = \tau$ [see (2.3.18)], where τ is a parameter. The solution $u(x,t)$ is constant along the characteristics. The envelope of this family of characteristic curves is readily found to be $x - t - 1/4t = 0$, which is identical to the curve obtained above. Since two neighboring curves of the characteristic family that intersect on the envelope correspond to two different values of τ, the solution $u(x,t)$ becomes double valued at the envelope. In fact, the radical term in the solution (2.3.23) vanishes on the curve $x = t + 1/4t$ in the (x,t)-plane, so that $u(x,t)$ splits into the two branches given in $u(x,t) = x/t - [1 \pm \sqrt{1 + 4t(t - x)}]/2t^2$, along the envelope $x = t + 1/4t$. An indication of how the wave $u(x,t)$ propagates is given in Figure 2.4.

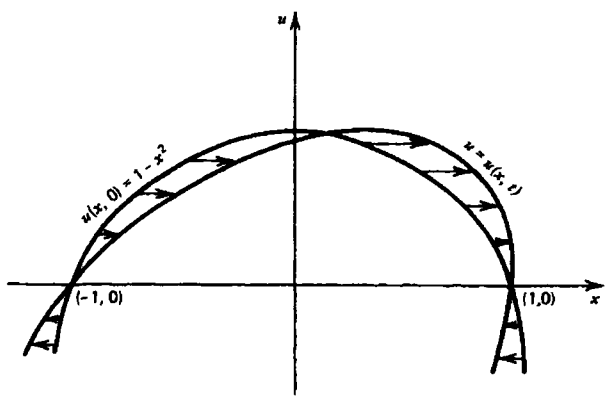

Figure 2.4 The wave motion.

If the initial data (2.3.22) are restricted to the interval $-\infty < x \leq a$ where $a > 1$, the foregoing discussion implies that the resulting wave will not break immediately. However, some care must be exercised in determining the breaking time. On applying the formula (2.3.17) to this problem, we obtain $t = 1/2\tau$, so that the wave breaks when $t = 1/2a$. Yet the solution (2.3.23) with $x \leq a$ and $t > 0$ predicts the earlier breaking time $t = \frac{1}{2}(a - \sqrt{a^2 - 1})$. This discrepancy may be resolved by observing that for this problem, (2.3.23) is valid only in the region $x \leq a + (1 - a^2)t$, $t > 0$. That is, the region is bounded on the right by the characteristic base curve $x = a + (1 - a^2)t$ with $t > 0$ that passes through the point $(x, t) = (a, 0)$. In that region the breaking time is easily found to be $t = 1/2a$. The point at which the wave breaks is $x = a/2 + 1/2a$.

In general, $u(x, 0)$ must be specified over the entire x-axis, whereas in the foregoing it is given only for $x \leq a$. The simplest approach is to set $u(x, 0) = 1 - a^2$ for $x > a$. The resulting initial value $f(x)$ is continuous for all x, but it is not differentiable at $x = a$. The solution $u(x, t)$ is given by (2.3.23) to the left of the characteristic $x = a + (1 - a^2)t$, whereas $u(x, t) = 1 - a^2$ to the right of the characteristic. The solution is continuous across the characteristic, but the first derivatives of u have jump discontinuities there. [Generalized solutions of (2.3.9) with jumps in the first derivatives are considered below.] The breaking time for this wave is again $t = 1/2a$. This result is valid for all a and not just for $a > 1$. However, the wave does not break unless a is positive.

3. The initial condition is

$$u(x, 0) = f(x) = \sin x. \tag{2.3.24}$$

The implicit form of the solution is given as

$$u = f(x - tu) = \sin(x - tu), \tag{2.3.25}$$

and the characteristic curves are $x - t \sin \tau - \tau = 0$. Using (2.3.17), we have

$$t = -1/\cos \tau, \tag{2.3.26}$$

so that the breaking time occurs when $t = 1$, since when

$$\tau = (2n + 1)\pi, \qquad n = 0, \pm 1, \pm 2, \ldots, \tag{2.3.27}$$

we have $\cos \tau = -1$. At all other values of τ, t either exceeds unity or is negative. When $t = 1$ the wave breaks at the points $x = (2n + 1)\pi$.

A qualitative picture of the wave motion is indicated in Figure 2.5, where the interval $0 \leq x \leq 2\pi$ is considered. The critical nature of the point $(x, u) = (\pi, 0)$ is apparent since the wave form rotates in a clockwise motion near that point.

In a generalization of the foregoing we replace (2.3.24) by $u(x, 0) = f(x) = a + b \sin x$, $a > b > 0$. Then $f(x)$ is positive for all x and each point on the initial wave form moves to the right. From (2.3.17) we find that $t = -1/b \cos \tau$, and t

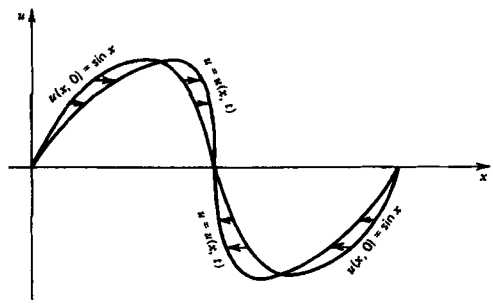

Figure 2.5 The wave motion.

has a minimum when τ is an odd multiple of π, as in (2.3.27). The breaking time is $t = 1/b$, and the wave breaks simultaneously at the points $x = (2n + 1)\pi + a/b$.

Generalized Solutions and Shock Waves

If the initial wave form $f(x)$ for (2.3.9) is a smooth function, the solution $u(x, t)$ remains smooth until the time when the wave breaks, if, indeed, it breaks at all. However, if $f(x)$ is continuously differentiable for $x \neq x_0$ but is merely continuous at $x = x_0$ and has a jump in the first derivative there, it follows from the results of Example 2.7 that $u = f(x_0)$ on the line $x = x_0 + tf(x_0)$. The jumps in the first derivatives of u across this line satisfy the equation

$$[u_t] + u[u_x] = [u_t] + f(x_0)[u_x] = 0, \tag{2.3.28}$$

which is derived exactly as was (2.2.19) for the linear case. The solution $u(x, t)$, which is continuous but has a jump in the first derivatives across a characteristic line, is called a *generalized solution*. It is shown in the exercises that the jumps in the first derivatives satisfy ordinary differential equations along the characteristic base curve.

For example, if $f(x) = x$ for $x \leq 1$ and $f(x) = 1$ for $x > 1$, we have $u(x, t) = x/(t + 1)$ for $x \leq t + 1$ and $u(x, t) = 1$ for $x > t + 1$. Then, $u = 1$ on the line $x = t + 1$, while $[u_t] = 1/(t + 1)$ and $[u_x] = -1/(t + 1)$ across the line $x = t + 1$, so that (2.3.28) is satisfied.

If $f(x)$ has a jump discontinuity at $x = x_0$, the concept of solution must be generalized even further. To do so, we express (2.3.9) in *conservation form*,

$$(u)_t + \left(u^2/2\right)_x = 0. \tag{2.3.29}$$

This is also referred to as a *conservation law*, for if we integrate with respect to x from a to b, we obtain

$$\frac{d}{dt}\int_a^b u\, dx = \frac{1}{2}\left(u(a, t)^2 - u(b, t)^2\right). \tag{2.3.30}$$

This can be interpreted to mean that the rate of change of $u(x,t)$ over the interval $[a,b]$ equals the difference in the flux $u(x,t)^2/2$ at $x=a$ and $x=b$. Thus, $u(x,t)$ is a conserved quantity.

We integrate (2.3.29) over an arbitrary closed and bounded region R in the (x,t)-plane with a piecewise smooth boundary curve S. Then, Green's theorem in the plane yields

$$\int_S u\,dx - \frac{1}{2}u^2\,dt = 0, \tag{2.3.31}$$

with integration taken in the positive direction. If $u(x,t)$ is continuously differentiable, this *integral relation* is equivalent to (2.3.9).

Suppose that a solution $u(x,t)$ of (2.3.9) has a jump discontinuity across the curve $x=x(s),\ t=t(s)$. Let the boundary curve S in (2.3.31) collapse onto an arbitrary portion of $[x(s),t(s)]$ and conclude that [see (2.2.23)]

$$x'(s)[u] - \frac{1}{2}t'(s)[u^2] = 0. \tag{2.3.32}$$

Represent the discontinuity curve as $x=x(t)$ (i.e., let $t=s$) and put $[u]=u_2-u_1$, where u_2 and u_1 are the limits of u as the curve is approached from the right and the left, respectively. Since $[u^2]=u_2^2-u_1^2$, (2.3.32) reduces to

$$x'(t) = \frac{1}{2}(u_1+u_2). \tag{2.3.33}$$

If the discontinuity in u originates at the point x_0 when $t=t_0$, we obtain the initial condition $x(t_0)=x_0$ for the equation (2.3.33).

Next, let $U(t)=(u_1+u_2)/2$ and write (2.3.33) as $x'(t)=U(t)$. Thus, the discontinuity in the solution moves with the velocity U, which is the average of the two (limiting) velocities u_1 and u_2 on either side of the discontinuity curve. Since $u(x,t)$ is constant on each characteristic, the discontinuity curve is intersected by the characteristics $x=tu_1+\tau_1$ and $x=tu_2+\tau_2$ with $\tau_1<\tau_2$. These lines can intersect only if $u_1>u_2$. Since U is the average of u_1 and u_2, we must have

$$u_1 > U > u_2. \tag{2.3.34}$$

Without the discontinuity curve, the solution would be multivalued and this curve prevents that from happening. In applications of (2.3.9) to fluid mechanics, as in Section 10.3, (2.3.34) is often referred to as the *entropy condition*.

The propagating jump or discontinuity in the solution is referred to as a *shock wave* and the velocity U of the jump is called the *shock velocity*. When a wave breaks at $x=x_0$ at the time t_0, the solution $u(x,t)$ becomes multivalued and loses its validity. By introducing a shock wave for which the shock curve is determined from (2.3.33) with $x(t_0)=x_0$, the solution can be extended beyond the breaking time. In the following examples we show that this can be done for a number of problems related to those considered in Example 2.8. In each problem, the initial wave form

is a piecewise linear bounded analog of a function considered in the example. This permits the construction of explicit solutions and shock waves for the initial value problems. Their behavior is qualitatively similar to that of the solutions of the more general problems. Additionally, a Maple procedure is presented in Section 2.5 that automates the construction of shock waves for problems with piecewise linear initial data.

Example 2.9. Compression and Expansion Waves. We consider waves of the form studied in part 1 of Example 2.8. But here, as in the examples that follow, the initial wave form $u(x,0) = f(x)$ is assumed to be bounded for all x. In all cases the initial values are given for the equation (2.3.9).

The initial wave form $f(x)$ yields the solution $u(x,t)$:

$$f(x) = \begin{cases} 1, & x \le 0, \\ 1 - x/a, & 0 < x < a, \\ 0, & x \ge a, \end{cases} \quad u(x,t) = \begin{cases} 1, & x \le t, \\ (a-x)/(a-t), & t < x < a, \\ 0, & x \ge a. \end{cases} \tag{2.3.35}$$

The distance between the two constant wave heights $u = 1$ and $u = 0$ decreases until the wave breaks at the time $t = a$. The wave form appears to undergo compression as t increases, so it is called a *compression wave.*

At the time $t = a$, all the characteristics that issue from the interval $0 \le x \le a$ intersect at the point $x = a$. Since $f(x)$ assumes all the values of x between zero and 1 in that interval, the solution $u(x,t)$, which is constant on each characteristic, assumes all the values between zero and 1 when $t = a$. That is, it is infinitely multivalued. Technically, however, $u(x,a)$ merely has a jump discontinuity at $x = a$. We have $u(x,a) = 1$, $x < a$ and $u(x,a) = 0$, $x > a$. Thus, the solution can be extended beyond $t = a$ by introducing a *shock wave.*

The wave $u(x,t)$ breaks at $x = a$ when $t = a$, with $u_1 = 1$ and $u_2 = 0$. Thus, $U = \frac{1}{2}$ and the entropy condition (2.3.34) is met. The equation (2.3.33) becomes $x'(t) = \frac{1}{2}$ with $x(a) = a$, with the solution $x(t) = (a+t)/2$, for $t \ge a$. Consequently, we obtain the *shock wave* for $t \ge a$,

$$u(x,t) = \begin{cases} 1, & x < (a+t)/2, \\ 0, & x > (a+t)/2. \end{cases} \tag{2.3.36}$$

The continuous (generalized) solution (2.3.35) remains valid until $t = a$. Then it is replaced by the shock wave (2.3.36). The shock speed is $U = \frac{1}{2}$, so that the shock wave form is a unit step that moves to the right with speed $\frac{1}{2}$.

The initial wave form $f(x)$ yields the solution $u(x,t)$,

$$f(x) = \begin{cases} 0, & x \le 0, \\ x/a, & 0 < x < a, \\ 1, & x \ge a, \end{cases} \quad u(x,t) = \begin{cases} 0, & x \le 0, \\ x/(a+t), & 0 < x < a+t, \\ 1, & x \ge a+t. \end{cases} \tag{2.3.37}$$

For this problem, as t increases, the distance between the two constant wave heights $u = 0$ and $u = 1$ increases. Therefore, this is called an *expansion wave*. This wave does not break and the solution is valid for all time.

It is of interest to consider the limits of the two foregoing problems as a tends to zero. In both cases this means that in the limit, the initial function $f(x)$ is piecewise constant and has a jump discontinuity at $x = 0$. This is often referred to as a *Riemann problem*. In the first case, $f(x) = 1$ for $x < 0$ and $f(x) = 0$ for $x > 0$. On solving this initial value problem by the method of characteristics, we find that the characteristics intersect at $x = 0$ as soon as t increases from zero. Thus the solution is double valued and the wave breaks immediately. In the second case, $f(x) = 0$ for $x < 0$ and $f(x) = 1$ for $x > 0$. Again solving by the method of characteristics, we find that $u(x,t) = 0$ for $x < 0$ and $u(x,t) = 1$ for $x > t$. In this problem the characteristics do not intersect so that the wave does not break. However, since none of the characteristics pass through the points (x,t) in the sector $0 \leq x \leq t$, the solution remains undetermined there.

As noted, in the first problem the wave breaks immediately at $x = 0$ when $t = 0$. Since $u_1 = 1$ and $u_2 = 0$, the *shock wave* is found to be

$$u(x,t) = \begin{cases} 1, & x < t/2, \\ 0, & x > t/2. \end{cases} \tag{2.3.38}$$

This also follows from (2.3.36) on taking the limit as a tends to zero.

For the second problem we obtain a continuous solution valid for all time on letting a tend to zero in the original solution (2.3.47). This yields

$$u(x,t) = \begin{cases} 0, & x \leq 0, \\ x/t, & 0 < x < t, \\ 1, & x \geq t. \end{cases} \tag{2.3.39}$$

It is possible to construct a shock wave for the second problem, defined as

$$u(x,t) = \begin{cases} 0, & x < t/2, \\ 1, & x > t/2. \end{cases} \tag{2.3.40}$$

However, this solution must be rejected because the shock wave construction fails to satisfy the entropy condition (2.3.34), since the wave speed on the left of the shock (i.e., $u_1 = 0$) is smaller than that on the right (i.e., $u_2 = 1$). This shows that the absence of the entropy condition (2.3.34) could lead to nonunique solutions of problems with discontinuous data. It is also shown in Section 10.3 that shock waves for the inviscid Burgers' equation that fail to satisfy the entropy condition (2.3.34) cannot be limits, as the viscosity tends to zero, of traveling wave solutions of the viscous Burgers' equation.

Example 2.10. Triangular Waves. If the initial condition of part 2 of Example 2.8 is replaced by

$$f(x) = \begin{cases} 1 - x^2, & |x| \leq 1, \\ 0, & |x| > 1, \end{cases} \tag{2.3.41}$$

the resulting wave breaks at the point $x = 1$ when $t = \frac{1}{2}$. The jump in the solution is zero at that time and it increases gradually. As a result, the shock wave for this problem has a variable speed, in contrast to the situation encountered in Example 2.9. Rather than study this problem, we consider a simpler qualitatively similar problem.

Let $u(x, 0) = f(x)$ be given as

$$f(x) = \begin{cases} 1, & |x| < 1, \\ 0, & |x| > 1. \end{cases} \tag{2.3.42}$$

The compressive part of this wave at $x = 1$ breaks immediately and requires the introduction of a shock wave with shock speed $U = \frac{1}{2}$. The expansive part of the wave at $x = -1$ can be dealt with by introducing a continuous wave, as was done Example 2.9.

Since the shock originates at $x = 1$, the equation of the shock front is $x = t/2 + 1$. The jump at $x = -1$ yields $u(x, t) = (x + 1)/t$ for $-1 < x < t - 1$ (see Exercise 2.3.18). The full solution to the problem is

$$u(x, t) = \begin{cases} 0, & x \leq -1, \\ (x + 1)/t, & -1 < x < t - 1, \\ 1, & t - 1 < x < \frac{1}{2}t + 1, \\ 0, & x > \frac{1}{2}t + 1. \end{cases} \tag{2.3.43}$$

This solution is valid only until the time $t = 4$, when the expansion wave $u = (x+1)/t$ overtakes the shock wave. This occurs at the point $x = 3$.

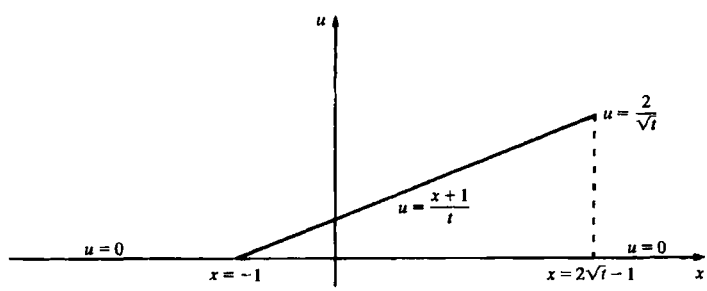

Figure 2.6 A triangular wave.

Within the sector $-1 < x < t - 1$, the characteristics are the lines $x + 1 = \lambda t$ that issue from the point $(x, t) = (-1, 0)$ and on which $u = (x + 1)/t = \lambda$. The characteristics that issue from the x-axis for $x > 3$ are all vertical lines on which

$u = 0$. These two sets of characteristics begin to intersect at $x = 3$ when $t = 4$ and the solution becomes double valued at that time. A new shock wave is required and the shock front is determined from $x'(t) = -\frac{1}{2}(x(t) + 1)/t$ with $x(4) = 3$. The solution is $x(t) = 2\sqrt{t} - 1$ and the shock speed is $U = (1/t)^{1/2}$.

Thus, the solution for $t \geq 4$ is

$$u(x,t) = \begin{cases} 0, & x \leq -1, \\ (x+1)/t, & -1 < x < 2\sqrt{t} - 1, \\ 0, & x > 2\sqrt{t} - 1. \end{cases} \tag{2.3.44}$$

The solution has the form of a triangular wave (see Figure 2.6). As t tends to infinity, the jump in the solution tends to zero. The wave form for the problem originally formulated in this example assumes a similar triangular shape as t gets large.

Example 2.11. N Waves. In view of the periodicity of the sinusoidal wave forms considered in Part 3 of Example 2.8, it is sufficient to study the problem over a single period. To simplify matters we consider an initial wave form that is constant outside a finite interval but has a quasisinusoidal shape within the interval.

Let $u(x,0) = f(x)$, with $f(x)$ defined as

$$f(x) = \begin{cases} -x, & -2 < x < -1, \\ x + 2, & |x| \leq 1, \\ 4 - x, & 1 < x < 2, \\ 2, & |x| \geq 2. \end{cases} \tag{2.3.45}$$

The wave form has two compressive parts whose slope is -1 and an expansive part with unit slope. Using the results of part 1 of Example 2.8, the solution is easily constructed. The wave $u(x,t)$ breaks at the time $t = 1$ and this occurs at the two points $x = 0$ and $x = 4$.

When $t \geq 1$, the solution obtained by the method of characteristics is given as

$$u(x,t) = \begin{cases} 2, & x < 2t - 2, \\ (x+2)/(t+1), & t - 1 < x < 3t + 1, \\ 2, & x > 2t + 2. \end{cases} \tag{2.3.46}$$

We observe that the characteristics intersect at $x = 0$ and $x = 4$ when $t = 1$ and two shock waves are required with shock fronts originating at these points. The equations for the shock speed (2.3.33) are identical for both shocks and are given as $x'(t) = 1 + (x+2)/(2+2t)$. With $x(1) = 0$, the solution is $x(t) = 2t - (2t+2)^{1/2}$, and for $x(1) = 4$, the solution is $x(t) = 2t + (2t+2)^{1/2}$.

Thus, the solution for $t \geq 1$ is

$$u(x,t) = \begin{cases} (x+2)/(t+1), & |x - 2t| < \sqrt{2t+2}, \\ 2, & |x - 2t| > \sqrt{2t+2}. \end{cases} \tag{2.3.47}$$

As shown in Figure 2.7, the wave form has the shape of an inverted N. Therefore it is called an *N wave*. As t tends to infinity, the magnitudes of the two jumps decrease on the order of $1/\sqrt{t}$.

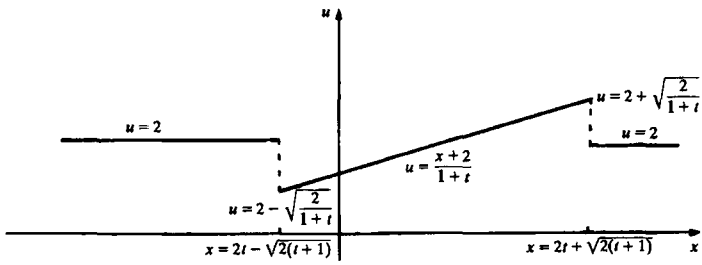

Figure 2.7 An N wave.

If $f(x) = a + b\sin(x)$ with $a > b > 0$, the solution takes the form of a periodic set of N waves after the breaking time. Whereas the foregoing N waves all travel to the right with increasing t, if $f(x) = \sin x$ the shock waves are all stationary because of the symmetry of the data with respect to the x-axis. However, they still assume the general form of N waves after the breaking time.

The equation

$$u_t(x,t) + c(u(x,t))u_x(x,t) = 0, \tag{2.3.48}$$

where $c(u(x,t))$ is an arbitrary function of $u(x,t)$, is the quasilinear equivalent of the linear equation (2.2.6) and is more representative of the type of first order quasilinear wave equations studied in applications. If we multiply across by $c'(u)$ in (2.3.48), we obtain $c_t + cc_x = 0$, and this is exactly (2.3.9) with u replaced by c. Thus, we do not expect the behavior of the solutions of (2.3.48) to be qualitatively much different from those of (2.3.9), as long as the solutions remain smooth. However, if the solution becomes multivalued and shock waves are required, the transformed version of (2.3.48) does not determine the correct shock fronts unless $c(u) = \hat{c} + \tilde{c}u$. To see this, let the conservation law associated with (2.3.48) be given as $u_t + (C(u))_x = 0$, with $C'(u) = c(u)$. Then the shock speed U for this equation is $U = [C(u)]/[u]$. For the transformed equation we have $U = \frac{1}{2}(c_2 + c_1)$, where c_2 and c_1 are the values of $c(u)$ on the right and left sides, respectively, of the shock. These shock velocities are unequal, in general, unless $c(u) = \hat{c} + \tilde{c}u$ with constant $\hat{c}$ and $\tilde{c}$, as is easily verified directly. Nevertheless, the simpler transformed version of (2.3.48) remains valid until the breaking time and can be used to determine that time.

Exercises 2.3

2.3.1. Using implicit differentiation, verify that $u = f(x - tu)$ is a solution of the wave equation (2.3.9).

2.3.2. Consider the damped quasilinear wave equation $u_t(x,t) + u(x,t)u_x(x,t) + cu(x,t) = 0$, where c is a positive constant. (a) Using the method of characteristics, construct an implicit solution of the initial value problem with $u(x,0) = f(x)$. Discuss the wave motion and the effect of the damping. (b) Determine the breaking time of the solution by finding the envelope of the characteristic curves and by using implicit differentiation. With τ as the parameter on the initial line, show that unless $f'(\tau) < -c$, no breaking occurs. Contrast this result with that for the undamped case discussed in the text.

2.3.3. Let $f(x) = ax$ in Exercise 2.3.2. Obtain an explicit representation of the solution and show directly that the wave does not break if $a \geq -c$. If $a < -c$, the wave breaks at the time $t = (1/c)\log[a/(a+c)]$.

2.3.4. Consider the one-dimensional form of Euler's equations for isentropic flow (8.5.42), (8.5.47), and (8.5.49) and assume that the pressure p is a constant. The equations reduce to $\rho_t + \rho u_x + u\rho_x = 0$, $u_t + uu_x = 0$. Let $u(x,0) = f(x)$ and $\rho(x,0) = g(x)$. By first solving the equation for u and then the equation for ρ, obtain the implicit solution $u = f(x - ut)$, $\quad \rho = g(x - ut)/[1 + tf'(x - ut)]$.

2.3.5. Obtain explicit expressions for the solution given in Exercise 2.3.4 if (a) $f(x) = x$, $g(x) = 1$; (b) $f(x) = x$, $g(x) = x$.

2.3.6. Solve the initial value problem for the equation $u_t(x,t) + cu_x(x,t) + u^2(x,t) = 0$, $u(x,0) = x$, where c is a constant.

2.3.7. Obtain the solution of the initial value problem $u_t + c(u)u_x = 0$, $u(x,0) = f(x)$, where c is a function of u, in the implicit form $u = f[x - tc(u)]$. Discuss the solution in the cases where $c'(u) > 0$ and $c'(u) < 0$.

2.3.8. Using the implicit form of the solution obtained in Exercise 2.3.7, determine the breaking time of the wave.

2.3.9. Solve the initial value problem $u_t(x,t) + u^2(x,t)u_x(x,t) = 0$, $u(x,0) = x$. Determine the breaking time of the solution and compare it with that obtained through the result of Exercise 2.3.8.

2.3.10. Solve the initial value problem $u_t(x,t) + u(x,t)u_x(x,t) = 0$, $u(x,0) = f(x)$, using the method of characteristics. (a) Using the parameters s and τ as defined in the text, show that the solution can be expressed as $x(s,\tau) = \frac{1}{2}[f(\tau) + \tau]e^s - \frac{1}{2}[f(\tau) - \tau]e^{-s}$, $t(s,\tau) = s$, $u(s,\tau) = \frac{1}{2}[f(\tau) + \tau]e^s + \frac{1}{2}[f(\tau) - \tau]e^{-s}$. (b) Obtain the solution in the form $u = u(x,t)$ when (i) $f(x) = 1$ and (ii) $f(x) = x$.

2.3.11. Consider the initial value problem $u_t(x,t) + u(x,t)u_x(x,t) = 0$, $u(x,0) = u_0 + \epsilon F(x)$, where u_0 is a constant, $0 < \epsilon \ll 1$, and $F(x)$ is uniformly bounded for all x. With $\epsilon = 0$, $u(x,t) = u_0$ represents a constant solution of the given equation. The

initial condition represents a small perturbation around the constant u_0, so we look for the solution in the form of a perturbation series around the constant state $u = u_0$. (a) Let $u(x,t) = u_0 + \epsilon u_1(x,t) + \epsilon^2 u_2(x,t) + \cdots$ and substitute this series formally into the equation for $u(x,t)$. Collecting like powers of ϵ and equating their coefficients to zero, obtain a recursive system of equations for the functions $u_i(x,t)$ $(i \geq 1)$. [Determine only the equations for $u_1(x,t)$ and $u_2(x,t)$.] Show that the appropriate initial conditions for the $u_i(x,t)$ are $u_1(x,0) = F(x)$, $u_i(x,0) = 0$, $i \geq 2$. (b) Show that the solutions for $u_1(x,t)$ and $u_2(x,t)$ are $u_1(x,t) = F(x - u_0 t)$, $u_2(x,t) = -tF'(x - u_0 t)F(x - u_0 t)$. (c) Since $u = u_0 + \epsilon u_1 + \epsilon^2 u_2 + \cdots$, show that when $-\epsilon t F' \approx 1$, the term $\epsilon^2 u_2$ is of the same order of magnitude as ϵu_1 and the terms in the series do not get smaller as was assumed. Compare the time t at which the series breaks down with the breaking time given in the text, as applied to the initial value problem. (d) Obtain the foregoing results for $u_1(x,t)$ and $u_2(x,t)$ by inserting the series for $u(x,t)$ directly into the implicit form of the solution as given in (2.3.15), expanding the function f in a series in ϵ, and comparing like powers of ϵ.

2.3.12. Show that the initial condition $u(x,t) = t^2/2$ on $x = t^3/6$ for the equation $u_t(x,t) + u(x,t)u_x(x,t) = t$ yields a characteristic initial value problem. Obtain the implicit solution $x = ut - t^3/3 + F(u - t^2/2)$, where $F(0) = 0$ but $F(z)$ is otherwise arbitrary.

2.3.13. Given the initial condition $u(x,t) = 1 - 4/t^2$ on $x = t + 1/4t$ for equation (2.3.9), express the initial curve in parametric form by setting $t = \tau$. Show that $\Delta(\tau) = 0$ but that the initial curve is not characteristic. Obtain two continuous solutions for this problem, neither of which is differentiable on the initial curve. *Hint*: See Example 2.8, Part 2.

2.3.14. Let $x = \frac{1}{2}\tau^2$, $t = \tau$, and $u = \tau$ be the initial curve for equation (2.3.9). Verify that $\Delta(\tau) = 0$ but that this is not a characteristic initial value problem. Obtain two solutions for this problem.

2.3.15. Solve the equation of Exercise 2.3.12 with the initial curve given as in Exercise 2.3.14. Show that the initial curve is not characteristic but that $\Delta(\tau) = 0$ there. Obtain an implicit solution of this problem.

2.3.16. Solve equation (2.3.9) with the initial value $u(0,t) = -t$. Express the initial curve as $x = 0$, $t = \tau$, and $u = -\tau$ and show that $\Delta(\tau) = 0$ only when $\tau = 0$. Use the method of characteristics or (2.3.15) to obtain the solution

$$u(x,t) = -\frac{1}{2}t + \frac{1}{2}(t^2+4x)^{1/2}, \quad t < 0, \quad u(x,t) = -\frac{1}{2}t - \frac{1}{2}(t^2+4x)^{1/2}, \quad t > 0.$$

Observe that $u(x,t)$ is not differentiable on $t^2 + 4x = 0$ and is not continuous at $t = 0$, $x > 0$.

2.3.17. Given the initial values for (2.3.9), $u(x,0) = f(x)$, $x < a$, and $u(x,0) = g(x)$, $x > a$, with $f(a) < g(a)$, obtain the continuous solution (in implicit form) $u(x,t) = f(x - tu)$, $x \leq tf(a) + a$, $u(x,t) = (x - a)/t$, $tf(a) + a < x < tg(a) + a$,

and $u(x,t) = g(x - tu)$, $x \geq tg(a) + a$. *Hint*: Consider the characteristics that issue from the point $(x,t) = (a,0)$.

2.3.18. Let $f(x) = A$ and $g(x) = B$ $(A < B)$ in Exercise 2.3.17. Obtain an explicit expression for the solution $u(x,t)$ for all x.

2.3.19. If the initial condition for (2.3.9) is $u(x,0) = A$, $x < a$ and $u(x,0) = B$, $x > a$, with $A > B$, show that the wave breaks immediately. Determine the form of the shock wave and show that it is stationary if $B = -A$.

2.3.20. Let $u(x,0) = A$, $x < -A$, $u(x,0) = -x$, $-A < x < A$, and $u(x,0) = -A$, $x > A$, with $A > 0$, be the initial values for (2.3.9). Solve for $u(x,t)$ and determine the breaking time. Construct the appropriate shock wave and show that it is stationary.

2.3.21. By introducing a shock wave when necessary, obtain a solution of the initial value problem for (2.3.9) with $u(x,0) = 1$, $x \leq -1$, $u(x,0) = -x$, $-1 < x < 0$, and $u(x,0) = 0$, $x \geq 0$, that is valid for all time.

2.3.22. Let $u(x,0) = 2$, $x < 0$, $u(x,0) = 1$, $0 < x < 1$ and $u(x,0) = 0$, $x > 1$ be the initial values for (2.3.9). Show that the shock that originates at $x = 0$ overtakes the shock from $x = 1$ at the time $t = 1$ at the point $x = 3/2$. Introduce a new shock wave and extend the validity of the solution for all time.

2.3.23. Using graphical means or otherwise, show that if $x = \pi$ and $t \geq 1$ in (2.3.26), there is one root τ_1 for equation (2.3.26) that lies in the interval $0 < \tau < \pi$. Verify that $\tau_0 = \pi$ and $\tau_2 = \pi - \tau_1$ are also roots of this equation. Since $u(x,t)$ is constant on each characteristic, show that $u(\pi,t)$ is triple valued for $t \geq 1$. Show that $u(x,t)$ is equal in magnitude but opposite in sign on the two characteristics that intersect the line $x = \pi$ from the left and from the right. Conclude that the corresponding shock wave is stationary.

2.3.24. Multiply (2.3.9) by u and obtain the conservation law $(u^2/2)_t + (u^3/3)_x = 0$. From the corresponding integral relation determine the equation for the shock front that takes the place of (2.3.33). Show that if $u(x,0) = 1$, $x < 0$ and $u(x,0) = 0$, $x > 0$, the shock speed $U = 2/3$. On the basis of the conservation law (2.3.29), however, the shock speed $U = 1/2$. Thus, even though both conservation laws are equivalent to (2.3.9) if the solutions are smooth, they give rise to different shock waves. The correct form of the conservation law must be known to guarantee the correct choice for the discontinuity solutions.

2.3.25. Verify that if $c(u)$ in (2.3.48) is a linear function of u, the shock velocity has the form given in the text.

2.3.26. Differentiate (2.3.9) with respect to x and show that on a characteristic curve along which $x'(t) = u$, u_x satisfies the equation $d(u_x)/dt + u_x^2 = 0$. Let $u_x = A$ at $t = 0$ and show that the solution of this ordinary differential equation becomes singular when $t = -1/A$. Consequently, any point on the initial wave form where

the slope A is negative is associated with a breaking time $t = -1/A$. Determine that this result agrees with that given in equation (2.3.17).

2.3.27. Generalize the method of characteristics to deal with the three-dimensional quasilinear equation $au_x + bu_y + cu_z = d$, where a, b, c, and d, are functions of x, y, z, and u. Show that the characteristic equations are $\{dx/ds = a,\ dy/ds = b,\ dz/ds = c,\ du/ds = d\}$ and discuss appropriate initial conditions.

2.3.28. Solve the initial value problem $u_x(x,y,z) + u_y(x,y,z) + u_z(x,y,z) = u^2(x,y,z)$, $u(x,y,0) = x + y$.

2.4 NONLINEAR FIRST ORDER PARTIAL DIFFERENTIAL EQUATIONS

In this section the method of characteristics is extended to deal with *nonlinear first order partial differential equations* in two independent variables. Although solutions of linear and quasilinear equations can be obtained as special cases of the general method developed below, these equations are of sufficient special interest to merit the separate discussion accorded them. A general discussion of the method of characteristics is followed by two examples. The eiconal equation of geometrical optics is analyzed in some detail.

Method of Characteristics

In its most general form the first order PDE can be written as

$$F(x,t,u(x,t),u_x(x,t),u_t(x,t)) = 0, \tag{2.4.1}$$

where $u_x(x,t) = \partial u(x,t)/\partial x$ and $u_t(x,t) = \partial u(x,t)/\partial t$. Let $p = u_x$ and $q = u_t$, and consider an integral surface $u = u(x,t)$ that satisfies (2.4.1). Its normal vector has the form $[u_x, u_t, -1] = [p, q, -1]$, and (2.4.1) requires that at the point (x,t,u), the components p and q of the normal vector satisfy the equation

$$F(x,t,u,p,q) = 0. \tag{2.4.2}$$

An integral surface $u = u(x,t)$ of (2.4.1) is not known a priori, so we think of (2.4.2) as characterizing a collection of admissible solutions or integral surfaces $u = u(x,t)$. That is, their normal vectors at the points (x,t,u) must satisfy (2.4.2). Each normal vector determines a tangent plane to the surface, and (2.4.2) is seen to generate a one-parameter family of tangent planes (to possible integral surfaces) at each point in (x,t,u)-space. We require that $F_p^2 + F_q^2 \neq 0$, so that (2.4.2) is not independent of p and q.

For example, if (2.4.1) has the form $u_x u_t - 1 = 0$, then $F = pq - 1 = 0$ and $q = 1/p$. As p ranges through all real values, $q = 1/p$ determines a one-parameter family of normal vectors $[p, q, -1] = [p, 1/p, -1]$ at each point (x,t,u). Similarly,

if $F = xu + tu_t^2 - (\sin x)u_x + 1 = 0$, then at (x, t, u) we have $F(x, t, u, p, q) = xu + tq^2 - (\sin x)p + 1 = 0$, which yields a set of values $q = q(p)$. Note that q cannot always be specified as a single-valued function of p. However, we shall always assume that one branch of the possible set of solutions $q = q(p)$ has been chosen.

In general, the (tangent) planes determined by p and q envelop a cone known as the *Monge cone*. Then if $u = u(x, t)$ is a solution of (2.4.1), it must be tangent to a Monge cone at each point (x, t, u) on the surface. The intersection of the Monge cones with the surface determines a field of directions on the surface known as the *characteristic directions*, as shown in Figure 2.8.

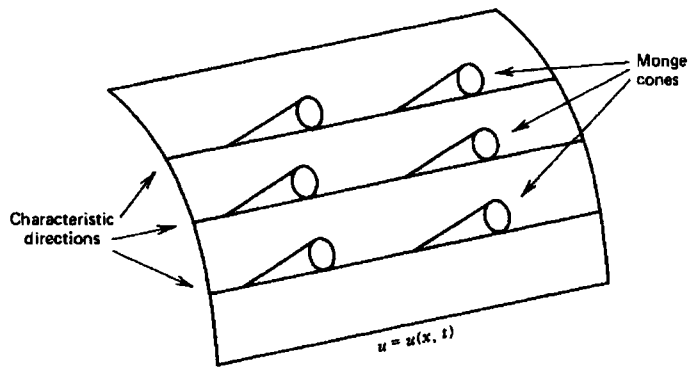

Figure 2.8 The Monge cones.

We find the characteristic directions as follows. The planes determined by the set of normal directions $[p, q, -1]$ at a point (x_0, t_0, u_0) satisfy the equations

$$H(x, t, u, p, q) = u - u_0 - p(x - x_0) - q(p)(t - t_0) = 0. \tag{2.4.3}$$

It is assumed in (2.4.3) that $F(x_0, t_0, u_0, p, q) = 0$ is solved for q as a function of p and one of the set of possible solutions expressed as $q = q(p)$ is selected. The envelope of the planes (2.4.3) (where p is a parameter) is determined by eliminating p from (2.4.3) and the equation

$$\frac{\partial H}{\partial p} = -(x - x_0) - \frac{dq}{dp}(t - t_0) = 0. \tag{2.4.4}$$

If we solve for $p = p(x, t)$ from (2.4.4) and substitute the result in (2.4.3), we obtain the equation of the Monge cone.

For example, if $F = pq - 1 = 0$, then $q = 1/p$ and $dq/dp = -1/p^2$. From (2.4.4) we obtain $1/p^2 = q^2 = (x - x_0)/(t - t_0)$. Substituting in (2.4.3) gives the equation of the Monge cone as $(u - u_0)^2 = 4(x - x_0)(t - t_0)$.

Continuing our discussion of the general case, we apply the chain rule to $F(x_0, t_0, u_0, p, q) = 0$, and obtain

$$\frac{dF}{dp} = \frac{\partial F}{\partial p} + \frac{dq}{dp}\frac{\partial F}{\partial p} = F_p + q'(p)F_q = 0. \tag{2.4.5}$$

Since $q'(p) = -F_p/F_q$, substitution in (2.4.4) yields

$$\frac{x - x_0}{F_p} = \frac{t - t_0}{F_q}. \tag{2.4.6}$$

Using (2.4.6) in (2.4.3) gives

$$\frac{u - u_0}{t - t_0} = p\,\frac{x - x_0}{t - t_0} + q = p\frac{F_p}{F_q} + q = \frac{pF_p + qF_q}{F_q}. \tag{2.4.7}$$

Combining (2.4.6) and (2.4.7) yields

$$\frac{u - u_0}{pF_p + qF_q} = \frac{x - x_0}{F_p} = \frac{t - t_0}{F_q}. \tag{2.4.8}$$

The (x, t, u) are the running variables in the tangent planes (2.4.3) so that the denominators in (2.4.8) are all constants evaluated at (x_0, t_0, u_0) on the integral surface. As a result, a fixed direction is determined by (2.4.8) on each tangent plane (for particular values of p and q). The direction is given by the vector $[F_p, F_q, pF_p + qF_q]$. As p and q range through all their values, these directions, known as the *characteristic directions*, determine the family of lines (2.4.8) that generate the Monge cone at (x_0, t_0, u_0).

We now construct equations for curves $x = x(s)$, $t = t(s)$, and $u = u(s)$ that have a characteristic direction at each point (i.e., the direction of the vector $[F_p, F_q, pF_p + qF_q]$). The ODEs for these curves are given as

$$\frac{dx}{ds} = F_p, \quad \frac{dt}{ds} = F_q, \quad \frac{du}{ds} = pF_p + qF_q. \tag{2.4.9}$$

Since p and q can vary from point to point, the curves determined from (2.4.9) must be chosen such that they lie on a single surface $u = u(x, t)$. To achieve this result, we assume that a surface $u = u(x, t)$ is given, so that p and q (i.e., u_x and u_t) are known. Then we determine the values that p and q must have along curves $x = x(s)$, $t = t(s)$, $u = u(x(s), t(s))$ that lie on that surface. Since we must also have $p = p(x(s), t(s))$ and $q = q(x(s), t(s))$ on that surface, we obtain

$$\frac{dp}{ds} = p_x F_p + p_t F_q, \quad \frac{dq}{ds} = q_x F_p + q_t F_q, \tag{2.4.10}$$

on using (2.4.9).

Further, if $u = u(x,t)$ is a solution of (2.4.1), we have $F(x,t,u,p,q) = 0$, which implies that

$$\frac{dF}{dx} = F_x + F_u p + F_p p_x + F_q q_x = 0, \quad \frac{dF}{dt} = F_t + F_u q + F_p p_t + F_q q_t = 0. \quad (2.4.11)$$

But $u_x = p$ and $u_t = q$ imply that $u_{xt} = p_t = u_{tx} = q_x$. Substituting $p_t = q_x$ into (2.4.11) and inserting the result in (2.4.10) yields

$$\frac{dp}{ds} = -F_x - F_u p, \quad \frac{dq}{ds} = -F_t - F_u q. \quad (2.4.12)$$

The five equations (2.4.9) and (2.4.12) now constitute a completely self-contained system of equations for the functions $x(s), t(s), u(s), p(s)$, and $q(s)$. [The integral surface $u = u(x,t)$ no longer needs to be given a priori in order to specify the values of p and q.] They are known as the *characteristic equations* for the differential equation (2.4.1).

The *initial value problem* for (2.4.1) requires that the integral surface $u = u(x,t)$ contain a curve C, the *initial curve*, which we give parametrically as $x = x(\tau)$, $t = t(\tau)$, $u = u(\tau)$. In terms of the functions $x(s)$, $t(s)$, and $u(s)$, these values are taken to correspond to initial values given at $s = 0$, as was done in previous sections. However, in contrast to the situation for linear and quasilinear equations, we must now also specify initial values for $p(s)$ and $q(s)$ if we hope to obtain a unique solution of the characteristic equations and, consequently, of the initial value problem for (2.4.1). The initial values $p(\tau)$ and $q(\tau)$ on the curve C cannot be arbitrary, since p and q must be components of the normal vector to the integral surface $u = u(x,t)$.

If we let $p = p(\tau)$ and $q = q(\tau)$ be the initial values of p and q on the curve C, these values must be determined from the equation

$$F(x(\tau), t(\tau), u(\tau), p(\tau), q(\tau)) = 0 \quad (2.4.13)$$

and the *strip condition*

$$\frac{du(\tau)}{d\tau} = p(\tau)\frac{dx(\tau)}{d\tau} + q(\tau)\frac{dt(\tau)}{d\tau}. \quad (2.4.14)$$

These equations follow from the fact that $p(\tau)$ and $q(\tau)$ must be components of the normal vectors to the surface $u = u(x,t)$ evaluated along the initial curve C. The data $p(\tau)$ and $q(\tau)$ are determined by solving (2.4.13) and (2.4.14) simultaneously. There may be more than one set of solutions for $p(\tau)$ and $q(\tau)$. However, in each problem we select a specific set of initial values $x(\tau)$, $t(\tau)$, $u(\tau)$, $p(\tau)$, and $q(\tau)$.

Given the characteristic equations and initial conditions for x, t, u, p, and q, together with the requirement (see the Appendix)

$$\frac{dx}{d\tau}F_q - \frac{dt}{d\tau}F_p \neq 0 \quad (2.4.15)$$

on the initial curve C, we can obtain a unique solution of the initial value problem for (2.4.1) in a neighborhood of the initial curve C. We note that the assignment of x, t, and u together with p and q on the curve C means that we are specifying an initial strip, that is, a space curve $(x(\tau),\ t(\tau),\ u(\tau))$ together with a family of tangent planes at each point having a normal vector with components $[p(\tau),\ q(\tau),\ -1]$, as shown in Figure 2.9. This explains why (2.4.14) is called the *strip condition.*

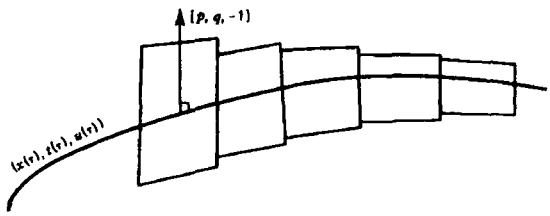

Figure 2.9 The initial strip.

The solutions $x(s,\tau)$, $t(s,\tau)$, and $u(s,\tau)$ of the characteristic equations for fixed τ determine a space curve. The functions $p(s,\tau)$ and $q(s,\tau)$ determine a tangent plane with normal vector $[p,\ q,\ -1]$ at each point of the space curve. The combination of a curve and its tangent planes is called a *characteristic strip.* The integral surface $u = u(x,t)$ is constructed by piecing together these characteristic strips to form a smooth surface. Its equation is found by solving for s and τ as functions of x and t from $x = x(s,\tau)$ and $t = t(s,\tau)$ and inserting the results in $u = u(s,\tau)$.

Example 2.12. A Nonlinear Wave Equation. We consider the equation

$$F(x,t,u(x,t),u_x(x,t),u_t(x,t)) = u_t(x,t) + u_x^2(x,t) = 0, \tag{2.4.16}$$

with the initial condition

$$u(x,0) = ax, \tag{2.4.17}$$

where a is a constant. The equation (2.4.16) is of the general form $u_t + c(u,u_x)u_x = 0$, of which two special cases, c = constant and $c = c(u)$, were considered in the previous sections. We might expect that (2.4.16) has some features in common with the linear and quasilinear wave equations considered earlier and refer to (2.4.16) as a *wave equation.* The nonlinearity of the x-derivative term in (2.4.16) implies, as we demonstrate, that the solution $u = u(x,t)$ is not constant along the characteristic curves. However, the velocity of specific points x on the wave is related to the nature of the wave form at those points, as was the case for the quasilinear equation (2.3.9).

To solve this initial value problem, we parameterize the initial curve as

$$x = \tau, \qquad t = 0, \qquad u = a\tau. \tag{2.4.18}$$

From (2.4.13)–(2.4.14) we obtain

$$q(\tau) + p(\tau)^2 = 0, \qquad a = p(\tau), \tag{2.4.19}$$

which yields as the initial data for p and q,

$$p(\tau) = a, \qquad q(\tau) = -a^2. \tag{2.4.20}$$

Since $x'(\tau)F_q - t'(\tau)F_p = 1$, condition (2.4.15) is also satisfied.

The characteristic equations for this problem are found to be

$$x'(s) = 2p, \quad t'(s) = 1, \quad u'(s) = q + 2p^2, \quad p'(s) = 0, \quad q'(s) = 0. \tag{2.4.21}$$

Using the data (2.4.18) and (2.4.20), we obtain $x(s,\tau) = 2as + \tau$, $t(s,\tau) = s$, $u(s,\tau) = a^2 s + a\tau$, $p(s,\tau) = a$, $q(s,\tau) = -a^2$. Solving for $s = s(x,t)$, $\tau = \tau(x,t)$ from the first two equations above gives $s = t$, $\tau = x - 2at$, and the solution of (2.4.16)–(2.4.17) becomes

$$u(x,t) = a(x - at). \tag{2.4.22}$$

The solution $u(x,t) = a(x - at)$ represents a plane wave moving with velocity a. If $a > 0$, it moves to the right, whereas if $a < 0$, it moves to the left.

Although the solution $u = a(x - at)$ behaves like the solution of a linear unidirectional wave equation with velocity a, the wave $u = a(x - at)$ is not constant on the characteristic base curves, $x - 2at = \tau$. Points on these curves travel with velocity $2a$, and points on the wave $u = a(x - at)$ have velocity a.

Replacing the parameter s by t in (2.4.21) yields $dx/dt = 2p$, $du/dt = q + 2p^2$, $dp/dt = 0$, $dq/dt = 0$. Since p = constant and $p = u_x$, we conclude that points x move with a velocity equal to twice the slope of the wave form $u = u(x,t)$. However, the fact that $u \neq$ constant on the characteristics implies that this velocity is not that of points on the wave moving parallel to the x-axis. This is consistent with the foregoing results. It does indicate that if the initial wave form $u(x,0) = f(x)$ has variable slope, points x where $f(x)$ has positive slope move to the right, whereas points where $f(x)$ has negative slope move to the left. Also, the greater the magnitude of the slope, the greater the velocity. This process continues throughout the wave motion and can lead to the breakdown of the solution even for smooth initial wave forms. This difficulty did not arise for the problem (2.4.16)–(2.4.17) since the initial slope is constant.

In the exercises another initial value problem for (2.4.16) is considered, where the problem of the breaking of the wave $u = u(x,t)$ does occur.

Geometrical Optics: The Eiconal Equation

The *eiconal equation of geometrical optics* in two space dimensions has the form

$$u_x^2(x,y) + u_y^2(x,y) = n^2, \tag{2.4.23}$$

and we shall assume that n = constant. We present a derivation of the eiconal equation in the exercises that exhibits its connection with wave propagation problems. Here we replace the variable t of (2.4.1) by the variable y so that $q = u_y$, and y takes the place of t in (2.4.1)–(2.4.15).

To begin, we take as the initial condition for (2.4.23)

$$u(x,y)|_{x=y} = u(y,y) = ay, \tag{2.4.24}$$

where a = constant. Parametrically, this can be expressed as

$$x = \tau, \qquad y = \tau, \qquad u = a\tau. \tag{2.4.25}$$

The initial values $p(\tau)$ and $q(\tau)$ on the curve (2.4.25) are determined from (2.4.13)–(2.4.14) as

$$p^2 + q^2 - n^2 = 0, \qquad a = p(\tau) + q(\tau). \tag{2.4.26}$$

The first equation states that $[p/n,\ q/n]$ is a unit vector. Let θ = constant, with $0 \le \theta < 2\pi$, and set

$$p(\tau) = n\cos(\theta), \qquad q(\tau) = n\sin(\theta), \tag{2.4.27}$$

so that the first equation in (2.4.26) is satisfied. Then the second equation implies that

$$\cos(\theta) + \sin(\theta) = \sqrt{2}\sin\left(\theta + \frac{\pi}{4}\right) = \frac{a}{n}. \tag{2.4.28}$$

[We select one of the solutions of (2.4.28). For example, if $a/n = 1$, we can have $\theta = 0$ or $\theta = \pi/2$.] The condition (2.4.15) yields $2(\sin(\theta) - \cos(\theta)) = 0$. Since this expression vanishes at $\theta = \pi/4$ or $\theta = 5\pi/4$, we must exclude those values of θ and the related values of a; that is, $a = \pm n\sqrt{2}$.

The characteristic equations are

$$\frac{dx}{ds} = 2p, \quad \frac{dy}{ds} = 2q, \quad \frac{du}{ds} = 2p^2 + 2p^2, \quad \frac{dp}{ds} = 0, \quad \frac{dq}{ds} = 0. \tag{2.4.29}$$

Thus, p and q are constant on the characteristics, so that $p(s,\tau) = p(0,\tau) = n\cos(\theta)$, $q(s,\tau) = q(0,\tau) = n\sin(\theta)$, and we obtain the straight-line characteristics,

$$x(s,\tau) = 2ns\cos(\theta) + \tau,\ y(s,\tau) = 2ns\sin(\theta) + \tau,$$
$$u(s,\tau) = 2n^2 s + [\sin(\theta) + \cos(\theta)]n\tau, \tag{2.4.30}$$

where (2.4.28) was used. Solving for s and τ as functions of x and y from (2.4.30) yields

$$s = \frac{y - x}{2n(\sin(\theta) - \cos(\theta))}, \qquad \tau = \frac{x\sin(\theta) - y\cos(\theta)}{\sin(\theta) - \cos(\theta)}. \tag{2.4.31}$$

Inserting this into (2.4.30) gives the *plane wave* solution

$$u(x,y) = n(x\cos(\theta) + y\sin(\theta)). \tag{2.4.32}$$

It is readily verified that (2.4.32) satisfies the eiconal equation and the initial condition (2.4.24). A more general plane wave solution of the eiconal equation can immediately be constructed from (2.4.32) in the form

$$u(x,y) = u_0 + n[(x - x_0)\cos(\theta) + (y - y_0)\sin(\theta)], \tag{2.4.33}$$

where x_0, y_0, and u_0 are constants.

A second initial value problem for the eiconal equation, whose values are given at a point rather than on a curve, leads to a *singular solution* of (2.4.23) which is called a *cylindrical wave*. Let the initial point be $(x,y,u) = (x_0,y_0,u_0)$, so that the initial values are $x(\tau) = x_0$, $y(\tau) = y_0$, $u(\tau) = u_0$. The *strip condition* (2.4.14) is satisfied automatically for all p and q. Thus $p(\tau)$ and $q(\tau)$ are determined from the single equation

$$F(x_0,y_0,u_0,p(\tau),q(\tau)) = p(\tau)^2 + q(\tau)^2 - n^2 = 0. \tag{2.4.34}$$

Let $p(\tau) = n\cos(\tau)$, $q(\tau) = n\sin(\tau)$, and (2.4.34) is satisfied. Continuing as in the preceding initial value problem we again find that $p(s,\tau)$ and $q(s,\tau)$ are constant on the characteristics and

$$x(s,\tau) = 2ns\cos(\tau) + x_0,\ y(s,\tau) = 2ns\sin(\tau) + y_0,\ u(s,\tau) = 2n^2 s + u_0. \tag{2.4.35}$$

Squaring and summing the equations for x and y in (2.4.35) and solving for s readily yields, if we take the positive square root,

$$u(x,y) = u_0 + n\sqrt{(x - x_0)^2 + (y - y_0)^2} \tag{2.4.36}$$

as a *singular solution* of the eiconal equation. It clearly satisfies (2.4.23) everywhere except at the initial point (x_0,y_0), where the derivatives u_x and u_y are singular. In fact, (2.4.36) determines a cone with vertex at (x_0,y_0,u_0), which is identical with the Monge cone through that point.

In the context of *geometrical optics*, if $u = u(x, y)$ is a solution of the eiconal equation, the level curves $u(x, y)$ = constant represent (light) *wave fronts*, whereas the characteristic base curves $x = x(s, \tau)$ and $y = y(s, \tau)$, for fixed τ, represent *light rays*. Thus the level curves of the solutions (2.4.32) or (2.4.33) are straight lines. In terms of the physical three-dimensional problem for which we are considering two-dimensional cross sections, the straight lines correspond to planes. Hence they are called *plane waves*. Similarly, the level curves of the solution (2.4.36) are circles that in the three-dimensional problem correspond to cylinders and therefore yield *cylindrical waves*. The tangent vectors to the characteristic base curves (i.e., the light rays) are found from (2.4.29) to be $[dx/ds,\ dy/ds] = 2[p,\ q]$. These curves are straight lines since p and q are constant for fixed τ. Further, the normal vectors to the level curves $u(x, y)$ = constant are $\nabla u = [u_x,\ u_y] = [p,\ q]$. This shows that the light rays are normal to the wave fronts.

The *intensity* of light in geometrical optics is characterized by the convergence or divergence of the light rays. As the rays converge, the intensity increases, whereas the intensity decreases as the rays diverge. The singular point (x_0, y_0) in the initial value problem for cylindrical waves is known as a *focal point* or a *focus*, since all the rays are seen to converge at that point or diverge from it. It represents a point of high intensity of light. Another region of high intensity of light is given by curves that are envelopes of the characteristic base curves or the light rays. Such curves are known as *caustic curves*. The solutions of the eiconal equations break down at foci or caustic curves. They suggest that the intensity of light tends to infinity there. Other methods not based solely on the eiconal equation for the description of optical effects, such as *wave optics*, yield valid descriptions of the optics in such regions. They are discussed in Chapter 10.

To complete our discussion of the eiconal equation, we show how the singular solution (2.4.36) may be used to solve a general initial value problem for the eiconal equation. The method we use is related to the *complete integral method* for the solution of first order partial differential equations. A complete integral of the equation (2.4.1) is a solution of the form $M(x, t, u, a, b) = 0$, where a and b are arbitrary constants. In our problem the solution (2.4.36), in fact, contains three arbitrary constants, x_0, y_0, and u_0. We now indicate how (2.4.36) can be used to solve (2.4.23) with the initial values (2.4.25) and (2.4.27). For simplicity, we assume that the constants a and n are chosen such that $a = n$. This implies that we either have $\theta = 0$ or $\theta = \pi/2$ in (2.4.32), so that $u(x, y) = nx$ or $u(x, y) = ny$.

We set $x_0 = \tau$, $y_0 = \tau$, and $u_0 = n\tau$ in (2.4.36) so that $(x_0, y_0, u_0) = (\tau, \tau, n\tau)$ represents the initial curve (2.4.25). This yields the family of solutions

$$u(x, y) = n\tau + n\sqrt{(x - \tau)^2 + (y - \tau)^2}. \tag{2.4.37}$$

Expressing the solution in implicit form as

$$H(x, y, u, \tau) = u - n\tau - n\sqrt{(x - \tau)^2 + (y - \tau)^2} = 0, \tag{2.4.38}$$

and differentiating with respect to τ gives

$$\frac{\partial H}{\partial \tau} = -n + n\frac{(x-\tau)+(y-\tau)}{\sqrt{(x-\tau)^2+(y-\tau)^2}} = 0, \tag{2.4.39}$$

where (2.4.38)–(2.4.39) determines the envelope of the family of solutions (2.4.37). We easily conclude from (2.4.39) that $(x-\tau)(y-\tau) = 0$, so that either $\tau = x$ or $\tau = y$, and we obtain

$$H = \begin{cases} u - nx - n|y-x| = 0, & \tau = x, \\ u - ny - n|y-x| = 0, & \tau = y. \end{cases} \tag{2.4.40}$$

Two continuously differentiable solutions can be obtained from (2.4.40) in the form $u(x,y) = nx$ and $u(x,y) = ny$, which correspond to the values $\theta = 0$ and $\theta = \pi/2$ in (2.4.36), respectively. Both solutions clearly satisfy the initial condition $u = nx = ny$ on the line $x = y$.

The existence of two solutions to the initial value problem (2.4.23)–(2.4.24) merits discussion. It signifies that merely specifying u on a curve without also specifying p and q (i.e., an initial strip) does not yield a unique solution, in general. The eiconal equation, $F = p^2+q^2-n^2 = 0$ is a quadratic expression for which $q = \pm\sqrt{n^2-p^2}$, so that for every choice of p there are generally two values of q. This means that, in general, there are two differentiable solutions passing through a given initial curve. The importance of these two solutions in applications of the eiconal equation is indicated in the exercises and is emphasized in our discussion of the eiconal equation in Chapter 10.

Exercises 2.4

2.4.1. Solve the initial value problem $u_x^2(x,t)u_t(x,t) - 1 = 0$, $u(x,0) = x$.

2.4.2. Use the method of characteristics to solve $u_t(x,t) + u_x^2(x,t) = t$ with $u(x,0) = 0$.

2.4.3. Obtain two solutions of (2.4.16) if $u(x,x) = 1$. Verify that (2.4.15) is satisfied for both solutions.

2.4.4. Find two complex-valued solutions of $u_x(x,t)u_t(x,t) = 1$ if $u(x,x) = 1$.

2.4.5. Solve (2.4.16) if $u(0,t) = t$.

2.4.6. Solve the initial value problem $u_t(x,t)+u_x^2(x,t)+u(x,t) = 0$, $u(x,0) = x$, using the method of characteristics.

2.4.7. Consider the wave equation in two dimensions $v_{xx}(x,y,t) + v_{yy}(x,y,t) = [1/c^2(x,y)]v_{tt}(x,y,t)$. (a) Let $v(x,y,t) = V(x,y)e^{-i\omega t}$ ($i = \sqrt{-1}$) and then show that $V(x,y)$ satisfies the reduced wave equation $V_{xx}(x,y) + V_{yy}(x,y) =$

$[\omega^2/c^2(x,y)]V(x,y)$. (The constant ω is the *angular* frequency of the solution.) Since $c(x,y)$ represents the (local) speed of wave propagation (as we shall see later in the text), we introduce a constant reference speed c_0, say, the speed of light, and define the *index of refraction* n as $n(x,y) = c_0/c(x,y)$ and the *wave number* k as $k = \omega/c_0$. Then the *reduced wave equation* takes the form $V_{xx}(x,y) + V_{yy}(x,y) + k^2n^2(x,y)V(x,y) = 0$. (b) Assume that $k \gg 1$ and look for a solution of the reduced wave equation in the form $V(x,y) = A(x,y)e^{iku(x,y)}$. Show that on substituting this form in the reduced wave equation, collecting like powers of k, and equating the coefficients of k^2 and k to zero, we obtain the eiconal equation (2.4.23) for $u(x,y)$ and $2A_x(x,y)u_x(x,y) + 2A_y(x,y)u_y(x,y) + [u_{xx}(x,y) + u_{yy}(x,y)]A(x,y) = 0$. This is known as a *transport equation* for $A(x,y)$ (which represents an amplitude term). (c) Show that if $n(x,y)$ and $A(x,y)$ are constants and $u(x,y)$ is the plane wave solution (2.4.36), $V(x,y)$ is an exact solution of the reduced wave equation. Otherwise, the function $V(x,y) = A(x,y)e^{iku(x,y)}$ represents what is known as the *geometrical optics approximation* to the solution of the reduced wave equation. (This approximation is discussed further in Section 10.1.)

2.4.8. Let $u(x,t) = A(t)e^{u(x,t)}$ in equation (1.1.15) and solve the resulting equation for $A(t)$ and $u(x,t)$ by first taking $u(x,t)$ to be a solution of $u_t(x,t) + cu_x(x,t) - \frac{1}{2}Du_x^2(x,t) = 0$. Let $u(x,t) = ax + bt$ and obtain the family of solutions $u(x,t) = \tau x + \left(\frac{1}{2}D\tau^2 - c\tau\right)t$. Determine the envelope of this family and thereby obtain the singular solution $u(x,t) = -(x-ct)^2/2Dt$. If $u(x,t)$ is the singular solution, solve for $A(t)$ and show that the resulting solution for $v(x,t)$ is of the form (1.1.20).

2.4.9. Solve $u_t(x,t) = c^2u_x^2(x,t)$ if $u(0,t) = -t$.

2.4.10. Let $u(x,y) = \sqrt{x^2+y^2}$. Solve the transport equation for $A(x,y)$ given in Exercise 2.4.7, subject to the condition $A(x,y) = 1$ on the circle $x^2 + y^2 = 1$. Show that $A(x,y)$ is singular at the origin $(0,0)$, which represents a focal point for $u(x,y)$.

2.4.11. Consider the eiconal equation $u_x^2(x,y) + u_y^2(x,y) = n^2(x,y)$ with $n(x,y) = \begin{cases} n_1, & x < 0, \\ n_2, & x > 0, \end{cases}$ with $n_2 > n_1$ and constant n_1 and n_2. (a) Given the boundary condition $u(0,y) = n_1 y\cos(\theta)$ where θ = constant, solve for all possible $u(x,y)$ in the regions $x < 0$ and $x > 0$. (*Hint*: They will be plane wave solutions.) These solutions play a role in the problem of a plane interface between two media with differing indices of refraction. (b) Obtain expressions for the directions of the rays found in part (a), as given by ∇u and evaluated on the interface $x = 0$. These formulas represent what are known as *Snell's laws* of reflection and refraction.

2.4.12. The characteristic equation for the wave equation (1.2.35) is given as $u_t^2(x,t) - \gamma^2u_x^2(x,t) = 0$ (see Section 3.2). Use the method of characteristics to obtain the plane wave solutions $u(x,t) = x \pm \gamma t$ and the singular solution $u(x,t)$ = constant on the pair of straight lines $x^2 = \gamma^2t^2$.

2.4.13. Extend the results of the text for nonlinear equations to n dimensions. With the equation given as $F(x_1, x_2, \ldots, x_n, u, u_{x_1}, \ldots, u_{x_n}) = 0$, obtain the characteristic equations (with $p_i = u_{x_i}$) $dx_i/ds = F_{p_i}$, $du/ds = \sum_{i=1}^{n} p_i F_{p_i}$, $dp_i/ds = -(F_u p_i + F_{x_i})$, where $i = 1, 2, \ldots, n$. Also, obtain conditions equivalent to (2.4.13)–(2.4.14).

2.4.14. Obtain plane wave solutions and singular solutions equivalent to (2.4.33) and (2.4.36), respectively, for the eiconal equation in three dimensions, $u_x^2(x, y, z) + u_y^2(x, y, z) + u_z^2(x, y, z) = n$, n =constant.

2.4.15. Solve the initial value problem $u_t(x, t) + u_x^2(x, t) = 0$, $u(x, 0) = -x^2$. Show that the solution breaks down when $t = \frac{1}{4}$.

2.4.16. Differentiate (2.4.16) with respect to x and let $v = u_x$. Show that v satisfies an equation of the form (2.3.48) and interpret the results of Example 2.12 accordingly.

2.4.17. Differentiate $F(x(s), t(s), u(s), p(s), q(s))$ with respect to s, and use (2.4.9) and (2.4.12) to conclude that F is constant along the characteristics.

2.4.18. Use differentiation and the equations of Exercise 2.4.13 to show that $F(x_1(s), \ldots, x_n(s), u(s), p_1(s), \ldots, p_n(s))$ is constant along the characteristics.

2.5 MAPLE METHODS

Maple contains a number of built-in procedures for the exact and numerical solution of ordinary and partial differential equations. The basic tool for the solution of ordinary differential equations (ODEs) and systems of ODEs is the procedure *dsolve*. Additionally, there is a package of procedures *DEtools* that contains specialized procedures for the manipulation and solution of ODEs. The related procedure and package for the solution of partial differential equations (PDEs) is *pdsolve* and *PDEtools*, respectively. In view of the greatly increased complexity inherent in the solution of PDEs as compared to the solution of ODEs, the ODE procedures in Maple yield much more information about the exact and numerical solution of initial and boundary value problems for ODEs than the corresponding ones for PDEs. When it comes to exact solutions of PDEs, the built-in Maple procedures mostly obtain general results, if possible, and do not solve initial and initial and boundary value problems directly. The situation is improved when finding numerical solutions for certain types of PDEs. Consequently, we have written a number of Maple procedures that deal with the problems considered in this chapter. They are presented and discussed below. To a large extent, the procedures employ the method of characteristics to generate solutions of the problems considered.

Linear First Order Partial Differential Equations

We begin by considering the basic Maple procedures for the solution of differential equations and apply them to some of the linear equations treated in this chapter. An initial value problem for the linear PDE

$$PDE1 = x\,\frac{\partial v(x,t)}{\partial x} + t\,\frac{\partial v(x,t)}{\partial t} = cv(x,t), \tag{2.5.1}$$

was considered in Example 2.5. On applying the procedure *pdsolve*(*PDE*1), where *PDE*1 refers to the foregoing PDE, Maple obtains the general solution

$$v\,(x,t) = F\left(\frac{t}{x}\right)x^c \tag{2.5.2}$$

with $F(z)$ as an arbitrary function. With the initial condition $v(x,1) = f(x)$, the PDE was shown to have the solution

$$v\,(x,t) = f\left(\frac{x}{t}\right)t^c. \tag{2.5.3}$$

It can be reproduced from (2.5.2) if we put $F(z) = f(1/z)z^c$.

The procedure *charstrip* in the *PDEtools* package determines the system of characteristic equations for a first order PDE. Thus, $charstrip(PDE1, v(x,t))$ yields the system of characteristic equations (2.2.28). This system of ODEs can then be solved with the initial data (2.2.26) using *dsolve*. This gives (2.2.29), and proceeding as in Example 2.5 and using Maple's *solve* procedure yields (2.5.3).

We have constructed a Maple procedure *LinPDE*1($[a(x,t), b(x,t), c(x,t)v + d(x,t)], [x,t,v], [x,t_0,f(x)]$) that solves the following initial value problem:

$$a\,(x,t)\,\frac{\partial v(x,t)}{\partial x} + b\,(x,t)\,\frac{\partial v(x,t)}{\partial t} = c\,(x,t)\,v(x,t) + d\,(x,t)\,, \quad v(x,t_0) = f(x) \tag{2.5.4}$$

if the characteristic equations can be solved and an exact solution can be determined. The output of the procedure exhibits the characteristic equations and their initial conditions. The solution of these equations is expressed in parametric form, if it can be found. Finally, the solution $v = v(x,t)$ of the given initial value problem, if it can be found, is exhibited. On applying this procedure to the PDE (2.5.1) with the initial value $v(x,1) = f(x)$ in the form $LinPDE1([x,t,cv],[x,t,v],[x,1,f(x)])$, the solution (2.5.3) is obtained. If the initial time is $t = 0$, the procedure indicates that the characteristic condition equals zero and no solution is found. As indicated in Example 2.6, this is not a characteristic initial value problem unless $f(x)$ has a special form.

The procedure *LinPDE*1 handles only initial value problems for linear first order PDEs with initial values of the form $v(x,t_0) = f(x)$. For initial value problems with the initial curve given parametrically as $x = g(\tau)$, $t = h(\tau)$, $v = f(\tau)$, the procedure $NumLinPDE1([a(x,t), b(x,t), c(x,t)v + d(x,t)], [x,t,v], [g(\tau), h(\tau), f(\tau)], \tau,$ *explicit*) attempts to find an exact solution of the first order PDE in (2.2.1) with

the specified initial values, and exhibits the solution if it can find it. If the fifth argument of the procedure is *numeric* and a sixth argument $[x_0,\ t_0,\ v_0]$ is added, the procedure uses a numerical method to determine the characteristic through the point $(x_0,\ t_0,\ v_0)$. (In both cases, the method of characteristics is used.) The Maple output is a procedure that permits the evaluation of the solution values along the characteristic. This can be done as long as Maple's numeric ODE solver can find a solution.

We reconsider the PDE (2.5.1). If the initial condition is $v(x,1) = f(x)$, the procedure $NumLinPDE1([x,t,cv],[x,t,v],[\tau,1,f(\tau)],\tau,explicit)$ finds the solution (2.5.3). The third argument expresses the initial condition parametrically as $x = \tau,\ t = 1,\ v = f(\tau)$. If the initial condition is $v(x,0) = f(x)$, the parametric representation can be given as $x = \tau,\ t = 0,\ v = f(\tau)$. The characteristic condition in this case, as exhibited by the procedure, is zero, and no solution is found. However, this is not a characteristic initial value problem. But if we put $f(x) = x^c$, the output of the procedure states that this is a characteristic initial value problem with infinitely many solutions. It exhibits one of the solutions as being $v(x,t) = x^c$.

The initial condition $v(x, x-1) = f(x)$ for (2.5.1) cannot be dealt with by the procedure *LinPDE*1. If we represent the initial condition parametrically as $x = \tau,\ t = \tau - 1,\ v = f(\tau)$, the procedure *NumLinPDE*1 finds the solution

$$v\,(x,t) = f\left(\frac{x}{x-t}\right)(x-t)^c\,. \tag{2.5.5}$$

As shown in Example 2.6, the line $t = x/3$ is a characteristic base curve for (2.5.1). With the initial value $v(x, x/3) = 5x^c$, we have a characteristic initial value problem, whereas if the initial condition is $v(x, x/3) = \sin(x)$, there is no solution. On representing the base curve as $x = \tau,\ t = \tau/3$, *NumLinPDE*1 finds that $v(x,t) = 5\,x^c$ is one of infinitely many solutions in the first case, and that there is no solution in the second case.

The procedure $NumLinPDE1([1,\sin(x),v],[x,t,v],[\tau,0,\tau],\tau,explicit)$ determines the exact solution of the PDE

$$\frac{\partial v(x,t)}{\partial x} + \sin(x)\frac{\partial v(x,t)}{\partial t} = v(x,t), \tag{2.5.6}$$

with the initial condition $v(x,0) = x$, as

$$v\,(x,t) = \frac{\arccos\left(t + \cos\left(x\right)\right)e^x}{e^{\arccos(t+\cos(x))}}. \tag{2.5.7}$$

The procedure $NumLinPDE1([1,\sin(x),v],[x,t,v],[\tau,0,\tau],\tau,numeric,[1,0,1],$ $output = array([0,.5,1,1.2,2]))$ solves the characteristic equations for (2.5.6) with the initial condition $v(1,0) = 1$ numerically. [Note that this determines a single characteristic curve that passes through the point $(1,0,1)$ in the (x,t,v)-plane.] The output exhibits the solution values as an array. The values of s are given as a list

$[0, .5, 1, 1.2, 2]$ in the procedure and the result yields the values of x, t, v at these points.

$$\left[\begin{array}{c} [s, x(s), t(s), v(s)] \\ \left[\begin{array}{cccc} 0.0 & 1.0 & 0.0 & 1.0 \\ 0.50000 & 1.5000 & 0.46957 & 1.6487 \\ 1.0 & 2.0 & 0.95645 & 2.7183 \\ 1.2000 & 2.2000 & 1.1288 & 3.3201 \\ 2.0 & 3.0 & 1.5303 & 7.3891 \end{array}\right] \end{array}\right]. \tag{2.5.8}$$

The solution values for $v(x,t)$ agree with those determined from the exact solution (2.5.7).

Higher-dimensional linear first order PDEs can be solved by using the procedure *NumLinPDE1hd*. Exact and numerical solutions can be determined. As an example in three space dimensions and one time dimension, we have the PDE with constant coefficients

$$a\frac{\partial v}{\partial x} + b\frac{\partial v}{\partial y} + c\frac{\partial v}{\partial z} + d\frac{\partial v}{\partial t} = e\,v + f, \tag{2.5.9}$$

with $v = v(x,y,z,t)$ and the initial condition $v(x,y,z,0) = g(x,y,z)$. We invoke $NumLinPDE1hd([a,b,c,d], e, f, [x,y,z,t],\ v,\ [\lambda,\tau,\omega,0,g(\lambda,\tau,\omega)], [\lambda,\tau,\omega], explicit)$. The initial condition is expressed in parametric form. The solution is found to be

$$v(x,y,z,t) = \frac{f}{e}\left(\exp\left(\frac{et}{d}\right) - 1\right) + g\left(\frac{dx - at}{d}, \frac{dy - bt}{d}, \frac{dz - ct}{d}\right)\exp\left(\frac{et}{d}\right). \tag{2.5.10}$$

Equations with variable coefficients can also be treated and numerical solutions can be found.

Quasilinear First Order Partial Differential Equations

To solve a general quasilinear PDE for $u(x,t)$ with $u(x,t_0) = f(x)$ we use $NumQuasiPDE1(a(x,t,u), b(x,t,u), c(x,t,u), [\tau, t_0, f(\tau)], \tau, [x,t,u], explicit)$. The PDE to be solved is

$$a(x,t,u(x,t))\frac{\partial u(x,t)}{\partial x} + b(x,t,u(x,t))\frac{\partial u(x,t)}{\partial t} = c(x,t,u(x,t)). \tag{2.5.11}$$

The procedure attempts to find an explicit solution of the initial value problem, and this is clearly impossible in the general case. We specialize to the problem for the inviscid Burgers' equation (2.3.9) considered in Example 2.8. The initial condition is $u(x,0) = \alpha + \beta x$. We have $NumQuasiPDE1(u, 1, 0, [\tau, 0, \alpha + \beta\tau], \tau, [x,t,u], explicit)$. The initial curve is expressed in parametric form as

$x = \tau$, $t = 0$, $u = \alpha + \beta\tau$ and the solution is $u(x,t) = (\alpha + \beta x)/(1 + t\beta)$, in agreement with the result in Example 2.8.

We can solve the initial value problem for the PDE (2.3.48) with $u(x,0) = f(x)$ by using $NumQuasiPDE1(c(u), 1, 0, [\tau, 0, f(\tau)], \tau, [x, t, u], explicit)$. The output has the form $u = f\,(RootOf(-x + c(f(z))t + z))$, using the $RootOf$ procedure. On substituting $f(z) = u$ in the result, we obtain the implicit solution,

$$u = f(x - c(u)t). \tag{2.5.12}$$

This reduces to the implicit solution (2.3.15), if $c(u) = u$, in which case the PDE reduces to the inviscid Burgers' equation.

The procedure *ShockWave* constructs solutions and shock waves for the inviscid Burgers' equation (2.3.9) if the initial value is a piecewise linear function. We begin by considering the Riemann problems introduced in Example 2.9. If we invoke the procedure $ShockWave([[1, 0, \alpha]], [x, t], u)$, the second and third arguments specify the independent and dependent variables. The first argument is a list (of lists) that represents $[u|_{x<\alpha},\ u|_{x>\alpha}, \alpha]$, where the first two elements are the initial values of u for $x < \alpha$ and for $x > \alpha$ for a specified value of $x = \alpha$. The output exhibits the PDE (2.3.9) and the initial condition $u(x,0) = 1$, $x < \alpha$, $u(x,0) = 0$, $x > \alpha$. The statement *The solution or shock wave is valid for all time* is printed out. For this problem a shock wave is formed immediately at $x = \alpha$ when $t = 0$ and remains valid for all $t > 0$. The shock wave is represented as a piecewise-defined function

$$u\,(x,t) = \begin{cases} 1, & x \le t/2 + \alpha, \\ 0, & x > t/2 + \alpha. \end{cases} \tag{2.5.13}$$

The shock wave has a unit jump and travels to the right without change of shape with the shock speed $U = 1/2$. It reduces to the result obtained in Example 2.9 if $\alpha = 0$.

The procedure $ShockWave([[0, 1, \alpha]], [x, t], u)$ solves (2.3.9) with the initial condition $u(x,0) = 0$, $x < \alpha$, $u(x,0) = 1$, $x > \alpha$. For this problem, the solution is not a shock wave but a continuous (expansion wave) that is given as

$$u(x,t) = \begin{cases} 0, & x \le \alpha \\ (x - \alpha)/t, & x \le t + \alpha, \\ 1, & x > t + \alpha. \end{cases} \tag{2.5.14}$$

It reduces to the result given in Example 2.9 if $\alpha = 0$.

All the results found in Examples 2.10 and 2.11 can be reproduced by the $Shock\ Wave$ procedure, because the initial data are piecewise constant. As an example we consider the N-wave of Example 2.11. We use $ShockWave([[2, -x, -2], [-x, x + 2, -1], [x + 2, 4 - x, 1.], [4 - x, 2, 2]], [x, t], u)$. The output states that the solution is valid until $t = 1$ and that the wave breaks at $x = 0$ and $x = 4$. The solution is given as

$$u(x,t) = \begin{cases} 2, & x \le -2 + 2\,t, \\ x/(-1 + t), & x \le -1 + t, \\ (x + 2)/(1 + t), & x \le 3\,t + 1, \\ (-4 + x)/(-1 + t), & x \le 2 + 2\,t, \\ 2, & x > 2 + 2t. \end{cases} \tag{2.5.15}$$

On invoking the procedure $ShockWave(u(x,t),[x,t],u,t=1)$, where the solution (2.5.15) is entered as the first argument, together with a fourth argument $t=1$, (that gives the breaking time), the output is a shock wave (an N-wave) that remains valid for all $t \geq 1$. It is given as

$$u(x,t) = \begin{cases} 2, & x \leq 2t - \sqrt{2+2t}, \\ (x+2)/(1+t), & x \leq 2t + \sqrt{2+2t}, \\ 2, & x > 2t + \sqrt{2+2t}. \end{cases} \tag{2.5.16}$$

The solution agrees with the result found in Example 2.11.

As seen from the foregoing, the first argument in $ShockWave$ can either be a piecewise function, defined via the Maple command *piecewise*, or a list of lists that specifies the points where the piecewise function undergoes a change of definition. The definition to the right and left of each point must be specified, and any number of points may be considered. However, the function must be piecewise linear.

Nonlinear First Order Partial Differential Equations

The procedure $NumNonLinPDE1(F(x,t,u,p,q)=0,[x(\tau),t(\tau),u(\tau),p(\tau),q(\tau)],\tau,[x,t,u,p,q],explicit)$ solves the initial value problem for the PDE $F(x,t,u,p,q)=0$ [as defined in (2.4.2)] with the initial condition given parametrically as $x=x(\tau)$, $t=t(\tau)$, $u=u(\tau)$, $p=p(\tau)$, $q=q(\tau)$. As indicated in Section 2.4, it is necessary to select a specific $q=q(p)$ from the solution(s) of $F(x,t,u,p,q)=0$. To find a possible set of initial values for p and q in terms $F(x,t,u,p,q)=0$, we must first apply the procedure with only three terms $[x(\tau),t(\tau),u(\tau)]$ in the second argument. The output is a possible set of values $p(\tau)$, $q(\tau)$. On selecting one of these values, the full procedure can be used to determine an explicit solution if possible. A numerically determined characteristic can also be found, but we don't consider this here. Additionally, the procedure $NumNonLinPDE1hd$ that deals with nonlinear first order PDEs is not discussed here.

As a simple example we consider the initial value problem of Example 2.12. $NumNonLinPDE1(q+p^2=0,[\tau,0,a\tau],\tau,[x,t,u,p,q],explicit)$ yields the output $q(\tau)=-a^2$, $p(\tau)=a$ as the initial values for p and q. Then, we replace $[\tau,0,a\tau]$ in the foregoing procedure by $[\tau,0,a\tau,a,-a^2]$. This gives the solution $u(x,t)=ax-a^2t$, in agreement with (2.4.22).

Next, we consider the eiconal equation (2.4.23) and construct a plane wave solution. We invoke the procedure $NumNonLinPDE1(p^2+q^2-n^2=0,[\tau,\tau,n(\sin(\theta)+\cos(\theta))\tau],\tau,[x,y,u,p,q],explicit)$. The output yields two possible sets of values $p=n\cos(\theta)$, $q=n\sin(\theta)$ and $p=n\sin(\theta)$, $q=n\cos(\theta)$. We select the first set as in (2.4.27) and replace the second argument in the procedure by $[\tau,\tau,n(\sin(\theta)+\cos(\theta))\tau,n\cos(\theta),n\sin(\theta)]$. Maple's output is the plane wave $u(x,y)=n\,(x\cos(\theta)+y\sin(\theta))$.

Exercises 2.5

2.5.1. Consider the PDE (2.5.1). (a) Apply the Maple procedure *pdsolve* to obtain the general solution (2.5.2). (b) With the initial condition $v(x,1) = f(x)$, use the procedures $LinPDE1$ and $NumLinPDE1$ to get the solution (2.5.3). (c) Use *charstrip* of the $PDEtools$ package to determine the characteristic equations for (2.5.1) and use *dsolve* to obtain a solution of the system.

2.5.2. Let $f(x) = \cos\left(x^2\right)$ and $c = 2$ in (2.5.3). (a) Use the *animate* procedure from the *plots* package to animate the solution (2.5.3) for $-2 \leq x \leq 2$ and $1 \leq t \leq 2$. (b) Apply the *plot3d* procedure to plot (2.5.3) for the range of values give in part (a).

2.5.3. Find the solution (2.5.5) of (2.5.1) using $NumLinPDE1$.

2.5.4. Use $NumLinPDE1$ to determine that (2.5.1) with the initial condition $v(x, x/3) = 5x^c$ yields a characteristic initial value problem for (2.5.1) with infinitely many solutions. However, if the initial condition is $v(x, x/3) = \sin(x)$, there is no solution.

2.5.5. Reproduce the table (2.5.8).

2.5.6. Solve the initial value problem (2.2.31) using the procedures $LinPDE1$ and $NumLinPDE1$.

2.5.7. Use $LinPDE1$ and $NumLinPDE1$ to obtain the solution (2.5.7).

2.5.8. Apply *pdsolve* to obtain a general solution of (2.2.38).

2.5.9. Consider the initial value problem $xu_x(x,y) + (x+y)u_y(x,y) = cu(x,y)$, $u(1,y) = f(y)$ with a constant c. (a). Find a general solution of the PDE using *pdsolve*. (b). Use $NumLinPDE1$ to solve the initial value problem.

2.5.10. Apply $NumLinPDE1hd$ to solve the initial value problem for (2.5.9) with the initial condition $v(x,y,z,0) = g(x,y,z)$. Obtain a general solution of (2.5.9) with *pdsolve*.

2.5.11. Use $NumLinPDE1hd$ to solve the initial value problem $zu_x(x,y,z) + yu_y(x,y,z) + u_z(x,y,z) = 0$, $u(x,y,0) = f(x,y)$.

2.5.12. Obtain the solution found at the end of Example 2.8, Part 1, by invoking $NumQuasiPDE1$.

2.5.13. Determine the implicit solution (2.5.12) by proceeding as indicated in the text.

2.5.14. Solve the initial value problem $u_t(x,t) + u(x,t)u_x(x,t) = t$, $u(x,0) = ax$, where a is a constant, using $NumQuasiPDE1$. Determine the breaking time for the solution. Show that the wave does not break if $a > 0$.

2.5.15. Consider the problem given in Example 2.8, part 2. (a) Invoke the procedure $NumQuasiPDE1$ and show that *allvalues* applied to the output yields two solutions. Determine that (2.3.23) is the correct solution. (b) The output of the procedure yields a parametric representation of the solution in the form $\{x(s) =$

$(1-\tau^2)s+\tau,\ t(s)=s,\ u(s)=1-\tau^2\}$. Use this parametric representation to animate the solution using $plots[animate]$ with $s=0..2$. We note that $s=t$ and the plot shows that the solution rapidly becomes double valued.

2.5.16. Use $NumQuasiPDE1$ to obtain an implicit solution of the initial value problem $(x+u(x,y))u_x(x,y)+(y+u(x,y))u_y(x,y)=0,\ u(x,0)=f(x)$. Find an explicit solution of the problem if $f(x)=x^2$ with $x>0$.

2.5.17. The Maple procedure $NumQuasiPDE1hd([1,1,1],u^2,[\tau_1,\tau_2,0,\tau_1+\tau_2],[\tau_1,\tau_2],[x,y,z,u],explicit)$ solves the initial value problem $u_x(x,y,z)+u_y(x,y,z)+u_z(x,y,z)=u^2(x,y,z),\ u(x,y,0)=x+y$. Use it to obtain the solution of the problem.

2.5.18. Use the procedure $ShockWave$ to obtain (2.5.13).

2.5.19. Obtain (2.5.14) by using the $ShockWave$ procedure.

2.5.20. Reproduce the results of Example 2.10 using the $ShockWave$ procedure.

2.5.21. Reproduce the results of Example 2.11 using the $ShockWave$ procedure.

2.5.22. Reproduce the results of Example 2.12 using the $NumNonLinPDE1$ procedure.

2.5.23. Reproduce the results of the subsection of Section 2.4 on the eiconal equation using the $NumNonLinPDE1$ procedure.

Appendix: Envelopes of Curves and Surfaces

Let $F(x,t,\tau)=0$ be a one-parameter family of curves in the (x,t)-plane with parameter τ. The envelope of the family of curves is determined from the equations

$$F(x,t,\tau)=0,\qquad \frac{\partial F}{\partial \tau}(x,t,\tau)=0.$$

Solving for $\tau=\tau(x,t)$ from the second equation and substituting into the first equation gives the *envelope* $F[x,t,\tau(x,t)]=0$. It is generally assumed that each curve in the family intersects the envelope tangentially at least once. Otherwise, the resulting curve is not referred to as an envelope. If F is a function of more than two variables and depends on a parameter τ, the foregoing procedure for constructing the envelope also applies.

As an example, let a curve be given as $t=f(x)$. If $[\tau,f(\tau)]$ is a fixed point on that curve, the equation of the tangent line to the curve at that point is $F(x,t,\tau)=t-f(\tau)-(x-\tau)f'(\tau)=0$. As the parameter τ varies, we have a one-parameter family of tangent lines to the curve. Now, $\partial F/\partial\tau=-(x-\tau)f''(\tau)=0$ implies that if $f''(\tau)\neq 0$ for all τ, then $\tau=x$. Inserting this in the equation $F(x,t,\tau)=0$ yields the envelope $F(x,t,\tau)|_{\tau=x}=t-f(x)=0$. Thus the given curve $t=f(x)$ is the envelope of its family of tangent lines. The family of curves is tangent to the envelope at, at least, one point.

If we consider the family of curves $t = |x| + \tau$, and express them as $F(x,t,\tau) = (t-\tau)^2 - x^2 = 0$, we find that $F_\tau = -2(t-\tau) = 0$. This yields $\tau = t$ and we conclude that $x = 0$. However, the t-axis is not an envelope of this family of broken lines, since none of them is tangent to the t-axis. It does represent the locus of singular (nondifferentiable) points of this family.

As another example we consider the pencil of straight lines through the origin $F = t - \tau x = 0$. Combined with the equation $F_\tau = 0$, we obtain $x = t = 0$. Although the origin is not an envelope for this family of lines, it is the point where they all intersect.

As an example of a three-dimensional problem, we consider the family of spheres $F(x,y,z,\tau) = (x-\tau)^2 + y^2 + z^2 - a^2 = 0$, whose radii equal a and whose centers are the points $(\tau, 0, 0)$, where τ is a parameter. Since $\partial F/\partial\tau = 2(x-\tau) = 0$, we find that $\tau = x$. The envelope of the family of spheres is, therefore, given by $F[x,y,z,\tau(x)] = y^2 + z^2 - a^2 = 0$. This is the equation of a circular cylinder of radius a whose axis coincides with the x-axis.

If the family of curves is given in parametric form as $x = x(s,\tau)$, $t = t(s,\tau)$, we let s be the running parameter along a single curve and τ be the parameter specifying a member of the family. Geometrically, the curves of the family are tangent to the envelope at their points of intersection, so we may characterize τ as a parameter along the envelope curve (each value of τ specifies a point of intersection on that curve). Now $[\partial x/\partial s,\ \partial t/\partial s]$ is a tangent vector of a member of the family for each fixed τ, whereas $[\partial x/\partial\tau,\ -\partial x/\partial\tau]$ is a normal vector of the envelope curve when s is evaluated at the envelope. Since the tangent vector of a member of the family of curves is orthogonal to the normal vector of the envelope at the point of intersection, we have for their dot product

$$\Delta(s,\tau) = \left[\frac{\partial x}{\partial s}, \frac{\partial t}{\partial s}\right] \cdot \left[\frac{\partial t}{\partial \tau}, -\frac{\partial x}{\partial \tau}\right] = x_s t_\tau - t_s x_\tau = 0.$$

This equation describes the envelope of the family.

The envelope equation is to be compared with the conditions (2.3.4) and (2.4.15). Noting (2.3.7) and (2.4.9) we see that both of the conditions (2.3.4) and (2.4.15) are equivalent to $\Delta(0,\tau) = x_s t_\tau - t_s x_\tau|_{s=0} \neq 0$. In terms of the initial value problem for the equations (2.3.1) and (2.4.1), the condition $\Delta(0,\tau) \neq 0$ requires that the initial curve C be neither a characteristic curve nor an envelope of characteristic curves. Otherwise, the solution may not exist, or if it does exist, need not be unique. Additionally, even if the initial curve satisfies the condition (2.3.4) or (2.4.15), when the family of characteristics forms an envelope, the solution breaks down, as is seen for the quasilinear equation (2.3.9) and is shown for the eiconal equation in Chapter 10.

For example, the circle $x^2 + t^2 = a^2$ is given parametrically as $x = a\cos(\tau)$, $t = a\sin(\tau)$. Its tangent lines are $x(s,\tau) = a\cos(\tau) - as\sin(\tau)$, $t(s,\tau) = a\sin(\tau) + as\cos(\tau)$, as is readily seen. Here s is the running parameter along a tangent line and τ characterizes a particular tangent line. We have $\Delta(s,\tau) = x_s t_\tau - t_s x_\tau = a^2 s$, so that $\Delta(s,\tau) = 0$ implies that $s = 0$ is the envelope. Then $x(0,\tau) = a\cos(\tau)$, $t(0,\tau) = a\sin(\tau)$ yields the circle $x^2 + t^2 = a^2$, as was expected.

CHAPTER 3

CLASSIFICATION OF EQUATIONS AND CHARACTERISTICS

The *telegrapher's* and *wave equations*, the *diffusion equation*, and *Laplace's equation*, derived in Chapter 1, are prototypes of the three basic types of partial differential equations encountered most often in applications. They are equations of *hyperbolic*, *parabolic*, and *elliptic* type, respectively. The classification of equations and systems of equations, in general, into these three (or possibly other) types is considered in this chapter.

Furthermore, formulations of initial and boundary value problems appropriate to equations of different types are given. Additionally, characteristics that were shown to play an important role in the theory of first order partial differential equations are reintroduced from a somewhat different point of view. Their significance in the theory of hyperbolic partial differential equations is demonstrated. Further, certain basic concepts relating to second order linear equations with constant coefficients are introduced and discussed. Finally, the important concept of adjoint differential operators is presented.

Partial Differential Equations of Appplied Mathematics, Third Edition. By Erich Zauderer

3.1 LINEAR SECOND ORDER PARTIAL DIFFERENTIAL EQUATIONS

We begin by considering the general second order linear partial differential equation in two independent variables,

$$A(x,y)u_{xx}(x,y) + 2B(x,y)u_{xy}(x,y) + C(x,y)u_{yy}(x,y) + D(x,y)u_x(x,y) + E(x,y)u_y(x,y) + F(x,y)u(x,y) = G(x,y), \qquad (3.1.1)$$

where the terms A, $B, \ldots,$ G are given real-valued functions. Special cases of (3.1.1) were derived in Chapter 1 (with y replaced by t in some instances). It will now be shown that the *principal part* of (3.1.1), that is, the terms containing second derivatives of u, can be transformed by a change of independent variables into a form similar to that of the wave, diffusion, and Laplace equations considered in Chapter 1. Consequently, these equations are, in fact, prototypes of second order linear equations of the general form (3.1.1). The transformed versions of these equations are referred to as *canonical forms*.

Without loss of generality, it may be assumed that $A(x,y) \neq 0$ in some region R and we divide by $A(x,y)$ in (3.1.1). Then, with $\partial_x = \partial/\partial x$ and $\partial_y = \partial/\partial y$, we express the principal part of the differential operator in (3.1.1) as follows:

$$\partial_x^2 + (2B/A)\partial_x\partial_y + (C/A)\partial_y^2 = (\partial_x - \omega^+\partial_y)(\partial_x - \omega^-\partial_y) + (\omega_x^- - \omega^+\omega_y^-)\partial_y, \qquad (3.1.2)$$

where $\omega^+(x,y)$ and $\omega^-(x,y)$ are defined as [on comparing both sides of (3.1.2)]

$$\omega^+(x,y) + \omega^-(x,y) = -\frac{2B(x,y)}{A(x,y)}, \qquad \omega^+(x,y)\omega^-(x,y) = \frac{C(x,y)}{A(x,y)}. \qquad (3.1.3)$$

Solving the system (3.1.3) for $\omega^+(x,y)$ and $\omega^-(x,y)$ gives

$$\omega^\pm(x,y) = \frac{-B(x,y) \pm \sqrt{B^2(x,y) - A(x,y)C(x,y)}}{A(x,y)}. \qquad (3.1.4)$$

The $\omega^\pm(x,y)$ are the roots of the quadratic equation $A(x,y)\omega^2 + 2B(x,y)\omega + C(x,y) = 0$.

We have obtained (3.1.2) by following the procedure used in Section 2.1, where the differential operators for the wave and Laplace's equations were factored. Since the coefficients A, B, and C are functions of x and y, there is an additional term involving ∂_y in (3.1.2). Instead of using (3.1.2) to express (3.1.1) as a system of two first order equations (as was done in Section 2.1), we use (3.1.2) to simplify the form of the given equation (3.1.1).

As we are generally interested in obtaining real-valued solutions of (3.1.1), the possibility of using the factorization (3.1.2) to simplify (3.1.1) depends on whether $\omega^\pm(x,y)$ are real- or complex-valued functions of x and y. This is determined by the sign of the discriminant $B^2(x,y) - A(x,y)C(x,y)$ in (3.1.4). We shall assume that $B^2(x,y) - A(x,y)C(x,y)$ is either of one sign or vanishes identically throughout the

given region R. On that basis we classify (3.1.1) as belonging to one of the following three types in the region R:

$$B^2(x,y) - A(x,y)C(x,y) > 0, \qquad \textit{hyperbolic type}, \tag{3.1.5}$$

$$B^2(x,y) - A(x,y)C(x,y) = 0, \qquad \textit{parabolic type}, \tag{3.1.6}$$

$$B^2(x,y) - A(x,y)C(x,y) < 0, \qquad \textit{elliptic type}. \tag{3.1.7}$$

Applying these criteria to the equations of Chapter 1, we find that the wave (and telegrapher's), diffusion, and Laplace's equations are of hyperbolic, parabolic, and elliptic type, respectively. We now determine canonical forms for (3.1.1) of each of the types (3.1.5)–(3.1.7).

Canonical Forms for Equations of Hyperbolic Type

When $B^2(x,y) - A(x,y)C(x,y) > 0$ in the region R, the functions $\omega^{\pm}(x,y)$ are real valued and distinct in that region. We may then express the operators $\partial_x - \omega^{\pm}\partial_y$ as directional derivatives. Along the family of curves

$$\frac{dy}{dx} = -\omega^{\pm}(x,y), \tag{3.1.8}$$

we have, for any differentiable function $v = v(x,y)$,

$$\frac{dv}{dx} = \frac{\partial v}{\partial x} + \frac{dy}{dx}\frac{\partial v}{\partial y} = (\partial_x - \omega^{\pm}\partial_y)v \tag{3.1.9}$$

on using (3.1.8).

The (one-parameter) families of curves determined by the solutions of (3.1.8) are called the *characteristic curves* of (3.1.1). They form two independent families of curves in the (x,y)-plane since $dy/dx = -\omega^{\pm}(x,y)$ with $\omega^{+}(x,y) \neq \omega^{-}(x,y)$ implies that they intersect nontangentially. We express these curves as

$$\xi = \xi(x,y), \qquad \eta = \eta(x,y), \tag{3.1.10}$$

where ξ = constant corresponds to the curves $y' + \omega^{+}(x,y) = 0$, while η = constant corresponds to the curves $y' + \omega^{-}(x,y) = 0$. For example, if the coefficients A, B, and C in (3.1.1) are constants, so are ω^{+} and ω^{-}. Then we can set $\xi = y + \omega^{+}x$ and $\eta = y + \omega^{-}x$.

We introduce ξ and η as new coordinates in (3.1.1), and this is referred to as the *characteristic coordinate system*. On the characteristic curves ξ = constant and η = constant, we have

$$0 = \xi_x + y'(x)\xi_y = \xi_x - \omega^{+}\xi_y, \qquad \omega^{+} = \xi_x/\xi_y, \tag{3.1.11}$$

$$0 = \eta_x + y'(x)\eta_y = \eta_x - \omega^{-}\eta_y, \qquad \omega^{-} = \eta_x/\eta_y. \tag{3.1.12}$$

It follows from our assumptions that $\xi_y \neq 0$ and $\eta_y \neq 0$. Since $\omega^+(x,y)$ and $\omega^-(x,y)$ are solutions of the quadratic equation $A(x,y)\omega^2+2B(x,y)\omega+C(x,y) = 0$ as shown above, we conclude that $\phi = \xi(x,y)$ and $\phi = \eta(x,y)$ both satisfy the *characteristic equation* for (3.1.1)

$$A(x,y)\phi_x^2(x,y) + 2B(x,y)\phi_x(x,y)\phi_y(x,y) + C(x,y)\phi_y^2(x,y) = 0. \quad (3.1.13)$$

Also, if $u = u(\xi,\eta) = u[\xi(x,y),\eta(x,y)]$, we have $\partial u/\partial x = u_\xi\xi_x + u_\eta\eta_x$ and $\partial u/\partial y = u_\xi\xi_y + u_\eta\eta_y$, so that

$$\frac{\partial u}{\partial x} - \omega^+\frac{\partial u}{\partial y} = (\partial_x - \omega^+\partial_y)u = u_\xi\xi_x + u_\eta\eta_x - \omega^+u_\xi\xi_y - \omega^+u_\eta\eta_y$$

$$= (\xi_x - \omega^+\xi_y)u_\xi + (\eta_x - \omega^+\eta_y)u_\eta = [(\eta_x - \omega^+\eta_y)\partial_\eta]u, \quad (3.1.14)$$

in view of (3.1.11). Similarly, (3.1.12) implies that

$$\frac{\partial u}{\partial x} - \omega^-\frac{\partial u}{\partial y} = (\partial_x - \omega^-\partial_y)u = [(\xi_x - \omega^-\xi_y)\partial_\xi]u. \quad (3.1.15)$$

Using (3.1.11)–(3.1.12) to express ξ_x and η_x in terms of ξ_y and η_y yields, in view of (3.1.14)–(3.1.15),

$$\partial_x - \omega^+\partial_y = -\eta_y(\omega^+ - \omega^-)\partial_\eta, \quad (3.1.16)$$

$$\partial_x - \omega^-\partial_y = \xi_y(\omega^+ - \omega^-)\partial_\xi. \quad (3.1.17)$$

Both operators are nonzero since $\omega^+ \neq \omega^-$. Thus

$$(\partial_x - \omega^+\partial_y)(\partial_x - \omega^-\partial_y) = [-\eta_y(\omega^+ - \omega^-)\partial_\eta][\xi_y(\omega^+ - \omega^-)\partial_\xi]$$

$$= -\xi_y\eta_y(\omega^+ - \omega^-)^2\partial_{\xi\eta}^2 - \eta_y(\omega^+ - \omega^-)\partial_\eta\left(\xi_y(\omega^+ - \omega^-)\right)\partial_\xi. \quad (3.1.18)$$

Substituting this result in (3.1.1), as modified in (3.1.2), and expressing everything in terms of ξ and η, we obtain (on dividing by the nonzero coefficient of $u_{\xi\eta}$)

$$u_{\xi\eta}(\xi,\eta) + a(\xi,\eta)u_\xi(\xi,\eta) + b(\xi,\eta)u_\eta(\xi,\eta) + c(\xi,\eta)u(\xi,\eta) = d(\xi,\eta). \quad (3.1.19)$$

This is one of the *canonical forms* for (3.1.1) when it is of *hyperbolic type*.

Although the factorization procedure was used to determine the principal part of the canonical form (3.1.19), it is generally simpler to use a more direct approach to obtain (3.1.19). Once the characteristic coordinates (3.1.10) have been found, we can transform (3.1.1) directly into the form (3.1.19), using the chain rule for differentiation. The principal part of the equation in $(\xi,\ \eta)$-coordinates is given as in (3.2.10). As shown in the exercises, the coefficient of $u_{\xi\eta}$ in (3.2.10) equals $(4/A)(AC - B^2)\xi_y\eta_y$, and this is nonzero in R.

The transformation

$$\xi = \alpha + \beta, \qquad \eta = \alpha - \beta, \tag{3.1.20}$$

in (3.1.19) yields an *alternative canonical form* for the *hyperbolic case*,

$$u_{\alpha\alpha} - u_{\beta\beta} + \tilde{a}u_{\alpha} + \tilde{b}u_{\beta} + \tilde{c}u = \tilde{d}, \tag{3.1.21}$$

where u and its derivatives and $\tilde{a}$, $\tilde{b}$, $\tilde{c}$, $\tilde{d}$, are functions of α and β. If A, B, and C in (3.1.1) are constants, the variables α and β are given as $\alpha = y - (B/A)x$ and $\beta = (1/A)(B^2 - AC)^{1/2}x$. In either of the canonical forms (3.1.19) or (3.1.21), the principal parts have constant coefficients. The telegrapher's and wave equations of Section 1.2 are of the form (3.1.21).

Canonical Forms for Equations of Parabolic Type

When $B^2(x,y) - A(x,y)C(x,y) = 0$ in the region R, we have $\omega^+(x,y) = \omega^-(x,y) = \omega(x,y)$ and (3.1.2) becomes

$$\partial_x^2 + (2B/A)\partial_x\partial_y + (C/A)\partial_y^2 = (\partial_x - \omega\partial_y)^2 + (\omega_x - \omega\,\omega_y)\partial_y, \tag{3.1.22}$$

with $\omega(x,y) = -B(x,y)/A(x,y)$, in view of (3.1.4). Since there is only one value of $\omega(x,y)$, we obtain only one family of characteristic curves, determined from

$$\frac{dy}{dx} = -\omega(x,y) = \frac{B(x,y)}{A(x,y)}. \tag{3.1.23}$$

Let

$$\xi = \xi(x,y), \qquad \eta = \eta(x,y). \tag{3.1.24}$$

Here ξ = constant is the family of characteristic curves determined from (3.1.23). η = constant is an arbitrarily chosen independent family of curves, chosen such that the Jacobian determinant $\xi_x\eta_y - \xi_y\eta_x$ of the transformation (3.1.24) is nonzero in the region R.

Again, $\phi = \xi(x,y)$ is a solution of the characteristic equation (3.1.13). For example, if A, B, and C in (3.1.1) are constants, so is ω, and we can set $\xi = y + \omega x$ and $\eta = \omega y + cx$, where $c \neq \omega^2$ but is an otherwise arbitrary constant. Proceeding as for the hyperbolic case, introducing the coordinates (3.1.24) and using the chain rule to transform (3.1.1), we obtain the *canonical form* [see (3.2.10) and the exercises]

$$u_{\eta\eta}(\xi,\eta) + a(\xi,\eta)u_{\xi}(\xi,\eta) + b(\xi,\eta)u_{\eta}(\xi,\eta) + c(\xi,\eta)u(\xi,\eta) = d(\xi,\eta), \tag{3.1.25}$$

where a, b, c, and d are specified functions of ξ and η. Again, the coefficients in the principal part of (3.1.25) are constant. The diffusion equation of Section 1.1 has the form (3.1.25).

Canonical Forms for Equations of Elliptic Type

When $B^2(x,y) - A(x,y)C(x,y) < 0$, $\omega^+(x,y)$ and $\omega^-(x,y)$ as defined in (3.1.4), are complex valued in the region R. Hence the characteristics defined by (3.1.8) are complex curves. Assuming that the terms A, $B, \ldots,$ G in (3.1.1) can be defined for complex arguments, the change of variables (3.1.10) may still be applied to bring (3.1.1) into the form (3.1.19). Again, the coefficient of $u_{\xi\eta}$ is given as $(4/A)(AC - B^2)\xi_y\eta_y$ and this is nonzero. However, the variables ξ and η are complex, so that the usefulness of the canonical form (3.1.19) is questionable. For example, if A, B, and C in (3.1.1) are constants, we obtain $\xi = y - (B/A)x + i(1/A)(AC - B^2)^{1/2}x$ and $\xi = y - (B/A)x - i(1/A)(AC - B^2)^{1/2}x$, where $i = \sqrt{-1}$.

The further transformation

$$\xi = \alpha + i\beta, \qquad \eta = \alpha - i\beta \tag{3.1.26}$$

in (3.1.19) introduces the real-valued variables α and β. (This follows since ω^+ and ω^- are complex conjugates in the elliptic case, when the terms A, B, and C are real as assumed.) If A, B, and C are constants, $\alpha = y - (B/A)x$ and $\beta = (1/A)(AC - B^2)^{1/2}x$. We then easily obtain the *canonical form*

$$u_{\alpha\alpha} + u_{\beta\beta} + a(\alpha,\beta)u_\alpha + b(\alpha,\beta)u_\beta + c(\alpha,\beta)u = d(\alpha,\beta) \tag{3.1.27}$$

for the elliptic case where u and its derivatives are functions of α and β and a, b, c, d are real valued functions. [The canonical form (3.1.27) can also be obtained by another method, which we do not consider, that does not require the introduction of complex variables.] Laplace's equation of Section 1.3 has the form (3.1.27).

Equations of Mixed Type

If the terms A, $B, \ldots,$ F in (3.1.1) are constants, the canonical forms for the hyperbolic, parabolic, and elliptic cases are valid over the entire domain of definition of the inhomogeneous term G. Each canonical form has constant coefficients. Then a further simplification of the equation can be achieved, as shown in the exercises. If one or more of the terms A, $B, \ldots,$ F in (3.1.1) is variable, and the partial differential equation is of hyperbolic, parabolic, or elliptic type throughout the region of interest, it can be brought into one of the foregoing canonical forms. However, if the equation is of more than one type within the region, it is said to be of *mixed type*. In that case, it can be brought into an appropriate canonical form within each subregion where it has fixed type. We consider an example of an equation that is of all three types.

Example 3.1. An Equation of Mixed Type. The equation

$$u_{xx}(x,y) + y\, u_{yy}(x,y) = 0 \tag{3.1.28}$$

is of *mixed type*, since it is hyperbolic for $y < 0$, parabolic for $y = 0$, and elliptic for $y > 0$, as follows immediately from the fact that $B^2(x,y) - A(x,y)C(x,y) = -y$. We obtain canonical forms for (3.1.28) in each of the three regions.

The parabolic region is the simplest to deal with, since for $y = 0$ the equation assumes the canonical form

$$u_{xx}(x, y) = 0. \tag{3.1.29}$$

In this case $\omega = 0$, so that the characteristic curve determined from $dy/dx = 0$ is $y = 0$. That is, the x-axis is the characteristic curve, and it represents a curve across which a transition from hyperbolic to elliptic type takes place. For this reason, the solution of initial and boundary value problems for (3.1.28) in a neighborhood of the x-axis is rather difficult since (3.1.28) is of three different types in this neighborhood.

In the hyperbolic region $y < 0$, we have

$$\omega^{\pm}(x, y) = \pm\sqrt{-y}. \tag{3.1.30}$$

The characteristic equations

$$y'(x) = -\omega^{\pm}(x, y) = \mp\sqrt{-y} \tag{3.1.31}$$

yield the two characteristic families

$$\xi = x + 2\sqrt{-y} = \text{constant}, \quad \eta = x - 2\sqrt{-y} = \text{constant}. \tag{3.1.32}$$

From (3.1.2) we obtain for the differential operator in (3.1.28)

$$\partial_x^2 + y\partial_y^2 = (\partial_x - \sqrt{-y}\,\partial_y)(\partial_x + \sqrt{-y}\,\partial_y) - \frac{1}{2}\partial_y, \tag{3.1.33}$$

and from (3.1.18),

$$(\partial_x - \sqrt{-y}\,\partial_y)(\partial_x + \sqrt{-y}\,\partial_y) = 4\partial_{\xi\eta}^2. \tag{3.1.34}$$

Since $\partial_y = \xi_y\partial_\xi + \eta_y\partial_\eta = -(1/\sqrt{-y})(\partial_\xi - \partial_\eta)$ and $\xi - \eta = 4\sqrt{-y}$, we obtain the canonical form for (3.1.28) in the hyperbolic region as

$$u_{xx}(x, y) + y\,u_{yy}(x, y) = 4\left[u_{\xi\eta}(\xi, \eta) + \frac{1}{2(\xi - \eta)}(u_\xi(\xi, \eta) - u_\eta(\xi, \eta))\right] = 0. \tag{3.1.35}$$

The characteristic curves for the hyperbolic case are the two branches of the parabolas $y = -\frac{1}{4}(x - c)^2$, where c is a constant, as shown in Figure 3.1. The branches with the positive slopes give the curves ξ = constant, while η = constant gives the branches with the negative slopes. We note that both branches are tangent to the x-axis, which is the single characteristic curve in the parabolic region. In fact, the x-axis is the envelope of the characteristic curves for the hyperbolic region $y < 0$.

In the elliptic region $y > 0$, we have $\omega^{\pm}(x, y) = \pm\sqrt{-y} = \pm i\sqrt{y}$, so that the characteristic curves are complex. The equations (3.1.32) remain valid but they now take the form

$$\xi = x + 2i\sqrt{y} = \text{constant}, \qquad \eta = x - 2i\sqrt{y} = \text{constant}. \tag{3.1.36}$$

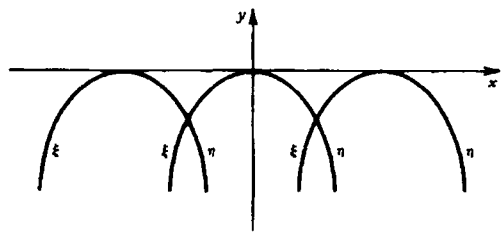

Figure 3.1 The characteristic curves.

Using (3.1.26) gives

$$\alpha = x, \qquad \beta = 2\sqrt{y}. \tag{3.1.37}$$

With α and β as the new variables, we have

$$u_{xx} = u_{\alpha\alpha}, \qquad u_y = \frac{1}{\sqrt{y}}\, u_\beta, \qquad u_{yy} = \frac{1}{y}\, u_{\beta\beta} - \frac{1}{2y^{3/2}}\, u_\beta = 0, \tag{3.1.38}$$

so that the canonical form is

$$u_{xx}(x,y) + y u_{yy}(x,y) = u_{\alpha\alpha}(\alpha,\beta) + u_{\beta\beta}(\alpha,\beta) - \frac{1}{\beta}\, u_\beta(\alpha,\beta) = 0. \tag{3.1.39}$$

We conclude our discussion of this example with the observation that the canonical forms (3.1.35) and (3.1.39) in the hyperbolic and elliptic cases have coefficients that are singular when $\xi = \eta$ and $\beta = 0$, respectively. Both singular regions correspond to the x-axis, across which the equation (3.1.28) undergoes a transition from hyperbolic to parabolic to elliptic type. The lack of validity of the hyperbolic and elliptic canonical forms at the x-axis is signaled by the singularity of the coefficients there. Although the study of equations of mixed type is of interest in a number of applications, we shall deal (almost) exclusively with equations that are of a single type in the region under consideration.

Exercises 3.1

3.1.1. Show that the coefficient of $u_{\xi\eta}$ in (3.2.10) equals $(4/A)(AC - B^2)\xi_y\eta_y$ in the hyperbolic case if ξ and η are defined as in Section 3.1.

3.1.2. Show that if $\phi = \xi$ and $\psi = \eta$, as defined in (3.1.24), are used as coordinates in transforming (3.1.1) in the parabolic case, the coefficient of $u_{\eta\eta}$ in (3.2.10) is the only one that is nonzero, and the equation (3.1.25) results.

3.1.3. Show that if the transformation (3.1.10) is introduced into (3.1.1) in the elliptic case, the coefficient of $u_{\xi\eta}$ in (3.2.10) equals $(4/A)(AC - B^2)\xi_y\eta_y$.

3.1.4. Show that the equation $u_{xx}(x,y) + 4u_{xy}(x,y) + 3u_{yy}(x,y) + 3u_x(x,y) - u_y(x,y) + 2u(x,y) = 0$ is of hyperbolic type. Determine its characteristic curves

and bring it into the canonical form (3.1.19). Introduce the further transformation $u(\xi,\eta) = \exp(\alpha\xi + \beta\eta)v(\xi,\eta)$, and choose the constants α and β to eliminate the first derivative terms in the resulting canonical form.

3.1.5. Show that the equation $u_{xx}(x,y) + 2u_{xy}(x,y) + u_{yy}(x,y) + 5u_x(x,y) + 3u_y(x,y) + u(x,y) = 0$ is of parabolic type and bring it into the canonical form (3.1.25). Using the transformation of the dependent variable given in Exercise 3.1.4, show that the terms involving u_ξ and u in the canonical form can be eliminated.

3.1.6. Show that the equation $u_{xx}(x,y) - 6u_{xy}(x,y) + 12u_{yy}(x,y) + 4u_x(x,y) - u(x,y) = \sin(xy)$ is of elliptic type and bring it into the canonical form (3.1.27). Proceeding as in the exercises above, show that the first derivative terms in the canonical form can be eliminated.

3.1.7. Determine the regions where *Tricomi's equation* $u_{xx}(x,y)+xu_{yy}(x,y) = 0$ is of elliptic, parabolic, and hyperbolic types. Obtain its characteristics and its canonical form in the hyperbolic region.

3.1.8. Show that the equation $u_{xx}(x,y)+yu_{yy}(x,y)+\frac{1}{2}u_y(x,y) = 0$ has the simple canonical form $u_{\xi\eta} = 0$ in the region where it is of hyperbolic type. Use this result to show that it has the general solution $u(x,y) = f(x + 2\sqrt{-y}) + g(x - 2\sqrt{-y})$, with f and g as arbitrary functions, in the hyperbolic region.

3.1.9. Classify the following equations as of hyperbolic, elliptic or parabolic type:
(a) $5u_{xx}(x,y) - 3u_{yy}(x,y) + \sin(x)u_x(x,y) + e^{xy^2}u_y(x,y) + u(x,y) = 0$.
(b) $3u_{yy}(x,y) - 10u_x(x,y) + 4u_y(x,y) + \cosh(x)u(x,y) = 0$.
(c) $10u_{xx}(x,y) + u_{yy}(x,y) - u_x(x,y) + [\log(1+x^2)]u(x,y) = 0$.

3.1.10. Classify the following equations as of hyperbolic, elliptic, or parabolic type:
(a) $e^{xy}u_{xx}(x,y) + \cosh(x)u_{yy}(x,y) + u_y(x,y) - u(x,y) = 0$.
(b) $[\log(1+x^2+y^2)]u_{xx}(x,y) - [2+\cos(x)]u_{yy}(x,y) = 0$.
(c) $u_{yy}(x,y) + [1+x^2]u_x(x,y) - u_y(x,y) + u(x,y) = 0$.

3.1.11. Show that if $B = 0$ in (3.1.1), the equation is of fixed type if $AC < 0$, $AC = 0$, or $AC > 0$ everywhere in the region where (3.1.1) is defined.

3.2 CHARACTERISTIC CURVES

To introduce and motivate our discussion of characteristic curves for second order partial differential equations, we begin by reconsidering linear first order equations and their characteristics. However, we now discuss them from a somewhat different point of view than that of Section 2.2.

First Order PDEs

The characteristic (base) curves of the linear first order PDE

$$a(x,y)u_x(x,y) + b(x,y)u_y(x,y) = c(x,y)u(x,y) + d(x,y), \qquad (3.2.1)$$

can be obtained [assuming that $a(x, y) \neq 0$] as solutions of

$$\frac{dy}{dx} = \frac{b(x, y)}{a(x, y)}. \tag{3.2.2}$$

We suppose that $y = h(x)$ is a solution of (3.2.2) and look for a solution $u(x, y)$ of (3.2.1) with the initial value $u(x, h(x)) = f(x)$. Using the chain rule and (3.2.1)–(3.2.2), we conclude that $a\, du/dx = af'(x) = cf(x) + d$ on the characteristic base curve $y = h(x)$. Unless $f(x)$ satisfies this *compatibility condition*, the initial value problem has no solution. If $f(x)$ does satisfy the compatibility equation, the curve $y = h(x)$, $u(x, h(x)) = f(x)$ in (x, y, u)-space is a characteristic curve for (3.2.1) and there are infinitely many solutions or integral surfaces $u = u(x, y)$ of (3.2.1) that contain the characteristic curve.

These results were obtained in Section 2.2 (with y replaced by t). We assume that a and b in (3.2.1) are not both zero. If $b \neq 0$ but a can vanish, we can express (3.2.1) in a form equivalent to (2.2.35) and adapt the discussion of that equation to obtain compatibility conditions in the present case. In this section, however, we deal with the case where a is nonzero.

It was also shown in Section 2.2 that a curve $y = h(x)$ along which a solution of (3.2.1) is continuous but whose derivatives are discontinuous across it must be a characteristic base curve [i.e., is, it satisfies (3.2.2)]. We demonstrate this again using a different method. This method can easily be generalized to deal with higher-order differential equations.

Let $y = h(x)$ be the curve of (derivative) discontinuity and represent it implicitly as $\phi(x, y) = y - h(x) = 0$. We consider the family of curves $\phi(x, y) =$ constant and the family of orthogonal trajectories $\psi(x, y) =$ constant [the orthogonal trajectories are solutions of the equation $y' = -1/h'(x)$]. Introducing the (ξ, η)-coordinate system defined by

$$\xi = \phi(x, y), \qquad \eta = \psi(x, y), \tag{3.2.3}$$

we obtain for $u = u(\xi, \eta) = u(\phi(x, y), \psi(x, y))$,

$$u_x = u_\xi \xi_x + u_\eta \eta_x = \phi_x u_\xi + \psi_x u_\eta, \; u_y = u_\xi \xi_y + u_\eta \eta_y = \phi_y u_\xi + \psi_y u_\eta. \tag{3.2.4}$$

Thus

$$au_x + bu_y = (a\phi_x + b\phi_y)u_\xi + (a\psi_x + b\psi_y)u_\eta = cu + d. \tag{3.2.5}$$

Now let $u_1(x, y)$ and $u_2(x, y)$ represent solutions of (3.2.5) defined in the regions $\xi \leq 0$ and $\xi \geq 0$, respectively, with $u_1 = u_2$ along the curve $\xi = y - h(x) = 0$. With $u = u_1$ for $\xi \leq 0$ and $u = u_2$ for $\xi > 0$, the solution u and the derivative u_η are continuous on the curve $\xi = 0$, since u_η is an interior derivative along the curve $\xi =$ constant. However, u_ξ represents a normal derivative across the curve $\xi = 0$ and u_ξ has a jump discontinuity there by assumption. Since u_1 and u_2 are solutions of (3.2.5) for $\xi \neq 0$, we have

$$[(a\phi_x + b\phi_y)u_\xi + (a\psi_x + b\psi_y)u_\eta - cu - d]\Big|_{u=u_1,\xi<0}^{u=u_2,\xi>0} = 0 \tag{3.2.6}$$

for the difference of the bracketed expression across the curve $\xi = 0$ along the orthogonal curves η = constant. In the limit as $\xi \to 0$ it follows that

$$(a\phi_x + b\phi_y)[u_\xi]\big|_{\xi=0} = 0, \tag{3.2.7}$$

where $[u_\xi]|_{\xi=0}$ is the jump in u_ξ across the curve $\xi = 0$. The other terms vanish since they are continuous at $\xi = 0$. By assumption $[u_\xi]|_{\xi=0} \neq 0$, so that (3.2.7) implies that

$$a\phi_x + b\phi_y|_{\xi=0} = -ah'(x) + b = 0, \tag{3.2.8}$$

since $\xi = 0$ corresponds to the curve $\phi(x, y) = y - h(x) = 0$. As a result, $h'(x) = y'(x) = b/a$, so that $y = h(x)$ is a solution of the characteristic equation (3.2.2), as was to be shown. That is, discontinuities in derivatives of continuous solutions $u = u(x, y)$ of (3.2.1) can occur only across characteristic (base) curves. However, this requires that a, b, c, and d are continuous (see Exercise 3.2.7).

We show that if initial data $u(x, y)$ are prescribed for (3.2.1) along a characteristic (base) curve $y = h(x)$, it is impossible to solve for u_x and u_y uniquely along that curve, if a solution exists at all. Thus if there is a solution, no unique tangent plane can be defined along the initial curve and it is not possible to specify a unique integral surface that contains the initial curve and satisfies the initial value problem. (This result has been obtained previously, based on the method of characteristics, but we now obtain it in a different way, one that can easily be generalized to deal with systems of equations.)

Given the initial value $u(x, h(x)) = f(x)$ on the curve $y = h(x)$ that satisfies (3.2.2), we have $du(x, h(x))/dx = u_x + h'(x)u_y = f'(x)$. Combining this with (3.2.1) evaluated on $y = h(x)$ yields the simultaneous linear system for u_x and u_y, $\{u_x + h'(x)u_y = f'(x),\ au_x + bu_y = cf(x) + d\}$. The determinant of coefficients of this system vanishes; that is,

$$D = \begin{vmatrix} 1 & h'(x) \\ a & b \end{vmatrix} = -ah'(x) + b = 0, \tag{3.2.9}$$

since $y'(x) = h'(x) = b/a$ in view of (3.2.2). Consequently, either u_x and u_y cannot be determined at all on $y = h(x)$ or, at best, they can be solved for nonuniquely if $f(x)$ satisfies a consistency condition. This condition is readily found to be $af'(x) = cf(x) + d$, which is identical with the compatibility condition given above. Conversely, the discussion above shows that a curve $y = h(x)$ on which u is prescribed and for which u_x and u_y either cannot be determined or can only be determined nonuniquely must be characteristic since we must then have $D = 0$ in (3.2.9). The determinant D vanishes only if $y = h(x)$ is a solution of (3.2.2).

Second Order PDEs

Characteristic curves for the second order equation (3.1.1) are determined by using the first of the two techniques developed for first order equations. (The second approach is presented in the exercises.) We specify the possible curves $y = h(x)$ across which the second derivatives of a solution $u(x, y)$ of (3.1.1) can have discontinuities. The solution $u(x, y)$ and its first derivatives are assumed to be continuous across these curves.

Let $\phi(x, y) = y - h(x) = 0$ and define the orthogonal family of curves $\xi =$ constant and $\eta =$ constant as in (3.2.3). In terms of the variables ξ and η, (3.1.1) takes the form

$$(A\phi_x^2 + 2B\phi_x\phi_y + C\phi_y^2)u_{\xi\xi} + 2\left(A\phi_x\psi_x + B(\phi_x\psi_y + \phi_y\psi_x) + C\phi_y\psi_y\right)u_{\xi\eta}$$

$$+(A\psi_x^2 + 2B\psi_x\psi_y + C\psi_y^2)u_{\eta\eta} + \cdots = G, \tag{3.2.10}$$

where the dots represent first derivative and undifferentiated terms in u. Since u and u_ξ are assumed to be continuous across $\xi = y - h(x) = 0$, we conclude that u_η, $u_{\xi\eta}$, and $u_{\eta\eta}$ are also continuous, since $\partial/\partial\eta$ is an interior derivative operator along $\xi = 0$.

Again we consider solutions $u_1(x, y)$ and $u_2(x, y)$ of (3.1.1) defined in the regions $\xi \leq 0$ and $\xi \geq 0$, respectively. Evaluating (3.2.10) in each of the regions above, taking the difference of these equations, and going to the limit as $\xi \to 0$, we obtain

$$(A\phi_x^2 + 2B\phi_x\phi_y + C\phi_y^2)[u_{\xi\xi}]|_{\xi=0} = 0, \tag{3.2.11}$$

where $[u_{\xi\xi}]|_{\xi=0}$ is the jump in the second derivative of u across the curve $\xi = 0$. All other terms vanish in view of their assumed continuity across that curve. Since $[u_{\xi\xi}]|_{\xi=0}$ is taken to be nonzero, we must have

$$A\phi_x^2 + 2B\phi_x\phi_y + C\phi_y^2 = 0. \tag{3.2.12}$$

This is the *characteristic equation* for (3.1.1) [see (3.1.13)].

Assuming (as was done in Section 3.1) that the coefficient A in (3.1.1) is nonzero, we can factor (3.2.12) as

$$A(\phi_x - \omega^+ \phi_y)(\phi_x - \omega^- \phi_y) = 0, \tag{3.2.13}$$

with ω^+ and ω^- defined as in (3.1.4). If ω^+ and ω^- are real valued, we see that a curve $\xi = \phi(x, y) = 0$ or, equivalently, $y = h(x)$ across which $u_{\xi\xi}$ is discontinuous must be one of the characteristic curves (3.1.8). This follows since $\phi = y - h(x)$ implies that $\phi_x - \omega^\pm \phi_y = -h'(x) - \omega^\pm$. Thus one of the two equations $y' = h' = -\omega^\pm$ must be satisfied in view of (3.2.13).

Using similar arguments, it can be shown that discontinuities in higher derivatives of a solution must also occur only across characteristic curves.

Example 3.2. The Wave, Diffusion, and Laplace's Equations. For the wave equation (1.2.35), the characteristic equation (3.2.12) has the form

$$\phi_t^2 - \gamma^2 \phi_x^2 = (\phi_t + \gamma\phi_x)(\phi_t - \gamma\phi_x) = 0. \tag{3.2.14}$$

Equating each factor to zero separately gives the straight-line characteristic curves $\phi(x,t) = x \pm \gamma t =$ constant, with slopes $dt/dx = \mp 1/\gamma$. The telegrapher's equation (1.2.25) has the same principal part as the wave equation (1.2.35), so it has the same characteristic curves.

For the diffusion equation (1.1.15) we have the characteristic equation $(D/2)\phi_x^2 = 0$, which implies that $\phi = \phi(t)$ = constant, so that the straight lines t = constant are the characteristics.

Laplace's equation (1.3.4) yields the characteristic equation

$$\phi_x^2 + \phi_y^2 = 0. \tag{3.2.15}$$

For real-valued $\phi(x,y)$ we must have $\phi_y = \phi_x = 0$, so that no real solutions other than $\phi \equiv$ constant exist. Thus, Laplace's equation has no real characteristic curves.

The foregoing results are typical for equations of hyperbolic, parabolic, and elliptic type. *Elliptic equations* have no real characteristic curves, so that solutions cannot have discontinuous derivatives. Consequently, solutions $u = u(x,y)$ are extremely smooth functions, and this is consistent with the nature of elliptic equations that describe equilibrium processes where everything has already smoothed itself out.

Parabolic equations of the type we generally consider have the lines t = constant as characteristics, and discontinuities in derivatives must occur across these lines. Since we are concerned with the evolution of solutions as t (i.e., the time variable) increases from some fixed value of t (usually, $t = 0$), if these discontinuities occur initially when $t = 0$, they cannot be spread into the region $t > 0$, and again solutions are smooth functions.

It is for *hyperbolic equations* that characteristics play the most significant role. Example 3.2 shows that for the wave and telegrapher's equations the characteristics extend from the (initial) line, say, $t = 0$ into the region $t > 0$. Thus in their role as curves across which discontinuities or singularities in the solution occur, they act as carriers of singular initial data for the solution, and the effects of these singularities are felt for all time. The importance of characteristics in the solution of various problems for hyperbolic equations is demonstrated throughout the book.

Exercises 3.2

3.2.1. If $\phi(x,y)$ = constant is a family of characteristics for (3.2.1), show that (3.2.5) reduces to $(a\psi_x + b\psi_y)u_\eta = cu + d$. Assuming that u is continuous across $\xi = \phi(x,y) = 0$ whereas u_ξ has a jump discontinuity there, demonstrate that the jump

$[u_\xi]$ at $\xi = 0$ satisfies the ordinary differential equation $(a\psi_x + b\psi_y)[u_\xi]_\eta = c[u_\xi]$ by differentiating (3.2.5) with respect to ξ. This result determines the variation of the jump $[u_\xi]$ along the characteristic $\xi = 0$.

3.2.2. Consider the equation $u_x(x, y) + u_y(x, y) = 0$, and the continuous solution $u(x, y) = \alpha_1(y - x)$ for $y \leq x$ and $u(x, y) = \alpha_2(y - x)$ for $y > x$. Let $\xi = y - x$. Show that for $\alpha_1 \neq \alpha_2$, $[u_\xi] = \alpha_2 - \alpha_1$ across $\xi = 0$ and that this jump satisfies the condition found in Exercise 3.2.1 for the variation of the jump.

3.2.3. Consider the equation $u_x(x, y) + u_y(x, y) - cu(x, y) = 0$ where c is a constant. Show that the appropriate (ξ, η) coordinates in (3.2.3) are $\xi = y - x$, $\eta = y + x$, where $\xi = y - x$ = constant are the characteristics. Obtain the equation $2u_\eta(\xi, \eta) - cu(\xi, \eta) = 0$ and the equation $2[u_\xi]_\eta = c[u_\xi]$ for the jump $[u_\xi]$. Verify that both $u(x, y) = \beta_1(y - x)\exp[\frac{1}{2}c(y + x)]$, $y \leq x$ and $u(x, y) = \beta_2(y - x)\exp[\frac{1}{2}c(y + x)]$, $y > x$ are solutions of the given equation for $y \neq x$ and that the jump in u_ξ across $\xi = y - x = 0$ satisfies the equation for the jump (β_1 and β_2 are constants).

3.2.4. If $\phi(x, y)$ = constant is a family of characteristics for (3.1.1), show that (3.2.10) takes the form $2\beta u_{\xi\eta} + \gamma u_{\eta\eta} + \delta u_\eta + \rho u_\xi + \lambda u + \sigma = 0$. If $u_{\xi\xi}$ has a jump across $\xi = \phi(x, y) = 0$ and the terms in the equation abobve are continuous across $\xi = 0$, conclude on differentiating the equation with respect to ξ that the jump $[u_{\xi\xi}]$ satisfies $2\beta[u_{\xi\xi}]_\eta + \rho[u_{\xi\xi}] = 0$ at $\xi = 0$, assuming smooth coefficients.

3.2.5. Given the hyperbolic equation $u_{xx}(x, y) - u_{yy}(x, y) = 0$, show that $u(x, y) = \beta_1(y - x)^2$, $y \leq x$ and $u(x, y) = \beta_2(y - x)^2$, $y > x$ are solutions with discontinuous second derivatives across the characteristic $y = x$ if the constants β_1 and β_2 are unequal. If $\xi = y - x$, show that $[u_{\xi\xi}]$ satisfies the appropriate form of the equation for the jump across $\xi = 0$ given in Exercise 3.2.4.

3.2.6. Let $v(x, t)$ be a continuous solution of the diffusion equation (1.1.15) and assume that $v_t(x, t)$ has a jump across a characteristic t = constant. Show that the jump $[v_t]$ satisfies the equation $[v_t] = 0$, so that no jump can exist under the given conditions.

3.2.7. Let the inhomogeneous term $d(x, y)$ in (3.2.1) have a jump across the curve $\xi = \phi(x, y) = 0$ and assume that the solution $u(x, y)$ is continuous there. Proceeding as in the text, show that the equation for the jump in u_ξ across $\xi = 0$ is $(a\phi_x + b\phi_y)[u_\xi] = [d]$. Thus, if $\xi = \phi(x, y) = 0$ is a characteristic curve, we must assume that u itself has a jump across $\xi = 0$ to avoid a contradiction.

3.2.8. Show that the results obtained in Exercise 2.2.11 are consistent with those that follow from Exercise 3.2.7.

3.2.9. By differentiating (3.2.10) as often as necessary with respect to ξ, show that jumps in $\partial^n u/\partial\xi^n$, $n \geq 3$ must also occur across characteristics, assuming the appropriate lower-order derivatives are continuous.

3.2.10. If u and the normal derivative $\partial u/\partial n$ are specified on a curve $y = h(x)$, show how to specify u_x and u_y on that curve. Then, assuming that $u_x[x, h(x)] = \alpha(x)$ and $u_y[x, h(x)] = \beta(x)$, where $\alpha(x)$ and $\beta(x)$ are known functions, obtain the

equations $\{u_{xx} + u_{xy}h' = \alpha',\ u_{xy} + u_{yy}h' = \beta',\ Au_{xx} + 2Bu_{xy} + Cu_{yy} = -D\alpha - E\beta - Ff + G\}$ for the specification of u_{xx}, u_{xy}, and u_{yy} along the curve $y = h(x)$. [We assume that $u[x, h(x)] = f(x)$.] Show that if $Ah'^2 - 2Bh' + C = 0$, the second derivatives of u cannot be specified uniquely on $y = h(x)$, if they can be determined at all. If we set $\phi = y - h(x)$, show that the equation above becomes the characteristic equation (3.2.12).

3.2.11. Consider the hyperbolic equation $u_{xx}(x, y) - u_{yy}(x, y) = 0$. Apply the method of the Exercise 3.2.10 and show that u_{xx}, u_{xy}, and u_{yy} can be determined nonuniquely on a characteristic if the compatibility condition $\alpha'h' - \beta' = 0$ is satisfied. Construct a set of initial conditions for u and $\partial u/\partial n$ on the characteristic $y - x = 0$ such that the compatibility conditions are satisfied, and show that this problem for the given equation has infinitely many solutions.

3.2.12. Show that the compatibility condition that guarantees that the problem in Exercise 3.2.10 has a (nonunique) solution along the characteristic $y = h(x)$ is $Ah'\beta' - 2B\beta' - A\alpha' + G - Ff - E\beta - D\alpha = 0$.

3.3 CLASSIFICATION OF EQUATIONS IN GENERAL

Classification of Second Order PDEs

We begin by considering the second order linear partial differential equation in n variables

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\frac{\partial^2 u}{\partial x_i\,\partial x_j} + \sum_{i=1}^{n} b_i\frac{\partial u}{\partial x_i} + cu + d = 0, \tag{3.3.1}$$

where u, a_{ij}, b_i, c, and d are functions of $x_1, \ldots, x_n$. This equation cannot, in general, be reduced to a simple canonical form over a full region, as was done in the case of two independent variables in Section 3.1. For the purpose of classifying (3.3.1) into different types, we shall generalize the factorization procedure of Section 3.1. It will then be seen that if the coefficients a_{ij} are constants or if we restrict ourselves to a single point in $(x_1, \ldots, x_n)$-space, it is possible to bring the *principal part* of (3.3.1) (i.e., the highest derivative terms) into a *canonical form.*

With $\partial_{x_i} = \partial/\partial x_i$, $i = 1, \ldots, n$, and $\boldsymbol{\partial}_x^T = [\partial_{x_1}, \ldots, \partial_{x_n}]$ as a row (gradient) vector (which is the transpose of the column vector $\boldsymbol{\partial}_x$), it is easily seen that we may express the principal part of the differential operator in (3.3.1) as

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\partial_{x_i}\partial_{x_j} = \boldsymbol{\partial}_x^T A\boldsymbol{\partial}_x + \cdots. \tag{3.3.2}$$

The dots stand for first derivative terms and the $n \times n$ matrix A has the coefficients a_{ij} as its elements. We note that if the a_{ij} are constants, the first derivative terms in (3.3.2) are absent. Since we assume that mixed partial derivatives of u are equal (i.e., $u_{x_ix_j} = u_{x_jx_i}$), (3.3.1) may be arranged so that its coefficients have the property that

$a_{ij} = a_{ji}$ and we assume that this has been done. Thus the matrix A is symmetric and it is assumed to be real valued in the present discussion.

It is well known from matrix theory that a real-valued symmetric matrix A has only real eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ (counted with their multiplicities) and that there exists a corresponding orthonormal set of n eigenvectors $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n$. Forming the matrix R with the eigenvectors $\mathbf{r}_i$ as its n columns, we find that R is an orthogonal matrix with the property

$$R^T A R = D = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}. \tag{3.3.3}$$

That is, using R, we can diagonalize the matrix A.

We now introduce the directional derivative operators

$$\partial_{\xi_i} = \mathbf{r}_i^T \boldsymbol{\partial}_x, \qquad i = 1, \ldots, n, \tag{3.3.4}$$

and form the vector operator

$$\boldsymbol{\partial}_\xi = R^T \boldsymbol{\partial}_x. \tag{3.3.5}$$

[The operators ∂_{ξ_i} may be compared with the operators ∂_ξ and ∂_η in (3.1.16)–(3.1.17).] Since R is an orthogonal matrix so that $R^T = R^{-1}$ (i.e., its transpose equals its inverse), (3.3.5) yields $\boldsymbol{\partial}_x = R\boldsymbol{\partial}_\xi$. Introducing this expression into (3.3.2) gives

$$\boldsymbol{\partial}_x^T A \boldsymbol{\partial}_x + \cdots = \boldsymbol{\partial}_\xi^T (R^T A R) \boldsymbol{\partial}_\xi + \cdots = \boldsymbol{\partial}_\xi^T D \boldsymbol{\partial}_\xi + \cdots = \sum_{i=1}^{n} \lambda_i \partial_{\xi_i}^2 + \cdots, \tag{3.3.6}$$

where the dots represent first derivative operators (since the elements in the foregoing matrices need not be constant) and where (3.3.3) has been used.

Recalling the classification method for equations in two independent variables given in Section 3.1, we find that in terms of the result (3.3.6), the *elliptic case* corresponds to the situation when λ_1 and λ_2 have the same sign. The *hyperbolic case* requires that λ_1 and λ_2 have opposite signs, whereas the *parabolic case* occurs when λ_1 or λ_2 equals zero. For the n-dimensional case in (3.3.1), we base our classification of the equation at a point P in $(x_1, \ldots, x_n)$-space on the result obtained in (3.3.6). Thus it is characterized by the properties of the eigenvalues of A.

With $\lambda_1, \lambda_2, \ldots, \lambda_n$ as the eigenvalues of A, we introduce the following *classification* for the equation (3.3.1):

$$\begin{cases} \lambda_i > 0, & \text{all } i, \\ \text{or} & \\ \lambda_i < 0, & \text{all } i, \end{cases} \qquad \textit{elliptic type.} \tag{3.3.7}$$

$$\begin{cases} \text{One of the } \lambda_i > 0 \text{ or } \lambda_i < 0, \\ \text{All other } \lambda_i \text{ have opposite sign,} \end{cases} \qquad \textit{hyperbolic type.} \tag{3.3.8}$$

$$\text{One or more of the } \lambda_i = 0, \qquad \textit{parabolic type.} \tag{3.3.9}$$

For the two- and three-dimensional cases, the classification above exhausts all possibilities. However, if $n \geq 4$ in (3.3.1), it may happen that two or more of the λ_i have one sign, while two or more of the remaining λ_i have the opposite sign. Such equations are said to be of *ultrahyperbolic type*. Since they do not occur often in applications, they are not studied in this book.

If (3.3.1) is of fixed type at every point in a region, the equation is said to be of that type in the region. If the type changes within a region, the equation is said to be of *mixed type*. When the principal part of (3.3.1) has constant coefficients, the equation is clearly of one type everywhere in $(x_1, \ldots, x_n)$-space.

As basic examples of second order partial differential equations we have

$$\nabla^2 u = u_{xx} + u_{yy} + u_{zz} = 0, \qquad \textit{elliptic type.} \tag{3.3.10}$$

$$u_{tt} - \nabla^2 u = u_{tt} - u_{xx} - u_{yy} - u_{zz} = 0, \qquad \textit{hyperbolic type.} \tag{3.3.11}$$

$$u_t - \nabla^2 u = u_t - u_{xx} - u_{yy} - u_{zz} = 0, \qquad \textit{parabolic type.} \tag{3.3.12}$$

These are the higher-dimensional Laplace, wave, and diffusion equations. The operator $\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$ is the *Laplacian operator* in three dimensions. A further example is given by *Schrödinger's equation*,

$$i\hbar u_t = -\frac{\hbar^2}{2m}\nabla^2 u + Vu, \tag{3.3.13}$$

where $i = \sqrt{-1}$, $\hbar$ is Planck's constant, m is the mass, and V is a given potential function. This equation is clearly of parabolic type.

Once we have exhibited how second order equations in n variables are to be classified, it is appropriate to ask whether there exist transformations of the independent variables (similar to those considered in Section 3.1) that bring (3.3.1) into a *canonical form*, say, of the type given in (3.3.6) but where the principal part has constant coefficients. It is shown in the exercises that this is not possible, in general, if $n \geq 3$ in (3.3.1).

If the coefficients a_{ij} of the principal part are all constants, simple canonical forms do exist. In fact [noting (3.3.4)], the linear transformation

$$\xi_i = \mathbf{r}_i^T \mathbf{x}, \qquad i = 1, \ldots, n, \tag{3.3.14}$$

where $\mathbf{x}^T = [x_1, \ldots, x_n]$, transforms the principal part of (3.3.1) into the form

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} u_{x_i x_j} = \sum_{i=1}^{n} \lambda_i u_{\xi_i \xi_i}, \tag{3.3.15}$$

where the λ_i are the (constant) eigenvalues of A. Clearly, an elementary further transformation of the variables $\xi_1, \ldots, \xi_n$, can reduce the principal part to a form in which the coefficients are either $0, +1$, or -1. The resulting partial differential equation is then in *canonical form*.

Characteristic Surfaces for Second Order PDEs

To determine the characteristics for (3.3.1) we look for surfaces $\phi(x_1, \ldots, x_n) =$ constant across which second derivatives of solutions of (3.3.1) can have discontinuities. This can be done by introducing n (independent) families of surfaces $\phi^{(i)}(x_1, \ldots, x_n) =$ constant (with $\phi^{(1)} \equiv \phi$) and the corresponding coordinate system $\eta_i = \phi^{(i)}(x_1, ..., x_n)$, $i = 1, \ldots, n$. Transforming from the x_i coordinates to the η_i coordinates in (3.3.1), we easily obtain, with $\eta_1 = \phi$,

$$\left[\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\phi_{x_i}\phi_{x_j}\right] u_{\phi\phi} + \cdots = 0, \tag{3.3.16}$$

where the dots represent second derivative terms in the variables $\eta_2, \ldots, \eta_n$, as well as first derivative terms and undifferentiated terms in all variables. Now $u_{\phi\phi}$ is a second derivative of u across the surface $\phi =$ constant. Thus if $\phi =$ constant satisfies the equation

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\phi_{x_i}\phi_{x_j} = 0, \tag{3.3.17}$$

$u_{\phi\phi}$ cannot be determined uniquely in terms of u_ϕ and u, as well as interior derivatives of u, given on the surface $\phi =$ constant. Consequently, the only surfaces across which u can have discontinuities in its second derivatives are solutions of (3.3.17). These surfaces are the *characteristic surfaces* for (3.3.1), and (3.3.17) is the *characteristic equation.*

For example, Laplace's equation (3.3.10) has the characteristic equation

$$\phi_x^2 + \phi_y^2 + \phi_z^2 = 0, \tag{3.3.18}$$

and this equation has no real solutions. Thus Laplace's equation has no real characteristics. This indicates that its solutions are smooth functions.

The three-dimensional wave equation (3.3.11) yields the characteristic equation

$$\phi_t^2 - \phi_x^2 - \phi_y^2 - \phi_z^2 = 0. \tag{3.3.19}$$

Among the important solutions of (3.3.19) are the *plane waves*,

$$\phi = \omega t - \kappa_1 x - \kappa_2 y - \kappa_3 z = \text{constant}, \tag{3.3.20}$$

where $\omega^2 = \kappa_1^2 + \kappa_2^2 + \kappa_3^2$, and a singular solution, the *characteristic cone*

$$\phi = (t - t_0)^2 - (x - x_0)^2 - (y - y_0)^2 - (z - z_0)^2 = 0, \tag{3.3.21}$$

where $P_0 = (x_0, y_0, z_0, t_0)$ is a given point.

The three-dimensional diffusion equation (3.3.12) has the characteristic equation

$$\phi_x^2 + \phi_y^2 + \phi_z^2 = 0, \tag{3.3.22}$$

which yields the characteristic surfaces

$$\phi(x,\ y,\ z,\ t) = t = \text{constant}. \tag{3.3.23}$$

Example 3.3. Classification of an Equation of Mixed Type. We consider the equation

$$u_{x_1x_1} + 2(1+cx_2)u_{x_2x_3} = u_{x_1x_1} + (1+cx_2)u_{x_2x_3} + (1+cx_2)u_{x_3x_2} = 0, \tag{3.3.24}$$

with constant c, and apply the classification procedure. From (3.3.2) we have

$$\partial_{x_1}^2 + (1+cx_2)\partial_{x_2}\partial_{x_3} + (1+cx_2)\partial_{x_3}\partial_{x_2} = \boldsymbol{\partial}_x^t A \boldsymbol{\partial}_x - c\partial_{x_3}, \tag{3.3.25}$$

where

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1+cx_2 \\ 0 & 1+cx_2 & 0 \end{bmatrix}, \qquad \boldsymbol{\partial}_x = \begin{bmatrix} \partial_{x_1} \\ \partial_{x_2} \\ \partial_{x_3} \end{bmatrix}. \tag{3.3.26}$$

The eigenvalues and the corresponding orthonormalized eigenvectors of A are

$$\lambda_1 = 1, \quad \mathbf{r}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}; \qquad \lambda_2 = (1+cx_2), \quad \mathbf{r}_2 = \begin{bmatrix} 0 \\ 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix};$$

$$\lambda_3 = -(1+cx_2), \quad \mathbf{r}_3 = \begin{bmatrix} 0 \\ 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}. \tag{3.3.27}$$

The orthogonal matrix R is

$$R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}, \tag{3.3.28}$$

and it may be verified that $R = R^T = R^{-1}$ and that

$$R^T A R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1+cx_2 & 0 \\ 0 & 0 & -1-cx_2 \end{bmatrix} = D. \tag{3.3.29}$$

Further,

$$\boldsymbol{\partial}_\xi = R\boldsymbol{\partial}_x = \begin{bmatrix} \partial_{x_1} \\ (\partial_{x_2} + \partial_{x_3})/\sqrt{2} \\ (\partial_{x_2} - \partial_{x_3})/\sqrt{2} \end{bmatrix} = \begin{bmatrix} \partial_{\xi_1} \\ \partial_{\xi_2} \\ \partial_{\xi_3} \end{bmatrix}, \tag{3.3.30}$$

so that since R is a constant matrix,

$$\partial_x^T A \partial_x = (R\partial_\xi)^T A(R\partial_\xi) = \partial_\xi^T D \partial_\xi$$

$$= \lambda_1 \partial_{\xi_1}^2 + \lambda_2 \partial_{\xi_2}^2 + \lambda_3 \partial_{\xi_3}^2 + \frac{c}{\sqrt{2}} \partial_{\xi_2} - \frac{c}{\sqrt{2}} \partial_{\xi_3}, \qquad (3.3.31)$$

where we have used (3.3.27) and (3.3.30).

In view of (3.3.11)–(3.3.12), we conclude that (3.3.24) is parabolic when $x_2 = -1/c$ (if $c \neq 0$) and is hyperbolic in the half-spaces $x_2 > -1/c$ and $x_2 < -1/c$. If $c = 0$, all the eigenvalues are constant with $\lambda_1 = \lambda_2 = 1$ and $\lambda_3 = -1$, so that (3.3.24) is hyperbolic everywhere. Then, on applying the transformation (3.3.14), we have

$$\xi_1 = x_1, \qquad \xi_2 = \frac{1}{\sqrt{2}}(x_2 + x_3), \qquad \xi_3 = \frac{1}{\sqrt{2}}(x_2 - x_3), \qquad (3.3.32)$$

and (3.3.24) takes on the canonical form

$$u_{\xi_1\xi_1} + u_{\xi_2\xi_2} - u_{\xi_3\xi_3} = 0. \qquad (3.3.33)$$

It is a wave equation in two dimensions with ξ_3 as the time variable.

First Order Systems of Linear PDEs: Classification and Characteristics

Next we consider the *classification* of first order systems of linear partial differential equations in two independent variables. They may be written as

$$A(x,y)\frac{\partial \mathbf{u}(x,y)}{\partial x} + B(x,y)\frac{\partial \mathbf{u}(x,y)}{\partial y} = C(x,y)\mathbf{u}(x,y) + \mathbf{d}(x,y), \qquad (3.3.34)$$

where $A(x,y)$, $B(x,y)$, and $C(x,y)$ are $n \times n$ matrix functions of x and y and $\mathbf{u}(x,y)$ and $\mathbf{d}(x,y)$ are n-component vectors. At least $A(x,y)$ or $B(x,y)$ is assumed to be nonsingular [i.e., $\det A(x,y) \neq 0$, $\det B(x,y) \neq 0$, or both determinants are not zero]. Our classification procedure is based on the properties of the characteristic curves that can occur for (3.3.34).

Now (3.3.34) represents a natural generalization of the linear first order equation (3.2.1) with the scalar dependent variable replaced by a vector variable. Consequently, if $A(x,y)$ is nonsingular, it is appropriate to formulate an initial value problem for (3.3.34) that assigns the value $\mathbf{u}(x,y) = \mathbf{f}(x)$ on the curve $y = h(x)$. Proceeding as in Section 3.2, the characteristic curves of (3.3.34) are defined to be those curves on which $\mathbf{u}_x(x,y)$ and $\mathbf{u}_y(x,y)$ cannot be specified uniquely (if, indeed, they can be determined at all) in terms of the given initial data. The alternative procedure presented in Section 3.2 whereby the characteristics were determined as the curves across which continuous solutions of (3.3.34) can have discontinuous first derivatives can also be applied here. It is considered in the exercises.

On the initial curve $y = h(x)$, we have $\mathbf{u}[x, h(x)] = \mathbf{f}(x)$. Thus

$$\frac{d\mathbf{u}(x, h(x))}{dx} = \mathbf{u}_x + h'(x)\mathbf{u}_y = \mathbf{f}'(x). \tag{3.3.35}$$

Solving for $\mathbf{u}_x$ in terms of $\mathbf{u}_y$ and substituting in (3.3.34) gives

$$(B - h'(x)A)\,\mathbf{u}_y = C\mathbf{f}(x) - A\mathbf{f}'(x) + \mathbf{d}, \tag{3.3.36}$$

where A, B, C, and $\mathbf{d}$ are evaluated on the curve $y = h(x)$. Now (3.3.36) represents a system of equations for $\mathbf{u}_y$, [on the curve $y = h(x)$]. The solution is unique only if the determinant of the coefficient matrix of is nonzero, i.e.,

$$\det(B - h'(x)A) \equiv |B - h'(x)A| \neq 0. \tag{3.3.37}$$

A curve $y = h(x)$ for which this determinant vanishes identically is called a *characteristic curve* for (3.3.34). If the initial conditions $\mathbf{u}(x, h(x)) = \mathbf{f}(x)$ are such that the system (3.3.36) has a (nonunique) solution, we have a *characteristic initial value problem*. The singularity of the matrix $B - h'A$ gives rise to a compatibility condition on the right side of (3.3.36) for a solution to exist, which, in turn, implies a restriction on the data $\mathbf{f}(x)$. Even if $\mathbf{f}(x)$ does not satisfy this condition, the curve $y = h(x)$ is called a *characteristic curve*.

To determine the full set of (possible) characteristic curves, we express the characteristic determinant (3.3.37) in the form

$$|B(x, y(x)) - y'(x)A(x, y(x))| = 0, \tag{3.3.38}$$

with $y = y(x)$ as the characteristic curve. The determinant is an nth-degree algebraic equation for $y'(x)$, and this is the *characteristic equation* for (3.3.34).

If all the roots $y'(x)$ of (3.3.38) are real and distinct in some region, the system (3.3.34) is said to be *strictly* or *totally hyperbolic* in that region. Denoting these roots by the functions $\omega_i(x, y)$, $i = 1, \ldots, n$, we obtain n families of characteristic curves as solutions of the equations

$$y'(x) = \omega_i(x, y), \qquad i = 1, \ldots, n. \tag{3.3.39}$$

If all the roots $\omega_i(x, y)$ of (3.3.38) are complex valued, the system (3.3.34) is said to be of *elliptic type*. Then there are no real characteristic curves as carriers of possible discontinuities in derivatives of solutions and the solutions are expected to be smooth functions. The case where some roots are real and others are complex will not be considered. When all the roots are real but one or more is a multiple root, additional conditions which we now consider must be given to determine if (3.3.34) is of hyperbolic or parabolic type.

If the matrix A in (3.3.34) is singular, (3.3.38) does not yield an nth-degree algebraic equation for $y'(x)$, as the lines x constant are characteristics in that case and they cannot be expressed in the form $y = y(x)$. Since A is singular, B must be nonsingular according to our assumptions. Then we consider the initial value problem

$\mathbf{u} = \mathbf{f}(y)$ on the curve $x = g(y)$. Proceeding as before and solving for $\mathbf{u}_x$ we find that if $x = x(y)$ is a solution of the *characteristic equation*

$$|A(x(y), y) - x'(y)B(x(y), y)| = 0, \tag{3.3.40}$$

the solution for $\mathbf{u}_x$ is not unique, if indeed it exists at all. This is an nth-degree algebraic equation for $x'(y)$ and each root determines a characteristic curve. The classification of the system (3.3.34) based on the roots $x'(y)$ of (3.3.40) parallels that given above based on the roots $y'(x)$ of (3.3.38) and will not be repeated here.

If we set $\phi = y - h(x) = 0$ or $\phi = x - g(y) = 0$, both characteristic equations (3.3.38) and (3.3.40) can be written as

$$|\phi_x A + \phi_y B| = 0. \tag{3.3.41}$$

This is often referred to as the *characteristic equation* for (3.3.34).

To complete the classification process we now assume that the matrix B in (3.3.34) is nonsingular throughout the region under consideration so that it has an inverse. Multiplying through by the inverse matrix B^{-1} in (3.3.34), we obtain an equation of the form

$$\mathbf{u}_y(x, y) + A(x, y)\mathbf{u}_x(x, y) = C(x, y)\mathbf{u}(x, y) + \mathbf{d}(x, y); \tag{3.3.42}$$

that is, we have effectively put $B = I$, the identity matrix. [With y replaced by t, (3.3.42) has the form of many time-dependent systems that occur in applications.] Then, with $B = I$ in (3.3.40) and λ_i $(i = 1, \ldots, n)$ as the eigenvalues of A counted with their multiplicities, we have $x'(y) = \lambda_i$. If all the eigenvalues are real and there exist n linearly independent eigenvectors $\mathbf{r}_i$ $(i = 1, \ldots, n)$ for the matrix A, the system (3.3.42) is of *hyperbolic type*. The multiplicity of the eigenvalues plays no role, but it is true that when all the eigenvalues are distinct, there exist n linearly independent eigenvectors. (Then the system is *strictly* or *totally hyperbolic*, as indicated above.) In particular, if A is a (real) symmetric matrix, there are always n linearly independent eigenvectors. We note that if there are multiple eigenvalues, there are fewer than n independent families of characteristic curves. However, if there are multiple (real) eigenvalues and fewer than n linearly independent eigenvectors for A, the system (3.3.42) is of *parabolic type*.

Systems of Hyperbolic Type

When the system (3.3.42) is of *hyperbolic type*, it can be transformed into the following *normal* or *characteristic form*. We form the matrix R whose column vectors $\mathbf{r}_1(x, y), \ldots, \mathbf{r}_n(x, y)$ are eigenvectors of $A(x, y)$. Then let

$$\mathbf{u}(x, y) = R(x, y)\mathbf{v}(x, y) \tag{3.3.43}$$

in (3.3.42). Since $R(x, y)$ is nonsingular, we have for $\mathbf{v}(x, y)$,

$$\mathbf{v}_y(x, y) + [R^{-1}(x, y)A(x, y)R(x, y)]\mathbf{v}_x(x, y) = \hat{C}(x, y)\mathbf{v}(x, y) + \hat{\mathbf{d}}(x, y), \tag{3.3.44}$$

where $R^{-1}(x,y)A(x,y)R(x,y)$ is a diagonal, not necessarily constant, matrix whose diagonal elements are the eigenvalues $\lambda_i(x,y)$ of $A(x,y)$. That is,

$$R^{-1}(x,y)A(x,y)R(x,y) = \begin{bmatrix} \lambda_1(x,y) & & 0 \\ & \ddots & \\ 0 & & \lambda_n(x,y) \end{bmatrix}. \tag{3.3.45}$$

This follows since $R^{-1}(x,y)A(x,y)R(x,y)$ is a similarity transformation which diagonalizes the matrix $A(x,y)$. Also, $\hat{C}(x,y) = R^{-1}(x,y)C(x,y)R(x,y) - R^{-1}(x,y)R_y(x,y) - R^{-1}(x,y)A(x,y)R_x(x,y)$, where $R_x(x,y)$ and $R_y(x,y)$ are derivatives of the matrix $R(x,y)$, and $\hat{\mathbf{d}}(x,y) = R^{-1}(x,y)\mathbf{d}(x,y)$. The representation (3.3.44) is often referred to as the *normal form* of the hyperbolic system.

Now each eigenvalue $\lambda_i(x,y)$ determines a characteristic direction $x'(y)$, via the equation $x'(y) = \lambda_i(x,y)$. As a result, the system (3.3.44) can be written in terms of the components $v_i(x,y)$ of $\mathbf{v}(x,y)$ as

$$\frac{dv_i}{dy} = \sum_{j=1}^{n} \gamma_{ij} v_j + \delta_j \quad \text{on} \quad \frac{dx}{dy} = \lambda_i, \qquad i = 1, \ldots, n. \tag{3.3.46}$$

The γ_{ij} and δ_i are determined from $\hat{C}(x,y)$ and $\hat{d}(x,y)$ in (3.3.44). Each $v_i(x,y)$ is differentiated along a characteristic curve determined from $x'(y) = \lambda_i(x,y)$. Therefore, this system is said to be in *characteristic form*.

In general, the system (3.3.46) must be solved simultaneously. If the equations in the system are not coupled, each equation can be solved separately along its characteristic curve. This is the case, for example, if $C = 0$ in (3.3.34) and A and B are constant matrices. Even if the equations cannot be solved separately, iteration methods can be applied that make use of the fact that the derivative terms are uncoupled.

Example 3.4. First Order Systems in Two Dependent Variables. In Section 2.1 the wave, diffusion, and Laplace's equations were reduced to first order systems. We now reconsider their classification in system form based on the foregoing discussion.

The system (2.1.2) for the *wave equation* may be written as

$$\mathbf{u}_t(x,t) + A\mathbf{u}_x(x,t) = C\mathbf{u}(x,t), \tag{3.3.47}$$

where

$$\mathbf{u}(x,t) = \begin{bmatrix} v(x,t) \\ u(x,t) \end{bmatrix}, \qquad A = \begin{bmatrix} -\gamma & 0 \\ 0 & \gamma \end{bmatrix}, \qquad C = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}. \tag{3.3.48}$$

Putting $y = t$ in (3.3.38), we have

$$|I - t'(x)A| = \begin{vmatrix} 1 + \gamma t'(x) & 0 \\ 0 & 1 - \gamma t'(x) \end{vmatrix} = 1 - [\gamma t'(x)]^2 = 0 \tag{3.3.49}$$

as the characteristic equation, because $B = I$, the identity matrix, in this case. Since $t'(x) = \pm 1/\gamma$, we obtain, as expected, the characteristic curves $x \pm \gamma t =$ constant. The roots of the characteristic equation are $\omega_1 = 1/\gamma$ and $\omega_2 = -1/\gamma$, and these are real and distinct. Thus the system (3.3.47) is *strictly hyperbolic.*

The *Cauchy-Riemann equations* (2.1.6) have the form

$$A\mathbf{u}_x(x, y) + B\mathbf{u}_y(x, y) = \mathbf{0}, \tag{3.3.50}$$

where

$$\mathbf{u}(x, y) = \begin{bmatrix} w(x, y) \\ u(x, y) \end{bmatrix}, \qquad A = I, \qquad B = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}. \tag{3.3.51}$$

Thus (3.3.38) becomes

$$|B - y'(x)A| = \begin{vmatrix} -y'(x) & -1 \\ 1 & -y'(x) \end{vmatrix} = 1 + (y'(x))^2 = 0. \tag{3.3.52}$$

Since $y'(x) = \pm i \ \ (i = \sqrt{-1})$, there are no real characteristics and the system (3.3.50) is *elliptic.*

The system (2.1.7) that represents the *diffusion equation* has the form

$$\hat{A}\mathbf{u}_x(x, t) + \hat{B}\mathbf{u}_t(x, t) = \hat{C}\mathbf{u}(x, t), \tag{3.3.53}$$

where

$$\mathbf{u}(x, t) = \begin{bmatrix} u(x, t) \\ v(x, t) \end{bmatrix}, \ \hat{A} = \begin{bmatrix} -D/2 & 0 \\ 0 & 1 \end{bmatrix}, \ \hat{B} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \ \hat{C} = \begin{bmatrix} -c & 0 \\ 1 & 0 \end{bmatrix}. \tag{3.3.54}$$

We multiply across in (3.3.53) by $\hat{A}^{-1} = \begin{bmatrix} -2/D & 0 \\ 0 & 1 \end{bmatrix}$ and obtain

$$\mathbf{u}_x(x, t) + B\mathbf{u}_t(x, t) = C\mathbf{u}(x, t), \tag{3.3.55}$$

where

$$B = \begin{bmatrix} 0 & -2/D \\ 0 & 0 \end{bmatrix}, \qquad C = \begin{bmatrix} 2c/D & 0 \\ 1 & 0 \end{bmatrix}. \tag{3.3.56}$$

This has the basic form of (3.3.35) and has the characteristic equation

$$|B - t'(x)I| = \begin{vmatrix} -t'(x) & -2/D \\ 0 & -t'(x) \end{vmatrix} = (t'(x))^2 = 0. \tag{3.3.57}$$

Thus $t'(x) = 0$ is a double root and the characteristic curves are $t =$ constant. Since the matrix B has only one linearly independent eigenvector $\mathbf{r} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ (as is easily shown), the system (3.3.55) is of *parabolic* type.

The system

$$\mathbf{u}_t(x,t) + \mathbf{u}_x(x,t) = C\mathbf{u}(x,t), \tag{3.3.58}$$

where

$$\mathbf{u}(x,t) = \begin{bmatrix} u(x,t) \\ v(x,t) \end{bmatrix}, \qquad C = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \tag{3.3.59}$$

is equivalent to the single equation

$$u_{xx}(x,t) + 2u_{xt}(x,t) + u_{tt}(x,t) - u(x,t) = 0. \tag{3.3.60}$$

Comparing (3.3.58) with (3.3.35), we set $t = y$ and $A = I$. The identity matrix I has the double real eigenvalues $\lambda_1 = \lambda_2 = 1$ and the linearly independent eigenvectors $\mathbf{r}_1 = \mathbf{i}$ and $\mathbf{r}_2 = \mathbf{j}$. Thus (3.3.58) is a *hyperbolic system* but it is not strictly hyperbolic. Yet according to the classification procedure of Section 3.1 [with $y = t$ in (3.1.1)], we have $A = B = C = 1$ in (3.3.60), so that $B^2 - AC = 0$ and (3.3.60) is of parabolic type. This example shows that the definition of hyperbolicity given above for systems of equations is somewhat more comprehensive than that given for single equations.

Higher-Order and Nonlinear PDEs

Next we turn to the consideration of a single linear partial differential equation of order m in n independent variables. We have

$$\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\cdots\sum_{i_m=1}^{n} a_{i_1,i_2,\cdots,i_m}\frac{\partial^m u}{\partial x_{i_1}\cdots\partial x_{i_m}} + \cdots = 0, \tag{3.3.61}$$

where we write only the principal part of the differential equation. The dots represent lower derivative and undifferentiated terms in $u(x_1,\ldots,x_n)$. Adapting the procedure given above for the second order equation (3.3.1), it is readily shown that the characteristic surfaces $\phi(x_1,\ldots,x_n)$ = constant are solutions of the characteristic equation

$$\sum_{i_1=1}^{n}\cdots\sum_{i_m=1}^{n} a_{i_1,\cdots,i_m}\phi_{x_{i_1}}\cdots\phi_{x_{i_m}} = 0. \tag{3.3.62}$$

These are surfaces in n-space across which mth-order derivatives of solutions of (3.3.61) can have discontinuities. If there are no real surfaces that satisfy (3.3.62), the equation (3.3.61) is of *elliptic type*. Classification into other types is somewhat more complicated and is not considered. However, specific equations of higher order are discussed in the text.

As an example we consider the *biharmonic equation* in two variables,

$$\nabla^2\nabla^2 u(x,y) = u_{xxxx}(x,y) + 2u_{xxyy}(x,y) + u_{yyyy}(x,y) = 0. \tag{3.3.63}$$

Its characteristic equation is

$$\phi_x^4 + 2\phi_x^2\phi_y^2 + \phi_y^4 = (\phi_x^2 + \phi_y^2)^2 = 0. \tag{3.3.64}$$

Clearly, it has no real solutions, so that (3.3.63) is *elliptic*.

Next, we consider a system of first order equations in n variables,

$$\sum_{i=1}^{n} A^i \frac{\partial \mathbf{u}}{\partial x_i} + B\mathbf{u} + \mathbf{c} = \mathbf{0}, \tag{3.3.65}$$

where A^i and B are $k \times k$ matrix functions of $x_1, \ldots, x_n$ and $\mathbf{u}$ and $\mathbf{c}$ are k-component column vector functions of $x_1, \ldots, x_n$. The characteristic surfaces for (3.3.65) are surfaces $\phi(x_1, \ldots, x_n) =$ constant on which the derivatives $\mathbf{u}_{x_i}$ $(i = 1, \ldots, n)$ cannot be specified uniquely (if at all) for given initial data $\mathbf{u} = \mathbf{f}$ on those surfaces. It can be shown that such surfaces must be solutions of the characteristic equations

$$\det\left[\sum_{i=1}^{n} A^i \frac{\partial \phi}{\partial x_i}\right] \equiv \left|\sum_{i=1}^{n} A^i \phi_{x_i}\right| = 0. \tag{3.3.66}$$

Again, if no real surfaces satisfying (3.3.66) exist, (3.3.65) is of *elliptic type*. We shall have occasion to characterize systems as being of *hyperbolic* or *parabolic type* in specific examples and exercises but do not classify them in general. We also do not discuss the classification of. systems of higher-order equations. Technically, they can always be reduced to first order systems.

Finally, we consider nonlinear equations and systems. If the principal parts of these equations are linear and the nonlinearities are confined to the lower order terms, the classification of these equations proceeds as above. For example, the equation

$$\nabla^2 u = e^u \tag{3.3.67}$$

is of *elliptic type*, while the system

$$\begin{cases} u_t(x,t) - \gamma u_x(x,t) = u^2(x,t) + v(x,t), \\ v_t(x,t) + \gamma v_x(x,t) = u(x,t)v(x,t), \end{cases} \tag{3.3.68}$$

where $\gamma =$ constant is of *hyperbolic type*.

As to equations where nonlinearities occur in the principal parts, we deal only with quasilinear equations. That is, the principal part is linear in the highest derivative terms, but it may have coefficients that contain lower derivative or undifferentiated terms. The classification proceeds as for the linear case, but it depends on the specific solution under consideration. The characteristics also depend on the specific solution. (This fact was observed in our discussion of first order quasilinear equations in Chapter 2.)

As an example we consider the system

$$\begin{cases} u_t(x,t) - v_x(x,t) = 0, \\ v_t(x,t) + c'(u(x,t))u_x(x,t) = 0, \end{cases} \tag{3.3.69}$$

where $c(u)$ is a given differentiable function. (This system occurs in a number of branches of continuum mechanics in one form or another. For example, it describes the motion of nonlinear elastic waves in a bar.) Identifying y with t, (3.3.69) has the form (3.3.42) with

$$\mathbf{u}(x,t) = \begin{bmatrix} u(x,t) \\ v(x,t) \end{bmatrix}, \quad A = \begin{bmatrix} 0 & -1 \\ -c'(u(x,t)) & 0 \end{bmatrix}, \quad C = 0, \ \mathbf{d} = \mathbf{0}. \tag{3.3.70}$$

The eigenvalues of the matrix A are given as $\lambda_1(x,t) = \sqrt{c'(u(x,t))}$ and $\lambda_2(x,t) = -\sqrt{c'(u(x,t))}$. If the solution $u(x,t)$ is such that $c'(u(x,t)) > 0$, we find that (3.3.69) is *hyperbolic*. If $c'(u(x,t)) = 0$, the system is clearly *parabolic*, while it is *elliptic* if $c'(u(x,t)) < 0$. In the hyperbolic case, the characteristic curves are determined from the equations $dx/dt = \pm\sqrt{c'(u(x,t))}$. They obviously depend on the particular solution $u(x,t)$ under consideration.

Putting $v(x,t) = w_t(x,t)$ and $u(x,t) = w_x(x,t)$ in (3.3.69), we obtain the second order quasilinear equation

$$w_{tt}(x,t) - c'(w_x(x,t))w_{xx}(x,t) = 0. \tag{3.3.71}$$

Considering a specific solution $w(x,t)$ in (3.3.71) and recalling the results of Section 3.1, we have

$$B^2(x,t) - A(x,t)C(x,t) = c'(w_x(x,t)). \tag{3.3.72}$$

Thus (3.3.71) is hyperbolic if $c'(w_x(x,t)) > 0$, parabolic if $c'(w_x(x,t)) = 0$, and elliptic if $c'(w_x(x,t)) < 0$. This is consistent with the results above for the system (3.3.69) since $u(x,t) = w_x(x,t)$.

Quasilinear First Order Systems and Normal Forms

We consider the quasilinear system

$$A(x,t,\mathbf{u})\mathbf{u}_x + B(x,t,\mathbf{u})\mathbf{u}_t = \mathbf{c}(x,t,\mathbf{u}), \tag{3.3.73}$$

where $\mathbf{u}$ and $\mathbf{c}$ are n-component vectors and the $n \times n$ matrices A and B, as well as $\mathbf{c}$, are functions of x, t, and $\mathbf{u}$. We assume that the matrix B is nonsingular and consider the eigenvalue problem

$$A^T\mathbf{r} = \lambda B^T\mathbf{r}, \tag{3.3.74}$$

where A^T and B^T are the transposes of the matrices A and B. Since B^T is nonsingular, this represents a generalized eigenvalue problem for A^T. [It can also be interpreted as a generalized eigenvalue problem for A with respect to B, in terms of left (row) eigenvectors $\mathbf{r}^T$ rather than right (column) eigenvectors $\mathbf{r}$. That is, we consider the eigenvalue problem $\mathbf{r}^T A = \lambda \mathbf{r}^T B$.] If all the eigenvalues λ_i are real and there exist n linearly independent eigenvectors $\mathbf{r}_i$, the system (3.3.73) is of *hyperbolic type*. If there are multiple (real) eigenvalues and fewer than n linearly independent

eigenvectors, the system (3.3.73) is of *parabolic type*. If all the eigenvalues are complex, it is of *elliptic type*. [We remark that this classification also applies to the linear system (3.3.34) if B is a nonsingular matrix.]

If the system (3.3.73) is *hyperbolic*, it is possible to reduce it to a simpler form known as the *characteristic normal form*. To do so, we multiply (3.3.73) on the left by the transpose of each of the n independent eigenvectors $\mathbf{r}_i$ determined above. This yields

$$\mathbf{r}_i^T B[\mathbf{u}_t + \lambda_i \mathbf{u}_x] \equiv \sum_{j=1}^{n} \sigma_{ij} \left[\frac{\partial u_j}{\partial t} + \lambda_i \frac{\partial u_j}{\partial x} \right] = \mathbf{r}_i^T \mathbf{c} \equiv \delta_i, \quad i = 1, ..., n. \tag{3.3.75}$$

Now the *characteristic curves* for this system are the solutions of $x'(t) = \lambda_i$. Thus, along each characteristic and for each eigenvector, we have

$$\sum_{j=1}^{n} \sigma_{ij} \frac{du_j}{dt} = \delta_i \qquad \text{on} \qquad \frac{dx}{dt} = \lambda_i, \qquad i = 1, ..., n. \tag{3.3.76}$$

This is the *characteristic normal form* for the system. In each of the n equations the components of $\mathbf{u}$ are differentiated in a single direction, the characteristic direction. Whereas a characteristic normal form can also be obtained for the linear system (3.3.34), it is not as strong a simplification of the linear system as was achieved in the *characteristic form* (3.3.46). In the characteristic form only a single unknown function is differentiated in each equation. In the following example and in the exercises, it will be shown that it can be possible to simplify the characteristic normal form even further for certain quasilinear hyperbolic systems.

Example 3.5. The Characteristic Normal Form. We construct the characteristic normal form for the quasilinear system (3.3.69) in the hyperbolic case. This system has the matrix form (3.3.73) with A and $\mathbf{u}$ defined as in (3.3.70). The matrix $B = I$ and the vector $\mathbf{c} = \mathbf{0}$. The eigenvalues for the problem (3.3.74) are the same as those for the eigenvalue problem for A. They were found to be $\lambda_1 = \sqrt{c'(u)}$ and $\lambda_2 = -\sqrt{c'(u)}$ in the preceding discussion of (3.3.69). [We assume that $c'(u)$ is positive.] Two generalized eigenvectors are easily determined to be $\mathbf{r}_1^T = [\sqrt{c'(u)}, -1]$ and $\mathbf{r}_2^T = [\sqrt{c'(u)}, 1]$.

To obtain the normal form, we multiply the equation $A\mathbf{u}_x + \mathbf{u}_t = \mathbf{0}$ by the eigenvectors $\mathbf{r}_1^T$ and $\mathbf{r}_2^T$. This yields

$$\sqrt{c'(u)} \left[u_t \pm \sqrt{c'(u)} u_x \right] \mp \left[v_t \pm \sqrt{c'(u)} v_x \right] = 0, \tag{3.3.77}$$

where the upper signs correspond to λ_1 and the lower signs to λ_2. Along the characteristics $x'(t) = \pm\sqrt{c'(u)}$ these equations can be written as

$$\sqrt{c'(u)} \frac{du}{dt} \mp \frac{dv}{dt} = 0 \qquad \text{on} \qquad \frac{dx}{dt} = \pm\sqrt{c'(u)}. \tag{3.3.78}$$

The equations (3.3.77) and (3.3.78) represent the *characteristic normal form* for the given system.

For this problem a further simplification is possible. Equation (3.3.78) can be written as

$$\frac{d}{dt}\left[\int^{u}\sqrt{c'(u)}\,du \mp v\right] = 0 \qquad \text{on} \qquad \frac{dx}{dt} = \pm\sqrt{c'(u)}. \tag{3.3.79}$$

This shows that the bracketed terms are constant on the corresponding characteristics. That is,

$$\int^{u}\sqrt{c'(u)}\,du \mp v = \text{constant} \qquad \text{on} \qquad \frac{dx}{dt} = \pm\sqrt{c'(u)}. \tag{3.3.80}$$

These terms are referred to as the *Riemann invariants* for this problem. Such invariants occur in a number of fluid mechanics problems and play an important role in their solution.

Exercises 3.3

3.3.1. Show that the equation $3u_{x_1x_1} - 2u_{x_1x_2} + 2u_{x_2x_2} - 2u_{x_2x_3} + 3u_{x_3x_3} + 5u_{x_2} - u_{x_3} + 10u = 0$ is of elliptic type by determining that the matrix A [see (3.3.2)] has the eigenvalues $\lambda_1 = 1$, $\lambda_2 = 3$, and $\lambda_3 = 4$. Determine a transformation (3.3.14) that yields the equation $u_{\xi_1\xi_1} + 3u_{\xi_2\xi_2} + 4u_{\xi_3\xi_3} + (9/\sqrt{6})u_{\xi_1} + (1/\sqrt{2})u_{\xi_2} - (6/\sqrt{3})u_{\xi_3} + 10u = 0$, for $u(\xi_1, \xi_2, \xi_3)$. Construct transformations of the independent and dependent variables that replace the coefficients of the second derivative terms by unity and eliminate the first derivative terms.

3.3.2. Show that the PDE $u_{x_1x_3}(x_1, x_2, x_3) = 0$ is of parabolic type. Use (3.3.14) to bring it into canonical form.

3.3.3. Classify the following equations into elliptic, parabolic, or hyperbolic type; (a) $u_{xx} + 2u_{yz} + \cos(x)u_z - \exp(y^2)u = \cosh(z)$; (b) $u_{xx} + 2u_{xy} + u_{yy} + 2u_{zz} + (1 + xy)u = 0$; (c) $7u_{xx} - 10u_{xy} - 22u_{yz} + u_{yy} - 16u_{xz} - 5u_{zz} = 0$; (d) $e^z u_{xy} - u_{xx} = \log[x^2 + y^2 + z^2 + 1]$.

3.3.4. Show that all linear second order equations of elliptic type with constant coefficients can be brought into the form $\sum_{i=1}^{n} u_{x_ix_i} + cu = F(x_1, \ldots, x_n)$.

3.3.5. Show that all linear second order equations of hyperbolic type (in $n + 1$ variables) with constant coefficients can be transformed into $\sum_{i=1}^{n} u_{x_ix_i} - u_{x_0x_0} + cu = F(x_0, x_1, \ldots, x_n)$.

3.3.6. Determine the regions where $u_{xx}(x, y, z) - 2x^2u_{xz}(x, y, z) + u_{yy}(x, y, z) + u_{zz}(x, y, z) = 0$ is of hyperbolic, elliptic, or parabolic type.

3.3.7. Demonstrate that the equations (a) $\nabla \cdot (p(\mathbf{x})\nabla u(\mathbf{x})) + q(\mathbf{x})u(\mathbf{x}) = F(\mathbf{x})$; (b) $\rho(\mathbf{x})u_t(\mathbf{x}, t) - \nabla \cdot (p(\mathbf{x})\nabla u(\mathbf{x}, t)) + q(\mathbf{x})u(\mathbf{x}, t) = F(\mathbf{x}, t)$, and (c) $\rho(\mathbf{x})u_{tt}(\mathbf{x}, t) -$

$\nabla \cdot (p(\mathbf{x})\nabla u(\mathbf{x},t)) + \lambda(\mathbf{x})u_t(\mathbf{x},t) + q(\mathbf{x})u(\mathbf{x},t) = F(\mathbf{x},t)$ are of elliptic, parabolic, and hyperbolic types, respectively, in a region R where $\rho(\mathbf{x})$ and $p(\mathbf{x})$ are positive.

3.3.8. Considering that if we disregard the order of differentiation, the function $u(\xi_1,\ldots,\xi_n)$ has exactly $\frac{1}{2}n(n-1)$ mixed partial derivatives of order n, show that if $n > 3$ there are, in general, an insufficient number of equations of transformation leading from (3.3.2) to (3.3.6) to eliminate the mixed partial derivative terms in $\xi_1,\ldots,\xi_n$ throughout a given region. If $n = 3$, the mixed partial derivative terms can be eliminated but the coefficients of the $u_{\xi_i\xi_i}$ terms cannot, in general, be made to equal plus or minus 1, as was done in the two-dimensional case.

3.3.9. Use the transformation (3.3.32) to bring (3.3.24) of Example 3.3 into the form $u_{\xi_1\xi_1} + [1 + (c/\sqrt{2})(\xi_2 + \xi_3)]u_{\xi_2\xi_2} - [1 + (c/\sqrt{2})(\xi_2 + \xi_3)]u_{\xi_3\xi_3} = 0$. Attempt a transformation of this equation to bring it into the form (3.3.33).

3.3.10. Show that $\mathbf{u}_y(x,y) + A\mathbf{u}_x(x,y) = \mathbf{0}$, with $A = \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix}$, is of parabolic type.

3.3.11. Show that the system $A\mathbf{u}_x(x,y) + B\mathbf{u}_y(x,y) = C\mathbf{u} + \mathbf{d}$, where $A = \begin{bmatrix} 1 & 0 \\ -3 & -1 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}$, is of elliptic type.

3.3.12. Determine the characteristic curves of the strictly hyperbolic system $\mathbf{u}_y(x,y) + A\mathbf{u}_x(x,y) = C\mathbf{u}(x,y)$, where

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 3 \\ 0 & 0 & -1 \end{bmatrix}, \qquad C = \begin{bmatrix} -2 & 1 & 5 \\ 0 & 3 & 7 \\ 1 & -3 & -10 \end{bmatrix}.$$

Reduce this system to the normal form (3.3.44).

3.3.13. Show that $\mathbf{u}_y(x,y) + A\mathbf{u}_x(x,y) = \mathbf{0}$, where $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$, is a hyperbolic system and reduce it to the normal form (3.3.44).

3.3.14. (a) Use the canonical form obtained in Exercise 3.3.13 to show that the general solution of the equation $\mathbf{u}_y(x,y) + A\mathbf{u}_x(x,y) = \mathbf{0}$ given in that exercise is $\mathbf{u}(x,y) = \begin{bmatrix} (f_1(x) + f_2(x-2y))/\sqrt{2} \\ (-f_1(x) + f_2(x-2y))/\sqrt{2} \\ f_3(x-2y) \end{bmatrix}$. (b) Solve the initial value problem for $\mathbf{u}_y(x,y) + A\mathbf{u}_x(x,y) = \mathbf{0}$ with $\mathbf{u}(x,0) = \begin{bmatrix} \sin x \\ 1 \\ e^x \end{bmatrix}$.

3.3.15. Show that $\mathbf{u}_y(x,y) + A\mathbf{u}_x(x,y) = \mathbf{0}$, where $A = \begin{bmatrix} 1 & 2e^x \\ 2e^{-x} & 1 \end{bmatrix}$ is strictly hyperbolic. Determine its characteristic curves and reduce it to the normal form (3.3.44).

3.3.16. If the constant coefficient system $\mathbf{u}_y(x,y) + A\mathbf{u}_x(x,y) = \mathbf{0}$ is hyperbolic [with $\mathbf{u}(x,y)$ as an n-component vector], show how to obtain a general solution by using the normal form (3.3.44).

3.3.17. Use the method given in the text that leads up to equation (3.2.7) to show that if $\mathbf{u}(x,y)$ is a continuous solution of (3.3.34) whose first derivatives have a jump across the curve $\xi = \phi(x,y) = 0$, the jump in the derivative across the curve satisfies the equation $(A\phi_x + B\phi_y)[\mathbf{u}_\xi] = \mathbf{0}$. Since the jump is nonzero by assumption, conclude that $\phi(x,y) = 0$ satisfies (3.3.41) and must be a characteristic curve.

3.3.18. Let $y = h(x)$ be a characteristic curve for the strictly hyperbolic system (3.3.34). Then the matrix $B - h'A$ is singular. Multiply on the left in (3.3.36) by a left nullvector of the singular matrix and determine a compatibility condition on the initial value $\mathbf{u} = \mathbf{f}(x)$ for the problem to be a characteristic initial value problem.

3.3.19. Consider the strictly hyperbolic system $\mathbf{u}_y(x,y) + A\mathbf{u}_x(x,y) = \mathbf{0}$ with $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$. Show that $y = -x$ and $y = x/3$ are characteristic curves for this system. If we set $\mathbf{u} = \mathbf{f}(x)$ on each of these curves, determine conditions on $\mathbf{f}(x)$ so that in each case we have a characteristic initial value problem. For appropriately chosen $\mathbf{f}(x)$, determine the (nonunique) solution in each case.

3.3.20. Reduce Euler's equations of one-dimensional isentropic fluid flow (see Section 8.5), $\{\rho_t + u\rho_x + \rho u_x = 0,\ u_t + (c^2/\rho)\rho_x + uu_x = 0\}$, to characteristic normal form. Also, determine the Riemann invariants for this system.

3.3.21. (a) Reduce the shallow water equations (see Section 10.3) $\{h_t + uh_x + hu_x = 0,\ u_t + uu_x + gh_x = 0\}$, where h is the height of the water above a horizontal bottom, u is its velocity, and g is the gravitational constant, to characteristic normal form. (b) Determine the Riemann invariants for this system.

3.3.22. Reduce Euler's equations of one-dimensional adiabatic flow (see Exercise 8.5.19), $\{\rho_t + u\rho_x + \rho u_x = 0,\ u_t + uu_x + (1/\rho)p_x = 0,\ p_t + up_x + c^2\rho u_x\}$, to characteristic normal form.

3.3.23. Show that $u_{ttt}(x,t) - c^2u_{xxt}(x,t) + \alpha(u_{tt}(x,t) - a^2u_{xx}(x,t)) = 0$ is of hyperbolic type and determine its characteristic curves. (It is of hyperbolic type if the characteristic curves are real and distinct.)

3.4 FORMULATION OF INITIAL AND BOUNDARY VALUE PROBLEMS

As seen in our discussion of the model equations derived in Chapter 1 and the first order PDEs of Chapter 2, we are not merely interested in finding arbitrary functions that satisfy the given differential equations. Rather, we ask for specific solutions that satisfy certain auxiliary conditions associated with the given problem.

The diffusion and telegrapher's or wave equations, which are of parabolic and hyperbolic types, respectively, both contain a time dependence. We have found it appropriate in our discussion in Chapter 1 to prescribe values of the solutions of these

equations at an initial time $t = 0$. For the diffusion equation, initial values for the density function $v(x,t)$ were assigned, and for the telegrapher's and wave equations $v(x,0)$ and $\partial v(x,0)/\partial t$ were prescribed. In case the values of x are unrestricted, this constitutes an *initial value problem* for each of these equations.

When the values of x are restricted to lie in a bounded or semi-infinite interval, $v(x,t)$, $\partial v(x,t)/\partial x$, or a linear combination of both must be prescribed on the boundaries for all $t \geq 0$, as seen in our discussion of the diffusion equation in Chapter 1. Such conditions are also appropriate for the telegrapher's or wave equation when x is similarly restricted. The combined prescription of $v(x,t)$ and/or its derivatives on the initial line $t = 0$ and on the boundary line(s) constitutes an *initial and boundary value problem* for each of these equations for $v(x,t)$.

Laplace's equation was shown in Chapter 1 to characterize an equilibrium or steady-state situation where time plays no role. The unknown function $v(x,y)$ was specified on the boundary of the region under consideration. This constitutes a *boundary value problem* for Laplace's equation.

As we have seen, the equations discussed in Chapter 1 are representative of second order equations of parabolic, hyperbolic, and elliptic type. As a general rule (in terms of the equations we shall consider), parabolic and hyperbolic equations are characteristic of problems that contain a time dependence. Initial value or initial and boundary value problems are appropriate for such equations, depending on whether spatial boundaries occur for these problems. Elliptic equations represent equilibrium or steady-state situations in regions with boundaries (in which time-dependent effects play no role), and boundary value problems are appropriate for such equations.

The number of initial and/or boundary conditions that should be assigned for a given differential equation depends on several factors. It may be stated that if the initial line or plane $t = 0$ is noncharacteristic and the equation contains k time derivatives, the function and its first $k - 1$ time derivatives must be prescribed at $t = 0$. This is valid if we deal with a vector or scalar function. The dependence of the number of boundary conditions on the order of the differential equations is more complicated and requires separate discussions for different classes of equations.

More generally, if the data for the problem are not given at $t = 0$ or on some spatial boundary for all time, it becomes necessary to examine if they have the character of initial or boundary data. This determination again depends on the given PDE. When data are given in a region that has the character of an initial curve or surface, the problem is known as a *Cauchy problem*. Also, for certain equations, boundary data are referred to as *Dirichlet* or *Neumann* or *Robin data*. They are discussed at appropriate places in the text. Further, we consider data given on characteristic curves and surfaces, and it is necessary to determine under what circumstances such problems can be solved. (Certain problems of this type have been encountered in Chapter 2 and in preceding sections of this chapter.)

Well-Posed Problems

The problem of deciding what form of initial and/or boundary data are appropriate for given partial differential equations is fairly complicated. A set of guidelines

was proposed by *Hadamard*, who listed three requirements that must be met when formulating an initial and/or boundary value problem. A problem for which the differential equation and the data lead to a solution satisfying these requirements is said to be *well posed* or *correctly posed*. If it does not meet these requirements, it is *incorrectly posed*.

Hadamard's conditions for a well-posed problem are:

1. *The solution must exist.*
2. *The solution must be determined uniquely.*
3. *The solution must depend continuously on the initial and/or boundary data.*

The first two conditions require that the equation plus the data for the problem must be such that one and only one solution exists. The third condition states that a slight variation of the data for the problem should cause the solution to vary only slightly. Thus, since data are generally obtained experimentally and may be subject to numerical approximations, we require that the solution be stable under small variations in initial and/or boundary values. We cannot permit wild variations to occur in the solution if the data are altered slightly. These are reasonable requirements for a problem arising in a physical context.

Thus for any given differential equation defined over a certain region, one must check whether the data assigned for the problem meet the *Hadamard criteria*. It was seen in Chapter 1 that the equations we derived were naturally associated with a set of initial and/or boundary conditions. These conditions are, in fact, appropriate for the differential equations that were considered and the problems given in Chapter 1 are well posed. It may well be argued that any problem arising in a physical context comes with built-in data, so that there is no need to decide which data are relevant. However, since any PDE representing a physical process is a mathematical model obtained, in general, under various simplifying assumptions, it is not a priori obvious that the formulation of the mathematical problem is reasonable or well posed.

The problems we consider in this book are sufficiently standard that their appropriate formulations are well understood. Nevertheless, we comment occasionally on questions of uniqueness or continuous dependence on data.

It should be noted that certain problems representing physical processes are, in fact, incorrectly posed in the sense of Hadamard, and this does not appear to be due to a weakness in the mathematical model. Such problems have been studied and methods for dealing with such problems are continuing to be developed. We do not consider such problems in this book.

We now present two examples of incorrectly posed problems and discuss further questions relating to this matter in Section 3.5.

Example 3.6. Incorrectly Posed Problems. Boundary value problems are, as a rule, not well posed for hyperbolic and parabolic equations. This follows because these are, in general, equations whose solutions evolve in time, and their behavior at later times is predicted by their previous states. Thus a boundary value problem that arbitrarily prescribes the solution at two or more separate times (at one or more given points in space) is not reasonable.

As a simple example, we consider the *hyperbolic equation* $u_{xy}(x,y) = 0$ in the square region $0 < x < 1$ and $0 < y < 1$ with boundary values assigned on the sides of the square. We show that this problem has no solution if the data are prescribed arbitrarily. Since $u_{xy}(x,y) = 0$ implies that $u_x(x,y)$ = constant, we have $u_x(x,0) = u_x(x,1)$. If we put $u(x,0) = f(x)$ and $u(x,1) = g(x)$, we have $u_x(x,0) = f'(x)$ and $u_x(x,1) = g'(x)$. Unless $f(x)$ and $g(x)$ are prescribed such that $f'(x) = g'(x)$, the boundary value problem cannot be solved, so it is *incorrectly posed.*

For *Laplace's equation*, the Cauchy problem is, in general, not well posed. This is shown by the following example credited to Hadamard. We consider the equation $u_{xx}(x,y) + u_{yy}(x,y) = 0$ in the region $y > 0$ with the Cauchy data $u(x,0) = 0$ and $u_y(x,0) = (\sin nx)/n$. The solution is easily obtained as $u(x,y) = (\sinh(ny)\sin(nx))/n^2$. Now, as $n \to \infty$, $u_y(x,0) \to 0$, so that for large n the Cauchy data $u(x,0)$ and $u_y(x,0)$ can be made arbitrarily small in magnitude. However, the solution $u(x,y)$ oscillates with an amplitude that grows exponentially like e^{ny} as $n \to \infty$. Thus arbitrarily small data can lead to arbitrarily large solutions and the solution is unstable. This violates the third condition of Hadamard, requiring continuous dependence of the solution on the data.

Exercises 3.4

3.4.1. Solve the Cauchy problem for Laplace's equation given in Example 3.6 by expanding $u(x,y)$ as a power series $u(x,y) = \sum_{k=0}^{\infty}[\partial^k u(x,0)/\partial y^k]y^k/k!$. Determine the derivatives $\partial^k u(x,0)/\partial y^k$ by differentiating Laplace's equation $u_{yy}(x,y) = -u_{xx}(x,y)$ along the initial line $y = 0$ and using the Cauchy data given there.

3.4.2. Construct the solution of the Cauchy problem for the heat equation $u_t(x,t) = c^2 u_{xx}(x,t)$ with the initial condition $u(x,0) = f(x)$, as a power series $u(x,t) = \sum_{n=0}^{\infty}[\partial^n u(x,0)/\partial t^n]t^n/n!$, assuming that $f(x)$ is sufficiently differentiable. [Determine the derivatives $\partial^n u(x,0)/\partial t^n$ from the heat equation and the initial condition.]

3.4.3. Use the method of Exercise 3.4.2 to show that the solution of $u_t(x,t) = c^2 u_{xx}(x,t)$, $u(x,0) = \cos(x)$ is given as $u(x,t) = \cos(x)\exp(-c^2 t)$.

3.4.4. Obtain a solution of the telegrapher's equation $u_{tt}(x,t) - \gamma^2 u_{xx}(x,t) + 2\lambda u_t(x,t) = 0$, where γ and λ are positive constants, in the form of a power series in t if the Cauchy data are given as $u(x,0) = \cos x$, $u_t(x,0) = 0$.

3.4.5. Show that the solution of the initial value problem for the wave equation (2.1.1) is continuously dependent on the initial data (2.2.13), by showing that over the finite time interval $0 \le t \le T$, d'Alembert's solution (2.2.17) is uniformly small in magnitude if the same is true for the initial values $f(x)$ and $g(x)$.

3.4.6. Consider the wave equation $u_{tt}(x,t) = u_{xx}(x,t)$ in the square $0 < x < \pi$, $0 < t < \pi$, with $u(x,t) = 0$ on the boundary. Show that $u(x,t) = \sin(x)\sin(t)$

is a solution of the equation that vanishes on the boundary. Explain why the boundary value problem for the wave equation in the given square with $u = f$ on the boundary is not well posed.

3.4.7. Use the solution $u(x,t) = (1/n)\exp(\rho n^2 t)\sin(nx)$ of the backward heat equation $\rho u_{xx} + u_t = 0$ with $\rho > 0$ to show that the initial and boundary value problem for that equation in the interval $0 < x < \pi$ with $t > 0$, where $u(x,t)$ is prescribed on $x = 0$, $x = \pi$, and $t = 0$, is not well posed.

3.4.8. Consider the exterior boundary value problem for Laplace's equation $\nabla^2 u(x,y,z) = 0$, with $\rho^2 = x^2 + y^2 + z^2 > a^2$ and $u(x,y,z)$ prescribed on the sphere $\rho = a$ as $u|_{\rho=a} = A$ (A = constant). Show that it has (at least) two solutions, $u = A$, $u = Aa/\rho$. (The solution can be shown to be unique if we require that $u \to 0$ as $\rho \to \infty$.)

3.4.9. Apply the divergence theorem to the expression $\nabla \cdot \nabla u(x,y)$ to show that the boundary value problem for $\nabla^2 u(x,y) = 0$ in the region A with $\partial u(x,y)/\partial n = f(x,y)$ on the boundary ∂A has no solution unless $\int_{\partial A} f\, d\sigma = 0$. Also, observe that if a solution does exist for this problem, it is not unique since the solution $u = a$ (a = constant) can be added to the given result.

3.5 STABILITY THEORY, ENERGY CONSERVATION, AND DISPERSION

In this section we consider certain general properties of partial differential equations that distinguish equations of different types beyond or apart from the classification process given in previous sections. We recall that classification depends only on the form of the *principal part* of the equation. Here we examine the role played by the *lower order terms* in the equation in determining the behavior of the solutions.

For simplicity we restrict our discussion to second order linear PDEs with constant coefficients. The general ideas carry over to higher order linear equations and systems of equations with constant coefficients. These matters are considered in the exercises. Also, certain aspects can be generalized to apply to equations with variable coefficients and to nonlinear equations as indicated later in the book.

Normal Modes and Well-Posedness

Given the linear second order homogeneous equation in two variables

$$Au_{xx} + 2Bu_{xt} + Cu_{tt} + Du_x + Eu_t + Fu = 0, \tag{3.5.1}$$

where $A, B, \ldots, F$ are real constants and $u = u(x,t)$, we look for exponential solutions of the form

$$u(x,t) = a(k)\exp[ikx + \lambda(k)t], \tag{3.5.2}$$

where $a(k)$ is a constant and $i = \sqrt{-1}$. The parameter k is assumed to be real and $\lambda(k)$ must be chosen such that (3.5.2) satisfies (3.5.1). The solutions $u(x,t)$ for each

k are called *normal modes* of the equation (3.5.1). Since the coefficients in (3.5.1) are assumed to be real, the real and imaginary parts of the normal modes (3.5.2) are also solutions of (3.5.1). (If α is a complex-valued quantity and is written as $\alpha = \alpha_1 + i\alpha_2$, where α_1 and α_2 are real-valued, then $\alpha_1 = \mathrm{Re}[\alpha]$ and $\alpha_2 = \mathrm{Im}[\alpha]$ are the real and imaginary parts of α, respectively.)

We assume that t represents time and discuss the behavior of the normal modes in the region $t \geq 0$. As shown in later chapters, the solution of an initial value problem for (3.5.1) with initial data at $t = 0$ can be represented as a superposition of normal modes either by the use of Fourier series or integrals. Thus the behavior of the solution of the initial value problem is largely determined by that of the normal modes.

Inserting (3.5.2) into (3.5.1) yields the quadratic equation

$$C\lambda^2 + (2iBk + E)\lambda + (-Ak^2 + iDk + F) = 0. \tag{3.5.3}$$

Solving for $\lambda = \lambda(k)$, we let $\lambda(k)$ represent either of the two possible solutions of (3.5.3). Then the normal mode can be written as

$$u(x,t) = a(k)\exp[i(kx + \mathrm{Im}[\lambda(k)]t)]\exp[\mathrm{Re}[\lambda(k)]t]. \tag{3.5.4}$$

The magnitude of $u(x,t)$ is given as

$$|u(x,t)| = |a(k)|\exp[\mathrm{Re}[\lambda(k)]t]. \tag{3.5.5}$$

We assume that $|a(k)|$ is bounded for all k and see that the growth of $|u(x,t)|$ as t increases from zero is determined by the expression $\mathrm{Re}[\lambda(k)]$.

We consider two possibilities. Either Re $[\lambda(k)]$ is bounded above for all real k or it is unbounded. We define the constant Ω to be the least upper bound (denoted as *lub*) of Re $[\lambda(k)]$ as k ranges through all its values; that is,

$$\Omega = lub\ \mathrm{Re}[\lambda(k)], \qquad -\infty < k < \infty. \tag{3.5.6}$$

If $\mathrm{Re}[\lambda(k)]$ is not bounded above, we set $\Omega = +\infty$ and this is the case we now study. The case when $\Omega < \infty$ is considered later.

Assuming that $\Omega = +\infty$, we put $a(k) = 1/\lambda(k)^2$ in the normal mode (3.5.2) and consider the initial value problem for the normal mode with data at $t = 0$. On evaluating (3.5.2) and its derivative at $t = 0$, we have $u(x,0) = \left[1/\lambda(k)^2\right] e^{ikx}$, $u_t(x,0) = [1/\lambda(k)]\, e^{ikx}$. Since $\mathrm{Re}[\lambda(k)]$ is unbounded, there exist values of k for which $|u(x,0)| = 1/|\lambda(k)|^2$ and $|u_t(x,0)| = 1/\lambda(k)$ are arbitrarily small, yet $|u(x,t)|$ [as given in (3.5.5)] is arbitrarily large for any $t > 0$. This follows since $1/\lambda(k)$ decays algebraically, whereas $\exp[\mathrm{Re}[\lambda(k)]t]$ grows exponentially. Thus *Hadamard's stability criterion* for the initial value problem for (3.5.1) is violated if $\Omega = +\infty$, and the initial value problem is not well posed. If $\Omega < +\infty$, it can be shown that the initial value problem for (3.5.1) is well posed.

Before considering the case where $\Omega < +\infty$, we apply the foregoing discussion to specific PDEs of elliptic, parabolic, and hyperbolic types.

For the *elliptic equation* $u_{xx}(x,t) + u_{tt}(x,t) + \rho u(x,t) = 0$ with ρ = constant, $\lambda(k)$ in the normal mode solution (3.5.2) has the form $\lambda(k) = \pm\sqrt{k^2 - \rho}$, and we readily find that $\Omega = +\infty$. Thus, the Cauchy problem for this PDE with data at $t = 0$ is not well posed.

The *parabolic equation* $\rho u_{xx}(x,t) + u_t(x,t) = 0$, where ρ = constant, has the form of the *diffusion* or *heat equation* if $\rho < 0$ and is known as the *backward diffusion* or *heat equation* if $\rho > 0$. One form can be obtained from the other if the time direction is reversed; that is, if t is replaced by $-t$. Normal mode solutions for this PDE yield $\lambda(k) = \rho k^2$. Now if $\rho < 0$, we have $\Omega = 0$, whereas if $\rho > 0$, we have $\Omega = +\infty$, so that the Cauchy problem is not well posed if $\rho > 0$.

The *hyperbolic equation* $u_{tt}(x,t) - u_{xx}(x,t) + \rho u(x,t) = 0$, where ρ = constant, has normal mode solutions (3.5.2) with $\lambda(k)$ given as $\lambda(k) = \pm i\sqrt{k^2 + \rho}$. For sufficiently large $|k|$, if $\rho < 0$, we have $\mathrm{Re}[\lambda(k)] = 0$. With $\rho \geq 0$, $\mathrm{Re}[\lambda(k)]$ vanishes for all k. Thus $\Omega < +\infty$ and the Cauchy problem is well posed.

In each of the foregoing cases we have $|\lambda(k)| \to \infty$ as $|k| \to \infty$, so that if data of the above form are chosen, $|u(x,0)|$ and $|u_t(x,0)|$ can be made arbitrarily small by choosing $|k| \gg 1$. But if $\Omega < +\infty$, $|u(x,t)|$ is also small since $\exp(\mathrm{Re}[\lambda(k)]t)$ does not grow unboundedly with k for fixed values of t.

Stability

The constant Ω defined in (3.5.6) is known as the *stability index* for the differential equation. We have already seen that when $\Omega = +\infty$, solutions of the initial value problem for (3.5.1) are unstable. That is, solutions whose magnitude is initially arbitrarily small can grow arbitrarily large even at finite times. In the following discussion we assume that $\Omega < +\infty$ and ask whether solutions of initially bounded magnitude can grow unboundedly as $t \to \infty$. Equations for which this can happen are said to be *unstable*. Otherwise, they are said to be *stable*. Thus even if the Cauchy problem is well posed, the equation may be unstable.

Physically, instability indicates that even in the absence of external effects due to forcing terms in (3.5.1), internal mechanisms generate a growth in the solution as time increases. If the Cauchy problem is not well posed (i.e., $\Omega = +\infty$), the validity of the mathematical model must be reexamined. If the Cauchy problem is well posed, instability plays an important role if the equation was derived by a linearization procedure under which it was assumed that solutions remain small in magnitude for all time. Examples of stability analyses where the present results play a role are given later in the book.

Given the normal mode solution (3.5.2) and its magnitude (3.5.5), we note the following results. If $\Omega < 0$, then $|u(x,t)| \to 0$ as $t \to \infty$ for all k. If $\Omega > 0$, there are normal modes with k near Ω, for which $|u(x,t)| \to \infty$ as $t \to \infty$. If $\Omega = 0$, there may be values of k for which $|u(x,t)|$ is bounded but does not tend to zero as $t \to \infty$, although $|u(x,t)| \to 0$ for the remaining values of k. [We remark that since Ω is a least upper bound, it may happen that $\mathrm{Re}[\lambda(k)]$ never equals zero even if $\Omega = 0$.]

Noting the preceding discussion, we introduce the following classification for (3.5.1). If $\Omega < 0$, the equation is said to be *strictly stable*. If $\Omega > 0$, the equation is *unstable*. If $\Omega = 0$, the equation is said to be *neutrally stable*, but it may be unstable in this case.

For example, the parabolic equation

$$u_t(x,t) - c^2 u_{xx}(x,t) + a u_x(x,t) + b u(x,t) = 0 \tag{3.5.7}$$

has $\lambda(k) = -k^2c^2 - b - ika$, so that $\Omega = -b$. It is stable if $b > 0$, unstable if $b < 0$, and neutrally stable if $b = 0$.

A simple example indicating the problems that can arise in the neutrally stable case is given by the equation

$$u_{xx}(x,t) + 2u_{xt}(x,t) + u_{tt}(x,t) = 0, \tag{3.5.8}$$

for which $\lambda(k) = -ik$ so that $\mathrm{Re}[\lambda(k)] = 0 = \Omega$. The equation has solutions

$$u(x,t) = ae^{ik(x-t)} + bte^{ik(x-t)}, \tag{3.5.9}$$

where a and b are arbitrary constants. In particular, the solution of the Cauchy problem with $u(x,0) = 0$ and $u_t(x,0) = 1$ is given by $u(x,t) = t$. Thus even though the data are uniformly bounded in magnitude, the solution grows unboundedly as $t \to \infty$. It is generally true that for a neutrally stable case, if there is instability, the growth of the solution will be algebraic rather than exponential. [We remark that the second term on the right of (3.5.9) is not a normal mode solution of (3.5.8) as defined in (3.5.2).]

It must again be emphasized that for the data at $t = 0$ of the normal mode solution (3.5.2), we have $|u(x,0)| = |a|$ and $|u_t(x,0)| = |a\lambda(k)|$. The data are uniformly bounded for all x and fixed k. However, $|u(x,t)| = |a| \exp[\mathrm{Re}[\lambda(k)t]]$ and if $\Omega < +\infty$, $|u(x,t)|$ can tend to infinity only as $t \to \infty$. This contrasts with the case where $\Omega = +\infty$ and $|u(x,t)|$ can become arbitrarily large for fixed t.

Energy Conservation and Dispersion

If $\mathrm{Re}[\lambda(k)] = 0$ for all k, (3.5.1) is said to be an equation of *conservative type*. Then we have $|u(x,t)| = |a| = |u(x,0)|$ in view of (3.5.5), so that the amplitude $|u|$ of the normal mode solution is constant in time. Since $|u|^2$ is generally a measure of the energy of the normal mode solution, we find that the energy is conserved in time and say that the equation (3.5.1) is *conservative*.

In the conservative case, $\lambda(k)$ may be expressed as $\lambda(k) = -i\omega(k)$, where $\omega(k)$ is real valued for all k. The normal mode solution then takes the form

$$u(x,t) = a(k) \exp\left(i[kx - \omega(k)t]\right). \tag{3.5.10}$$

Inserting (3.5.10) into (3.5.1) yields

$$\omega = \omega(k), \tag{3.5.11}$$

which is known as the *dispersion relation* for the differential equation. If $\omega(k)$ is real valued for all real k as we have assumed and, additionally, $\omega''(k) \not\equiv 0$ [i.e., $\omega(k)$ is not a linear function of k], the equation (3.5.1) is said to be of *dispersive type*. [We note that regardless of the assumptions made above on $\omega(k)$, (3.5.11) is known as the dispersion relation for (3.5.1).]

Recalling the discussion in Example 2.2, we find that (3.5.10), for fixed k, has the form of a wave that travels with velocity $dx/dt = \omega(k)/k$. The term $\theta = kx - \omega(k)t$ in (3.5.10) is called the *phase* of the normal mode and $dx/dt = \omega(k)/k$ is the *phase velocity*. If $\omega''(k) \equiv 0$, we assume that $\omega(0) = 0$ and find that $\omega(k) = ck$, where c = constant, so that the phase velocity $dx/dt = ck/k = c$ has a constant value c for all normal modes. Since the general solution can be constructed as a superposition of normal modes it will also represent a wave traveling with velocity c. [In general, $\omega(k)$ has two values, so that the general solution is a sum of two waves.]

However, if $\omega''(k) \not\equiv 0$, the phase velocity $dx/dt = \omega(k)/k$ will be different for different values of k so that different normal modes have different velocities. The general solution obtained as a superposition of the normal modes yields a *wave* that *disperses* since its components or modes all travel at different velocities. We shall see in Section 5.7 that the relevant velocity for *dispersive wave motion* is not the *phase velocity* $\omega(k)/k$, but the *group velocity*, defined as $d\omega(k)/dk$.

Dissipation

Finally, if the stability index $\Omega \leq 0$ and $\mathrm{Re}[\lambda(k)]$ is negative for all except a finite number of values of k, the equation (3.5.1) is said to be of *dissipative type*. In that case, (3.5.5) shows that $|u(x,t)|$ decays to zero as $t \to \infty$ for all but a finite number of values of k and the energy $|u(x,t)|^2$ is dissipated as $t \to \infty$. When $\Omega < 0$, all solutions tend to zero as $t \to \infty$. However, if $\Omega = 0$, those modes for which $\mathrm{Re}[\lambda(k)] = 0$ do not decay as $t \to \infty$ and they are expected to constitute the major contribution to the solution for large values of t. This fact is demonstrated when dissipative equations are studied in Section 5.7.

We now consider two examples of equations of dissipative and dispersive type. The telegrapher's equation derived in Chapter 1 is an equation of dissipative type. The Klein-Gordon equation which occurs in relativistic physics is an equation of dispersive type. Both these equations are commonly used as model equations of dissipative and dispersive type.

Example 3.7. The Telegrapher's and Klein-Gordon Equations. The *telegrapher's equation* (1.2.25), that is,

$$u_{tt}(x,t) - \gamma^2 u_{xx}(x,t) + 2\hat{\lambda} u_t(x,t) = 0 \tag{3.5.12}$$

[where λ has been replaced by $\hat{\lambda}$ to avoid confusion with $\lambda = \lambda(k)$], yields for the normal mode solutions (3.5.2) $\lambda(k) = -\hat{\lambda} \pm \sqrt{\hat{\lambda}^2 - \gamma^2 k^2}$. Since $\hat{\lambda} > 0$ by assumption, the real part of $\lambda(k)$ is easily seen to be negative for all $k \neq 0$. At $k = 0$

either $\lambda(0) = 0$ or $\lambda(0) = -2\hat{\lambda}$, so that $\Omega = 0$ for (3.5.12). Consequently, the telegrapher's equation is neutrally stable and is of dissipative type. Note that when $\hat{\lambda} = 0$, (3.5.12) reduces to the wave equation and $\lambda(k) = \pm i\gamma k$ in that case. Then $\mathrm{Re}[\lambda(k)] = 0$ for all k, so that the wave equation is also neutrally stable and is of conservative type. For this reason, when $\hat{\lambda} > 0$, (3.5.12) is often called the *damped wave equation.*

All normal mode solutions of (3.5.12) are damped as $t \to \infty$ except for the solution corresponding to $\lambda(0) = 0$. It is shown in Section 5.7 that the main contribution to the solution of the Cauchy problem for (3.5.12) as $t \to \infty$ comes from the normal modes with $k \approx 0$.

The *Klein-Gordon equation* has the form

$$u_{tt}(x,t) - \gamma^2 u_{xx}(x,t) + c^2 u(x,t) = 0, \tag{3.5.13}$$

where γ and c are constants. The normal modes (3.5.2) are $\lambda(k) = \pm i\sqrt{\gamma^2 k^2 + c^2}$, so that the dispersion relation is $\omega = \omega(k) = \sqrt{\gamma^2 k^2 + c^2}$. Since $\omega(k)$ is real valued for all k and $\omega''(k) \not\equiv 0$, we conclude that, in addition to being neutrally stable, the Klein-Gordon equation is of conservative and dispersive type. Again, it may be noted that if $c = 0$ in (3.5.13), it reduces to the wave equation for which $\omega(k) = \gamma k$. Since $\omega''(k) = 0$ in this case, the wave equation is not of dispersive type.

To conclude our discussion of stability theory, it must be emphasized that the foregoing definitions are relevant, in general, only for the initial value problem for (3.5.1). When initial and boundary value problems are considered, the parameter k may be restricted to a discrete set of values. The solution of the problem is then given as a superposition of a discrete set of normal modes that correspond to the aforementioned values of k. Stability is then defined in terms of the growth or decay of these normal modes. If one or more of the normal modes is unbounded as $t \to \infty$, the problem is unstable. If all the normal modes decay as $t \to \infty$, the problem is stable. Examples of stability analyses for initial and boundary value problems are given later in the text. The concept of Von Neumann stability for finite difference schemes, which is introduced in Chapter 11, is closely related to our presentation in this section.

Exercises 3.5

3.5.1. Show that the stability index Ω for the hyperbolic equation $u_{tt}(x,t) - \gamma^2 u_{xx}(x,t) - c^2 u(x,t) = 0$ is given as $\Omega = c$. (The constant c is assumed to be positive.) Thus although the Cauchy problem for this equation is well posed, the equation is unstable.

3.5.2. Consider the hyperbolic equation $u_{tt}(x,t) - c^2 u_{xx}(x,t) + u_t(x,t) - \alpha u_x(x,t) = 0$. By examining the solutions $\lambda = \lambda(k)$ of (3.5.3) as specialized to the above

equation, show that the equation is unstable if $c^2 < \alpha^2$. *Hint*: Look at small values of k.

3.5.3. Demonstrate that the following equations are all of dispersive type by considering normal mode solutions of the form (3.5.2). Also, determine the relevant dispersion relations. (a) $u_{tt}(x,t) - \gamma^2 u_{xx}(x,t) - \alpha^2 u_{xxtt}(x,t) = 0$; (b) $u_{tt}(x,t) + \gamma^2 u_{xxxx}(x,t) = 0$; (c) $u_t(x,t) - \alpha u_x(x,t) + \beta u_{xxx}(x,t) = 0$.

3.5.4. Show that the diffusion equation (1.1.15) is of dissipative type.

3.5.5. Given the hyperbolic equation $u_{ttt}(x,t) - \gamma^2 u_{xxt}(x,t) + u_{tt}(x,t) - c^2 u_{xx}(x,t) = 0$, obtain the relationship $\lambda = \lambda(k)$ for the normal mode solutions (3.5.2) of the equation above. Using the known criterion that the polynomial $P(\lambda) = \lambda^3 + a_1\lambda^2 + a_2\lambda + a_3$ has roots with negative real parts if $a_3 > 0$, $a_1 > 0$, and $a_1 a_2 > a_3$, conclude that the given hyperbolic equation is neutrally stable and is of dissipative type if $\gamma^2 > c^2$.

3.5.6. Given the system of equations $A\mathbf{u}_x(x,t) + B\mathbf{u}_t(x,t) + C\mathbf{u}(x,t) = \mathbf{0}$, where A, B, and C are constant $n \times n$ matrices and $\mathbf{u}(x,t)$ is an n-component vector, normal mode solutions are given as $\mathbf{u}(x,t) = \mathbf{a}(k)\exp[ikx + \lambda(k)t]$, where $\mathbf{a}(k)$ is a constant vector. Show that for $\mathbf{u}(x,t)$ to be a solution of the given equation, we must have $|ikA + AB + C| = 0$. Explain how concepts of well-posedness and stability may be introduced for the given system on the basis of the result above.

3.5.7. Apply the method of Exercise 3.5.6 to the system (1.2.19)–(1.2.20) and to the systems given in Example 3.4 [replacing y by t in the Cauchy- Riemann system (3.3.50)]. Discuss stability and well-posedness questions for these systems and compare the results with those obtained for the scalar equivalents of these systems.

3.5.8. With $\mathbf{k} = [k_1, k_2, k_3]$ and $\mathbf{x} = [x, y, z]$, obtain the appropriate equation $\lambda = \lambda(\mathbf{k})$ for the normal mode solutions $u(\mathbf{x},t) = \mathbf{a}(k)\exp[i\mathbf{k}\cdot\mathbf{x} + \lambda(k)t]$ of the following equations: (a) $u_{tt}(\mathbf{x},t) - \gamma^2\nabla^2 u(\mathbf{x},t) + \hat{\lambda}u_t(\mathbf{x},t) = 0$; (b) $u_{tt}(\mathbf{x},t) - \gamma^2\nabla^2 u(\mathbf{x},t) + c^2 u(\mathbf{x},t) = 0$; (c) $u_{tt}(\mathbf{x},t) + \nabla^2 u(\mathbf{x},t) = 0$; (d) $u_t(\mathbf{x},t) - \gamma^2\nabla^2 u(\mathbf{x},t) = 0$, where $\hat{\lambda} > 0$, all the coefficients are constants, and ∇^2 is the Laplacian operator in three dimensions. Discuss well-posedness and stability for these equations in the manner of Section 3.5.

3.6 ADJOINT DIFFERENTIAL OPERATORS

The concept of *adjoint differential operators* has been encountered in connection with each of the continuum limits of random walk problems considered in Chapter 1. In the study of time-dependent random walk problems it was noted that the forward and backward differential equations are adjoints of one another. For time-independent problems, it was indicated that Laplace's equation and its generalizations and the equations for the corresponding Green's functions involve adjoint differential operators. In each case the equations and their adjoints were derived independent of one another. However, a key relationship between them involves the use of Green's

theorem and its generalizations. This was demonstrated in Example 1.6 for Laplace's equation and the corresponding Green's function.

The basic result that is used in obtaining Green's theorem and its generalizations is the *divergence theorem*. Thus, if ∇ is the gradient operator in space and L is Laplace's operator $L = \nabla^2$, we have

$$wL[u] - uL[w] = w\nabla^2 u - u\nabla^2 w = \nabla \cdot [w\nabla u - u\nabla w], \tag{3.6.1}$$

where the term on the right is a *divergence expression*. On integrating (3.6.1) over a bounded region, the divergence theorem expresses the integral of the left side of (3.6.1) in the interior of the region in terms of the integral of the normal component of the vector on the right in (3.6.1) over the boundary.

However, if the operator L is given as $L = \partial/\partial t - \nabla^2$, we find that to obtain a divergence expression in the manner of (3.6.1), it is necessary to introduce a new operator $L^* = -\partial/\partial t - \nabla^2$ in terms of which

$$wL[u] - uL^*[w] = \tilde{\nabla} \cdot [-w\nabla u + u\nabla w,\ wu], \tag{3.6.2}$$

where $\tilde{\nabla} = [\nabla,\ \partial/\partial t]$ is the gradient operator in space-time. For two space dimensions, $\tilde{\nabla}$ has the form $[\partial/\partial x,\ \partial/\partial y,\ \partial/\partial t]$. [The divergence theorem in space-time can be applied to equation (3.6.2).] The operator L^* is called the *adjoint operator* of L.

As shown above, if L is Laplace's operator the expression $wL[u] - uL^*[w]$ is in divergence form with $L^* = L = \nabla^2$. Since the adjoint operator is identical to Laplace's operator, the Laplacian is a *self-adjoint operator*. Equations in which these differential operators appear are called *self-adjoint* or *nonself-adjoint equations* according as the operators are self-adjoint or not self-adjoint.

More precisely, if L^* is an operator for which $wL[u] - uL^*[w]$ is a divergence expression, it is the *formal adjoint operator* of L, and if $L^* = L$, the operator L is *formally self-adjoint*. (The same is true for the related differential equations.) As will be seen in our discussion of eigenvalue problems and Green's functions later in the book, the boundary (or initial and boundary) conditions associated with the given equation and its formal adjoint also play a role in determining whether or not the problem is self-adjoint. Thus it is possible for an operator to be formally self-adjoint but for the problem associated with this operator and its adjoint to be nonself-adjoint. But if the operator is not self-adjoint, the problem cannot be self-adjoint. In this section, however, we deal only with the adjointness of the operators themselves, and we now present a formula for the adjoint of the most general linear second order differential operator.

Scalar PDEs

The general *linear second order differential operator* L can be written as

$$L[u] = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\frac{\partial^2 u}{\partial x_i\,\partial x_j} + \sum_{i=1}^{n} b_i\frac{\partial u}{\partial x_i} + cu, \tag{3.6.3}$$

where the a_{ij}, b_i, and c are functions of $x_1, \ldots, x_n$. The *adjoint differential operator* L^* is given as

$$L^*[w] = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2(a_{ij}w)}{\partial x_i\, \partial x_j} - \sum_{i=1}^{n} \frac{\partial(b_i w)}{\partial x_i} + cu. \tag{3.6.4}$$

It can be verified directly that

$$wL[u] - uL^*[w] = \sum_{i=1}^{n} \frac{\partial P_i}{\partial x_i}, \tag{3.6.5}$$

where

$$P_i = \sum_{j=1}^{n} \left[a_{ij} w \frac{\partial u}{\partial x_j} - u \frac{\partial(a_{ij}w)}{\partial x_j} \right] + b_i uw. \tag{3.6.6}$$

The expression on the right of (3.6.5) is a *divergence expression*, and the n-dimensional form of the divergence theorem applied to the region G yields

$$\iint_G \{wL[u] - uL^*[w]\}\, dv = \int_{\partial G} \mathbf{P} \cdot \mathbf{n}\, ds, \tag{3.6.7}$$

where $\mathbf{P}$ is a vector with n components $P_1, \ldots, P_n$ and $\mathbf{n}$ is the exterior unit normal vector to the boundary ∂G. As has already been indicated, the possibility of using the divergence theorem as in (3.6.7) is a key reason for the importance of the adjoint operator L^*.

To see how the adjoint operator L^* is constructed, we consider a typical term. We have

$$\begin{aligned} wa_{11} \frac{\partial^2 u}{\partial x_1^2} &= wa_{11} \frac{\partial}{\partial x_1}\left[\frac{\partial u}{\partial x_1}\right] = \frac{\partial}{\partial x_1}\left[wa_{11}\frac{\partial u}{\partial x_1}\right] - \frac{\partial(a_{11}w)}{\partial x_1}\frac{\partial u}{\partial x_1} \\ &= \frac{\partial}{\partial x_1}\left[wa_{11}\frac{\partial u}{\partial x_1}\right] - \frac{\partial}{\partial x_1}\left[u\frac{\partial(a_{11}w)}{\partial x_1}\right] + u\frac{\partial^2(a_{11}w)}{\partial x_1^2}. \end{aligned} \tag{3.6.8}$$

This yields

$$wa_{11}\frac{\partial^2 u}{\partial x_1^2} - u\frac{\partial^2(a_{11}w)}{\partial x_1^2} = \frac{\partial}{\partial x_1}\left[a_{11}w\frac{\partial u}{\partial x_1} - u\frac{\partial(a_{11}w)}{\partial x_1}\right]. \tag{3.6.9}$$

By carrying out all the differentiations in (3.6.4) and comparing L^* with L, it is easy to determine that if $b_i = \sum_{j=1}^{n} \partial a_{ij}/\partial x_j$, $i = 1, \ldots, n$, the operator L is self-adjoint. As a result, a first order partial differential operator L [i.e., (3.6.3) with $a_{ij} = 0$ for all i and j] is never self-adjoint. In the self-adjoint case the operator L takes the form

$$L[u] = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial}{\partial x_i}\left(a_{ij}\frac{\partial u}{\partial x_j}\right) + cu. \tag{3.6.10}$$

We observe that if the operator L has constant coefficients, it cannot have any first derivative terms if it is to be self-adjoint. The foregoing results can be specialized, by putting $n = 1$, to apply to ODEs as well.

Example 3.8. Adjoint Differential Operators. In this example we consider three differential operators of elliptic, hyperbolic, and parabolic types (related PDEs are studied throughout the book) and determine their adjoints.

We begin with the *elliptic operator* L defined as

$$Lu = -\nabla \cdot (p\nabla u) + qu, \tag{3.6.11}$$

where $p > 0$. Since $wLu - uLw = \nabla \cdot [-pw\nabla u + pu\nabla w]$, we see that the operator L is self-adjoint.

For the *hyperbolic operator* $\tilde{L}$ defined as

$$\tilde{L} = \rho u_{tt} + Lu, \tag{3.6.12}$$

with Lu given as in (3.6.11), we have $w\tilde{L}u - u\tilde{L}w = \tilde{\nabla} \cdot [-pw\nabla u + pu\nabla w,\ \rho w u_t - \rho u w_t]$, where $\tilde{\nabla}$ is the space-time gradient operator. (The positive function ρ is assumed to be independent of t.) Thus, the operator $\tilde{L}$ is self-adjoint.

For the *parabolic operator* $\hat{L}$ defined as

$$\hat{L}u = \rho u_t + Lu, \tag{3.6.13}$$

where L is defined as in (3.6.11), we have $w\hat{L}u - u\hat{L}^* w = \tilde{\nabla} \cdot [-pw\nabla u + pu\nabla w,\ \rho w u]$, where $\tilde{\nabla}$ is the space-time gradient operator, and the adjoint operator $\hat{L}^*$ is given as $\hat{L}^* = -\rho u_t + Lu$. Thus, in the parabolic case, the operator $\hat{L}$ is not self-adjoint.

Systems of PDEs

Adjoints can also be defined for systems of equations. We consider only linear systems in two variables. Let the matrix operator M be defined as

$$M[\mathbf{u}(x,y)] = A(x,y)\mathbf{u}_x(x,y) + B(x,y)\mathbf{u}_y(x,y) + C(x,y)\mathbf{u}(x,y), \tag{3.6.14}$$

where $A(x,y)$, $B(x,y)$, and $C(x,y)$ are $n \times n$ matrix functions of x and y and $\mathbf{u}$ is an n-component vector. We define the adjoint operator M^* as $M^*[\mathbf{w}(x,y)] = -(A^T(x,y)\mathbf{w}(x,y))_x - (B^T(x,y)\mathbf{w}(x,y))_y + C^T(x,y)\mathbf{w}(x,y)$, where the superscript T denotes the transpose of the matrix or the vector. With this choice of M^* we obtain $\mathbf{w}^T M\mathbf{u} - \mathbf{u}^T M^*\mathbf{w} = \nabla \cdot [\mathbf{w}^T A\mathbf{u},\ \mathbf{w}^T B\mathbf{u}]$, where ∇ is the two-dimensional gradient vector. That is, we obtain a divergence expression and this motivates the definition of the adjoint operator. If $A = -A^T$, $B = -B^T$ (i.e., A and B are antisymmetric matrices), $A_x^T + B_y^T = 0$, and $C = C^T$, the operator M is *self-adjoint*. Some examples of systems of equations and their adjoints are considered in the exercises.

Quasilinear PDEs

If the equations or systems are quasilinear, the foregoing cannot be used to define adjoint operators because the dependence of the coefficients on the unknown function(s) invalidates the results, as is easily seen. Nevertheless, it is useful to obtain results similar to those given above for linear problems, for the purpose of defining *weak* or *generalized solutions* of quasilinear equations. This can be accomplished if the equation is given in *conservation form*. We consider only a single equation in two variables.

Let Lu be defined as

$$Lu(x,t) = \tilde{\nabla} \cdot \mathbf{u}(x,t) = u_1(x,t)_x + u_2(x,t)_t, \tag{3.6.15}$$

where $u_1(x,t) = f(x,t,u(x,t))$ and $u_2(x,t) = g(x,t,u(x,t))$ with $u(x,t)$ as the unknown function. ($\tilde{\nabla} = [\partial/\partial x, \partial/\partial t]$ is the gradient operator in space-time.) Then it is easily verified that

$$wLu + \mathbf{u} \cdot \tilde{\nabla} w = (wu_1)_x + (wu_2)_t = \tilde{\nabla} \cdot (w\mathbf{u}). \tag{3.6.16}$$

We do not explicitly define an expression that serves as the adjoint of Lu, but note that (3.6.16) does express the left side of the equation in divergence form. Thus, if we integrate (3.6.16) over a bounded region R in (x,t)-space and w vanishes on the boundary ∂R, we have

$$\iint_R wLu \, dx \, dt = -\iint_R \mathbf{u} \cdot \tilde{\nabla} w \, dx \, dt. \tag{3.6.17}$$

As an example, we consider the inviscid Burgers' equation in conservation form, $(u)_t + (\frac{1}{2}u^2)_x = 0$, as in (2.3.29). Here, $u_1 = \frac{1}{2}u^2$ and $u_2 = u$. Then

$$wLu + \mathbf{u} \cdot \tilde{\nabla} w = \left(\frac{1}{2} w u^2\right)_x + (wu)_t. \tag{3.6.18}$$

On the basis of (3.6.18), we can define *weak solutions* of $Lu = 0$ that are not differentiable everywhere as piecewise continuous solutions for which the integral on the right in (3.6.18) vanishes for each w. Weak solutions of linear equations are discussed from this point of view in Section 6.4.

Exercises 3.6

3.6.1. Verify the result (3.6.5).

3.6.2. Determine the adjoint operators for the following operators:
(a) $Lu(x,y) = e^x u_{xx}(x,y) + x^2 u_{xy}(x,y) + y u_y(x,y) - 10u(x,y)$.
(b) $Lu(x,y) = a(x,y)u_x(x,y) + b(x,y)u_y(x,y) + c(x,y)u(x,y)$.
(c) $Lu(x,y,t) = u_t(x,y,t) - c^2[u_{xx}(x,y,t) + u_{yy}(x,y,t)]$.

3.6.3. Show that the backward Kolmogorov equation (1.1.34) is the adjoint of the Fokker-Planck equation (1.1.32).

3.6.4. Show that the differential operator in (1.3.23) is the adjoint of the operator in (1.3.21).

3.6.5. Show that the equation derived in Exercise 1.2.16 is the adjoint of equation (1.2.44).

3.6.6. Show that the system derived in Exercise 1.2.15 is the adjoint of the system (1.2.41)–(1.2.42).

3.6.7 Determine the adjoints of the systems in Example 3.4.

3.6.8. Verify that (3.6.16) is valid for the equation (2.3.29).

3.6.9. Obtain an adjoint operator L^* for the second order linear ordinary differential operator L given as $Ly(x) = a(x)y''(x) + b(x)y'(x) + c(x)y(x)$.

3.6.10. Obtain an adjoint operator L^* for the third order operator L given as $Lu(x,t) = u_t(x,t) - u_x(x,t) + \gamma^2 u_{xxx}(x,t)$, such that $wLu - uL^*w$ is a divergence expression.

3.6.11. Show that the operator $Lu(x,t) = u_{tt}(x,t) + u_{xxxx}(x,t)$ is self-adjoint by expressing $wLu - uLw$ in divergence form.

3.6.12. Demonstrate that $Lu(\mathbf{x},t) = u_{tt}(\mathbf{x},t) + \nabla^2\nabla^2 u(\mathbf{x},t)$ is self-adjoint by expressing $wLu - uLw$ in divergence form.

3.7 MAPLE METHODS

We have constructed a number of *Maple procedures* that automate the *factorization* and *classification* techniques for scalar PDEs and systems of PDEs that were discussed in this chapter. Procedures that determine the *well-posedness* and *stability* of scalar PDEs and systems of PDEs are also presented. A procedure that determines adjoint operators is given in Chapter 7, where these operators play an important role in the construction of Green's functions.

Classification of Equations and Canonical Forms

The procedure *ClassPDE2d* determines the discriminant of a second order PDE [as in (3.1.5)–(3.1.7)] and *classifies* the equation as being of *hyperbolic, elliptic* or *parabolic type*, if possible. The PDE may be of *mixed type*.) It is given as $ClassPDE2d(a, 2b, c, [x, y])$ and has the output

$$\text{Principal Part} = a\,\frac{\partial^2 u}{\partial x^2} + 2\,b\,\frac{\partial^2 u}{\partial x \partial y} + c\,\frac{\partial^2 u}{\partial y^2},\ \text{Discriminant} = b^2 - ac. \quad (3.7.1)$$

Clearly, a, $2b$, and c are the coefficients of the second derivative terms in the PDE and $[x, y]$ are the independent variables.

As a specific example, we consider $ClassPDE2d(-x, 0, 1, [x, y])$. The principal part of the PDE is $-x\, u_{xx} + u_{yy}$ (it represents an equation of mixed type) and the output is Discriminant = x, *Hyperbolic*, if it is assumed that $x > 0$ by entering the Maple command $assume(x > 0)$. If the command $assume(x < 0)$ is entered, the output is Discriminant = x, *Elliptic*. If Maple can determine the sign of the discriminant, the PDE is classified.

The procedure $PDEFactor2d$ factors linear second order PDEs and finds a *canonical form*, if feasible. Maple's *dsolve* procedure is used to determine the characteristic curves. It has the form $PDEFactor2d(a, 2b, c, d, e, f, [x, y])$, where a, $2b$, c, d, e, f are the coefficients of the linear PDE $au_{xx} + 2bu_{xy} + cu_{yy} + du_x + eu_y + fu = 0$, with the independent variables $[x, y]$. The procedure exhibits the discriminant as given in (3.1.5)–(3.1.7) and should be used only if Maple can determine if the discriminant is positive, negative, or zero.

As an example. we apply the procedure to the equation (3.1.28) of *mixed type* considered in Example 3.1, that is,

$$u_{xx}(x, y) + yu_{yy}(x, y) = 0. \tag{3.7.2}$$

If $y < 0$, the PDE is of *hyperbolic type* and the output of $PDEFactor2d(1, 0, y, 0, 0, 0, [x, y])$ gives the characteristic curves $\xi = -x + 2\sqrt{-y}$, $\eta = -x - 2\sqrt{-y}$, which correspond to those given in (3.1.32)–(3.1.33) but are defined differently here. In addition, the variables α and β are specified as $\alpha = -x$, $\beta = 2\sqrt{-y}$. They are defined by $\xi = \alpha + \beta$, $\eta = \alpha - \beta$ as in (3.1.20).

In terms of the variables (ξ, η), the *canonical form* is given as $u_{\xi\eta}(\xi, \eta) + 1/(2(\xi - \eta))(u_\xi(\xi, \eta) - u_\eta(\xi, \eta)) = 0$, in agreement with (3.1.35). The alternative canonical form is $u_{\alpha\alpha}(\alpha, \beta) - u_{\beta\beta}(\alpha, \beta) + (1/\beta)\, u_\beta(\alpha, \beta) = 0$. These represent the two types of canonical forms for hyperbolic equations. The procedure also gives the factored form of the principal part of the PDE (3.7.2) as $\partial_x^2 + y\partial_y^2 = (\partial_x - \sqrt{-y}\, \partial_y)(\partial_x + \sqrt{-y}\, \partial_y)$.

With $y > 0$, the PDE is of *elliptic type* and the output of $PDEFactor2d(1, 0, y, 0, 0, 0, [x, y])$ is $\xi = \sqrt{y} + \frac{1}{2}ix$, $\eta = \sqrt{y} - \frac{1}{2}ix$; $\alpha = \sqrt{y}$, $\beta = x/2$. The variables (ξ, η) are complex and we select the real variables (α, β) defined as in (3.1.26). The corresponding *canonical form* is given as $u_{\alpha\alpha}(\alpha, \beta) + u_{\beta\beta}(\alpha, \beta) - (1/\alpha)\, u_\alpha(\alpha, \beta) = 0$. The result differs from the canonical form (3.1.40) because α and β are defined differently here, but the results are equivalent.

The PDE (3.7.2) is of parabolic type when $y = 0$. Rather than apply the procedure in that case we consider a PDE that is of *parabolic type* everywhere. We invoke the procedure $PDEFactor2d(y^2, 2y, 1, 0, 0, 0, [x, y])$ that corresponds to the equation $y^2 u_{xx}(x, y) + 2yu_{xy}(x, y) + u_{yy}(x, y) = 0$. The output states that the discriminant is zero and exhibits the factored principal part as $y^2\partial_x^2 + 2y\partial_{xy}^2 + \partial_{yy}^2 = y^2(\partial_x + \frac{1}{y}\, \partial_y)^2$. The transformation variables ξ and η of (3.1.24) are given as $\xi = -2x + y^2$, $\eta = -x + y^2$, with ξ = constant as the family of characteristics, and the canonical form is $u_{\eta\eta}(\xi, \eta) + 2/(2\eta - \xi)[u_\xi(\xi, \eta) + u_\eta(\xi, \eta)] = 0$.

The foregoing procedures deal only with second order PDEs in two independent variables. The procedure $ClassPDEScalar$ classifies linear second order PDEs

with constant coefficients for any number of independent variables and exhibits their canonical forms. The process presented in Section 3.3 is followed. The relevant eigenvalues and eigenvalues are exhibited. (Maple's *linalg* and *LinearAlgebra* packages are used.) The equation is classified into *hyperbolic, parabolic,* or *elliptic* type. The transformation variables are shown and the (canonical) form of the PDE is displayed. For the PDE in three independent variables

$$\sum_{i=1}^{3}\sum_{j=1}^{3} a_{ij}u_{x_ix_j} + \sum_{i=1}^{3} b_i u_{x_i} + cu = 0, \tag{3.7.3}$$

the classification procedure has the form $ClassPDEScalar([[a_{11}, a_{12}, a_{13}], [a_{21}, a_{22}, a_{23}], [a_{31}, a_{32}, a_{33}]], [b_1, b_2, b_3], c, [x_1, x_2, x_3])$ where the coefficients of the PDE exhibited in the first three arguments of the procedure.

We consider the PDE

$$2u_{xx} + 8u_{yx} - 12u_{zx} + 2u_{yy} - 12u_{zy} - 15u_{zz} + 3u_x - u_y + 2u_z + 5u = 0. \tag{3.7.4}$$

$ClassPDEScalar([[2, 8, -12], [0, 2, -12], [0, 0, -15]], [3, -1, 2], 5, [x, y, z])$ determines that (3.7.4) is of *hyperbolic type*. Further, it yields as the eigenvalues and eigenvectors $[-18, 1, [1, 1, 4]]$, $[9, 1, [-2, -2, 1]]$, $[-2, 1, [1, -1, 0]]$. That is, the eigenvalues are $\{-18,\ 9,\ -2\}$, each of which has multiplicity 1. The eigenvector that corresponds to the eigenvalue -18 is given as $[1, 1, 4]$, for example. The transformation variables are given as $\xi_1 = (\sqrt{2}/6)(x + y + 4z)$, $\xi_2 = (z - 2x - 2y)/3$, $\xi_3 = (\sqrt{2}/2)(x - y)$. The simplified (canonical) form is given as

$$-18\frac{\partial^2 u}{\partial {\xi_1}^2} + 9\frac{\partial^2 u}{\partial {\xi_2}^2} - 2\frac{\partial^2 u}{\partial {\xi_3}^2} + \frac{5\sqrt{2}}{3}\frac{\partial u}{\partial \xi_1} - \frac{2}{3}\frac{\partial u}{\partial \xi_2} + 2\sqrt{2}\frac{\partial u}{\partial \xi_3} + 5u = 0. \tag{3.7.5}$$

(An additional, elementary change of variables, whereby the coefficients of the second derivative terms in (3.7.5) are replaced by $+1$ or -1, is needed to actually obtain the canonical form. We do not carry this out.)

Classification and Solution of Linear Systems

system of linear PDEs in two independent variables

$$A(x, y)\mathbf{u}_x(x, y) + B(x, y)\mathbf{u}_y(x, y) = C(x, y)\mathbf{u}(x, y) + \mathbf{d}(x, y), \tag{3.7.6}$$

and finds normal and characteristic forms for the system in the hyperbolic case. [The matrices $A(x, y)$, $B(x, y)$, $C(x, y)$ are all $n \times n$ matrices and $\mathbf{u}(x, y)$ and $\mathbf{d}(x, y)$ are n-vectors. It is assumed that $B(x, y)$ has an inverse and y can be replaced by t or another variable in the system.] The procedure transforms the system (3.7.6) into a form with $B = I$, where I is the identity matrix. The analysis then proceeds as in Section 3.3. The procedure determines a set of eigenvalues and eigenvectors as in Section 3.3. If all the eigenvalues are complex, the system is *elliptic*. If they are real and each

eigenvalue of multiplicity k has exactly k linearly independent eigenvectors, the system is *hyperbolic*. If all the eigenvalues are real and at least one multiple eigenvalue of multiplicity k has fewer than k linearly independent eigenvectors, the system is *parabolic*. The procedure takes the form $ClassSyst(A, B, C, \mathbf{d}, [x_1, \ldots, x_n])$.

$ClassSyst([[1,0],[-3,-1]],[[2,1],[-1,-1]],[[0,0],[0,0]],[0,0],[x,y])$ has

$$A = \begin{bmatrix} 1 & 0 \\ -3 & -1 \end{bmatrix}, \qquad B = \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}, \tag{3.7.7}$$

in (3.7.6) with $C = 0$ and $\mathbf{d} = \mathbf{0}$. The system is transformed into (3.7.6) with $A = \begin{bmatrix} -2 & -1 \\ 5 & 2 \end{bmatrix}$, $B = I$, the identity matrix, $C = 0$, and $\mathbf{d} = \mathbf{0}$. The eigenvalues are the imaginary numbers i and $-i$, with unit multiplicity. The system is of *elliptic type* and the procedure identifies it as such, with eigenvalues and eigenvectors $[i, 1, [1, -2 - i]]$, $[-i, 1, [1, -2 + i]]$.

The use of $ClassSyst([[0,-1],[0,0]],[[1,0],[0,1]],[[0,0],[0,0]],[0,0],[x,t])$ yields $A = \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix}$ and $B = I$, the identity matrix and $C = 0$, $\mathbf{d} = \mathbf{0}$, with y replaced by t in (3.7.6). The procedure finds that 0 is a double eigenvalue and that the system is of *parabolic type*.

Invoking $ClassSyst([[0,-1],[-x^2,0]],[[1,0],[0,1]],[[0,0],[0,0]],[0,0],[x,t])$ yields $A = \begin{bmatrix} 0 & -1 \\ -x^2 & 0 \end{bmatrix}$, $B = I$, the identity matrix, and $C = 0$, $\mathbf{d} = \mathbf{0}$, with y replaced by t in (3.7.6). (It is assumed that $x \neq 0$.) The eigenvalues and eigenvectors are given as $[x, 1, [1, -x]]$, $[-x, 1, [1, x]]$. The system is of *hyperbolic type*. The matrix R in the transformation $\mathbf{u} = R\mathbf{v}$, whose columns are the eigenvectors found above, is given as $R = \begin{bmatrix} 1 & 1 \\ -x & x \end{bmatrix}$. The procedure determines the *normal form* (3.3.37) as

$$\frac{\partial \mathbf{v}(x,t)}{\partial t} + \begin{bmatrix} x & 0 \\ 0 & -x \end{bmatrix} \frac{\partial \mathbf{v}(x,t)}{\partial x} = \begin{bmatrix} -1/2 & 1/2 \\ -1/2 & 1/2 \end{bmatrix} \mathbf{v}(x,t) + \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{3.7.8}$$

With $[v_1(x,t),\ v_2(x,t)]$ and $[u_1(x,t),\ u_2(x,t)]$ as the components of $\mathbf{v}(x,t)$ and $\mathbf{u}(x,t)$, respectively, we obtain $\mathbf{v}(x,t) = \begin{bmatrix} u_1(x,t)/2 - u_2(x,t)/2x \\ u_1(x,t)/2 + u_2(x,t)/2x \end{bmatrix}$, $\mathbf{u}(x,t) = \begin{bmatrix} v_1(x,t) + v_2(x,t) \\ -xv_1(x,t) + xv_2(x,t) \end{bmatrix}$. Finally, the *characteristic form* is given as

$$\begin{cases} dv_1(t)/dt = -v_1(t)/2 + v_2(t)/2 & \text{on} \quad dx(t)/dt = x(t), \\ dv_2(t)/dt = -v_1(t)/2 + v_2(t)/2 & \text{on} \quad dx(t)/dt = -x(t). \end{cases} \tag{3.7.9}$$

The procedure $HypSystExplicit$ finds exact solutions for hyperbolic systems with constant coefficients that contain only principal parts. They have the form $A\mathbf{u}_x(x,t) + B\mathbf{u}_t(x,t) = \mathbf{0}$. On reducing the system to normal or characteristic form, the first order equations in the system are uncoupled and can be solved exactly.

The procedure is applied to an initial value problem for the system above with initial conditions given at the time $t = t_0$. It takes the form $HypSystExplicit(A, B, [x, t], t = t_0, IC)$, with IC as the initial data.

$HypSystExplicit(A, B, [x, t], t = 0, [f(x), g(x)])$, with A, B, and $\mathbf{u}$ as

$$A = \begin{bmatrix} 1 & 4 \\ -3 & 2 \end{bmatrix}, \; B = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}, \; \mathbf{u}(x,t) = \begin{bmatrix} u_1(x,t) \\ u_2(x,t) \end{bmatrix}, \; \mathbf{u}(x,0) = \begin{bmatrix} f(x) \\ g(x) \end{bmatrix}, \tag{3.7.10}$$

yields the exact solution

$$u_1(x,t) = f(x-7t), \; u_2(x,t) = \frac{3}{5}f(x-2t) + g(x-2t) - \frac{3}{5}f(x-7t). \tag{3.7.11}$$

Quasilinear Hyperbolic Systems in Two Independent Variables

The procedure $ClassHypSystChar(A, B, \mathbf{c}, \mathbf{u}, [x, t])$ determines a characteristic normal form for the quasilinear system (3.3.73), $A(x, t, \mathbf{u}(x, t))\mathbf{u}_x(x, t) + B(x, t, \mathbf{u}(x, t))\mathbf{u}_t(x, t) = \mathbf{c}(x, t, \mathbf{u}(x, t))$, where $\mathbf{u}$ and $\mathbf{c}$ are n-component vectors and A and B are $n \times n$ matrices.

We reconsider the quasilinear system considered in Example 3.5. To deal with this problem, we use the procedure $ClassHypSystChar([[0, -1], [-c, 0]], [[1, 0], [0, 1]], [0, 0], [u, v], [x, t])$. The term $[-c, 0]$ in the argument should be $[-c'(u), 0]$, in order to get the correct system (3.3.69). We use $[-c, 0]$ to enable Maple to determine the eigenvalues and eigenfunctions for this problem. As indicated above, the Maple procedures for doing so fail if the matrix contains general functions in its components. The output of the procedure exhibits the relevant eigenvalues and eigenvectors for this problem. Once the results given by the procedure are exhibited, the substitution $c = -c'(u)$ is used, and the correct eigenvalues, eigenfunctions, and the characteristic normal form given in (3.3.78) are obtained. We do not exhibit the output of the procedure.

Well-Posedness and Stability

We have constructed two procedures, *WellposedScal* and *WellposedSyst*, that determine if a linear scalar PDE or a linear system of first order PDEs with constant coefficients is stable or unstable. This includes the determination if the basic IVP for the PDE is well posed or not. If the PDE is neutrally stable, the dispersion relation for the PDE is exhibited if a real one exists.

For the scalar case the procedure takes the form $WellposedScal(PDE, u, [x, t])$, where PDE is a (homogeneous) partial differential equation to be analyzed, u is the dependent variable, and $[x, t]$ are the independent variables.

If we apply the procedure to (3.5.13), we use $WellposedScal(KGEquation, u, [x, t])$. The output exhibits the *Klein-Gordon equation* and the following information. As in (3.5.3) we obtain $\lambda^2 + \gamma^2 k^2 + c^2 = 0$, which is solved for λ to give $\lambda_1 = \sqrt{-\gamma^2 k^2 - c^2}$, $\lambda_2 = -\sqrt{-\gamma^2 k^2 - c^2}$. The output gives Real part of $\lambda = (0, 0)$, $lub = 0$. The statements *The initial value problem is neutrally stable* and

Dispersion Relation $= -\omega^2 + \gamma^2 k^2 + c^2 = 0$ are printed out. The results agree with those found in Example 3.7.

Next we consider $WellposedScal(PDE, u, [x, t])$ for the third order $PDE = \partial^3 u(x,t)/\partial t^3 + \partial^3 u(x,t)/\partial x^3 = 0$ and obtain $\lambda^3 - ik^3 = 0$, with $i = \sqrt{-1}$. The equation is solved for λ to give $\lambda_1 = -i(-1/2 + i\sqrt{3}/2)k$, $\lambda_2 = -i(-1/2 - i\sqrt{3}/2)k$, $\lambda_3 = -ik$. The results Real part of $\lambda = (\sqrt{3}k/2, -\sqrt{3}k/2, 0)$, $lub = \infty$, and *The initial value problem is not well posed* are printed out.

To deal with a system of PDEs $A\mathbf{u}_x(x,t) + B\mathbf{u}_t(x,t) + C\mathbf{u}(x,t) = \mathbf{0}$, we use $WellposedSyst(A, B, C, u, [x, t])$. We apply it to the system with

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}, \qquad B = I, \qquad C = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{3.7.12}$$

where I is the identity matrix, to obtain $(\lambda + ik)^2(\lambda - 2ik) = 0$, which is solved for λ to give; $\lambda_1 = 2ik$, $\lambda_2 = -ik$, $\lambda_3 = -ik$. The real parts for each of the three roots of λ and the least upper bound are exhibited as; Real part of $\lambda = (0, 0, 0)$, $lub = 0$. Then the statements *The initial value problem is neutrally stable* and Dispersion Relation $= -\omega^3 + 3k^2\omega + 2k^3 = 0$ are printed out.

Exercises 3.7

3.7.1. Consider the PDE $au_{xx} + 2bu_{xy} + cu_{yy} + du_x + eu_y + fu = 0$ with constant coefficients in the principal part. (a) Apply the procedure $ClassPDE2d(a, 2b, c, [x, y])$ to classify PDE, and invoke Maple's $assume$ facility, applied to the discriminant $b^2 - ac$ to obtain the results (3.1.5)–(3.1.7). (b) Use *PDEFactor2d* $(a, 2b, c, d, e, f, [x, y])$ to obtain a factorization of the principal part of the PDE and to obtain canonical forms for the PDE.

3.7.2. Reproduce the results of Example 3.1 by using $ClassPDE2d$ and $PDEFactor2d$.

3.7.3. Use $PDEFactor2d$ to obtain the result of Exercise 3.1.8.

3.7.4. Consider the PDE (3.7.4). Apply the procedure $ClassPDEScalar$ to determine its classification and obtain its canonical form.

3.7.5. Apply $PDEFactor2d$ to solve Exercise 3.1.4.

3.7.6. Apply $PDEFactor2d$ to solve Exercise 3.1.5.

3.7.7. Use $PDEFactor2d$ to solve Exercise 3.1.6.

3.7.8. Apply the procedure $ClassSyst$ to the system (3.7.6) with coefficients given in (3.7.7) and $C = 0$, $\mathbf{d} = \mathbf{0}$.

3.7.9. Apply $ClassSyst$ to the system (3.7.6) with $A = \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix}$, $B = I$, $C = 0$, and $\mathbf{d} = \mathbf{0}$.

3.7.10. Apply *ClassSyst* to the system (3.7.6) with $A = \begin{bmatrix} 0 & -1 \\ -x^2 & 0 \end{bmatrix}$, $B = I$, $C = 0$, and $\mathbf{d} = \mathbf{0}$.

3.7.11. Use *HypSystExplicit* to obtain (3.7.11) for the system (3.7.10).

3.7.12. Apply *ClassHypSystChar* to obtain the results of Example 3.5.

3.7.13. Apply *WellposedScal* to determine the stability property and the dispersion relation for the Klein-Gordon equation (3.5.13).

3.7.14. Apply *WellposedSyst* to $A\mathbf{u}_x(x,t) + B\mathbf{u}_t(x,t) + C\mathbf{u}(x,t) = \mathbf{0}$ with the coefficients (3.7.12).

CHAPTER 4

INITIAL AND BOUNDARY VALUE PROBLEMS IN BOUNDED REGIONS

4.1 INTRODUCTION

This chapter deals with *boundary value problems* for *elliptic equations* and *initial and boundary value problems* for *parabolic* and *hyperbolic equations* given over bounded spatial regions. The basic method for solving these problems is the technique of *separation of variables*. This method requires that eigenvalue problems for differential equations be studied and expresses solutions as series of eigenfunctions. Only *scalar problems* in one, two, or three space dimensions are treated in this chapter. The one-dimensional eigenvalue problem known as the *Sturm-Liouville problem* is studied in detail. Higher-dimensional eigenvalue problems are considered in Chapter 8.

The separation of variables method is applicable to the study of linear homogeneous equations with homogeneous boundary conditions. Techniques for solving *inhomogeneous problems*, such as *Duhamel's principle* and *eigenfunction expansions*, are also presented in this chapter. Further, it is demonstrated by means of a detailed example in Section 4.7 how eigenfunction expansions may be used in the solution of nonlinear problems.

Partial Differential Equations of Appplied Mathematics, Third Edition. By Erich Zauderer

Balance Law for Heat Conduction and Diffusion

To achieve a unity of presentation, we consider only second order equations of each of the three basic types given in a special form. This form, however, is representative of a large class of partial differential equations of physical interest. All equations with constant coefficients derived in Chapter 1 are either of this form or can be reduced to this form. To indicate how these equations, which may have variable coefficients, can arise in applications, we now present a brief derivation of a *parabolic equation* of the form to be considered. This derivation, in contrast to those given in Chapter 1, is more representative of the approaches generally used in deriving partial differential equations as models for physical processes.

Throughout this chapter, G will represent a bounded interval in one space dimension or a bounded region in two or three dimensions. ∂G will represent its boundary. (In one dimension ∂G is just the endpoints of the interval G.)

In the following derivation we restrict our discussion to three space dimensions and let $\hat{G}$ be an arbitrary closed region within G, with $\partial\hat{G}$ as the boundary surface. Let ds be an element of surface area of $\partial\hat{G}$ and dv be a volume element. The function $u(\mathbf{x}, t)$ where $\mathbf{x}$ is a point in space and t is time, is assumed to represent a scalar physical quantity such as the temperature or the concentration of a substance, for example.

With $\mathbf{n}$ as the exterior unit normal vector on the boundary $\partial\hat{G}$, the expression $p(\mathbf{x})\nabla u(\mathbf{x}, t) \cdot \mathbf{n}\, ds$ represents the flux or rate of flow of the quantity $u(\mathbf{x}, t)$ through ds at the time t. The positive function $p(\mathbf{x})$ is given and assumed to be time independent. The operator ∇ is the gradient operator in three-dimensional space. The time rate of change of $u(\mathbf{x}, t)$ in an element dv at the time t is given by $\rho(\mathbf{x})(\partial u(\mathbf{x}, t)/\partial t)\, dv$, where the positive quantity $\rho(\mathbf{x})$ is given and time independent. Additional effects occurring in the element dv at the time t are represented as $H(\mathbf{x}, t)\, dv$, with $H(\mathbf{x}, t)$ assumed to have the form $H(\mathbf{x}, t) = -q(\mathbf{x})u(\mathbf{x}, t) + \hat{F}(\mathbf{x}, t)$, where for notational convenience in our discussion we set $\hat{F}(\mathbf{x}, t) = \rho(\mathbf{x})F(\mathbf{x}, t)$. With $q(\mathbf{x})$ as a nonnegative time-independent function, $q(\mathbf{x})u(\mathbf{x}, t)$ represents (internal) effects due to changes proportional to $u(\mathbf{x}, t)$, whereas $\hat{F}(\mathbf{x}, t)$ represents external influences on the medium under consideration. [See Section 8.2 for a discussion of the consequences of permitting $q(\mathbf{x})$ to be negative.]

The basic physical *balance* or *conservation law* for the arbitrary region $\hat{G}$ is, in terms of the representation of $H(\mathbf{x}, t)$,

$$\iiint_{\hat{G}} \rho \frac{\partial u}{\partial t}\, dv = \int_{\partial\hat{G}} p\nabla u \cdot \mathbf{n}\, ds - \iiint_{\hat{G}} qu\, dv + \iiint_{\hat{G}} \rho F\, dv. \tag{4.1.1}$$

This states that the *time rate of change of* $u(\mathbf{x}, t)$ *in the region* $\hat{G}$ equals *the flux of* $u(\mathbf{x}, t)$ *through the boundary of* $\hat{G}$ and *the changes due to internal effects proportional to* $u(\mathbf{x}, t)$, as well as *external effects acting throughout* $\hat{G}$. That is, these effects must balance each other out.

Applying the *divergence theorem* to the surface integral in (4.1.1) yields

$$\int_{\partial\hat{G}} p\nabla u \cdot \mathbf{n}\, ds = \iiint_{\hat{G}} \nabla \cdot (p\nabla u)\, dv. \tag{4.1.2}$$

Inserting (4.1.2) into (4.1.1) and combining terms gives

$$\iint_{\hat{G}} \left[\rho \frac{\partial u}{\partial t} - \nabla \cdot (p \nabla u) + qu - \rho F \right] dv = 0. \tag{4.1.3}$$

Assuming a continuous integrand, the arbitrariness of the region $\hat{G}$ implies the vanishing of the integrand in (4.1.3) (see Exercise 8.1.9). Consequently, we obtain the *inhomogeneous parabolic equation*

$$\rho(\mathbf{x}) \frac{\partial u(\mathbf{x}, t)}{\partial t} - \nabla \cdot (p(\mathbf{x}) \nabla u(\mathbf{x}, t)) + q(\mathbf{x}) u(\mathbf{x}, t) = \rho(\mathbf{x}) F(\mathbf{x}, t), \tag{4.1.4}$$

which is valid at an arbitrary point in the region G. A similar argument in two space dimensions leads to (4.1.4) with ∇ interpreted as a two-dimensional gradient operator. In one space dimension, (4.1.4) has the form

$$\rho(x) \frac{\partial u(x, t)}{\partial t} - \frac{\partial}{\partial x} \left(p(x) \frac{\partial u(x, t)}{\partial x} \right) + q(x) u(x, t) = \rho(x) F(x, t). \tag{4.1.5}$$

In the context of *heat conduction*, $u(\mathbf{x}, t)$ is the *temperature* and $p(\mathbf{x})$ is the *thermal conductivity* of the medium. The relationship between the heat flux $\mathbf{J}(\mathbf{x}, t)$ and the temperature $u(\mathbf{x}, t)$, when expressed in the form $\mathbf{J}(\mathbf{x}, t) = -p(\mathbf{x}) \nabla u(\mathbf{x}, t)$, is called *Fourier's law*. (The minus sign indicates that heat flows from warmer to cooler regions.) The term $\rho(\mathbf{x})$ is the product of the *specific heat* c and the *density* $\hat{\rho}(\mathbf{x})$ of the medium. In the case of one-dimensional heat conduction in a rod, if there is lateral heat transfer between the rod and the surrounding medium, the function $q(x)$ results from *Newton's law of cooling*, which states that the heat flow through the (lateral) boundary is proportional to the difference between the internal temperature $u(x, t)$ of the rod and the external temperature.

If (4.1.4) or (4.1.5) describes the *diffusion* of a substance, $u(\mathbf{x}, t)$ equals the *concentration* of that substance. The relationship between the flow of the substance and its concentration $u(\mathbf{x}, t)$ has the same form as Fourier's law and is known as *Fick's law* in diffusion theory. The quantity $p(\mathbf{x})$ is called the *diffusion coefficient* or the *diffusivity*. In the simplest models of diffusion, $p(\mathbf{x}) = 1$ and $q(\mathbf{x}) = 0$. (We remark that in the case of heat conduction, if p and ρ are constants, the ratio p/ρ is called the *thermal diffusivity*.)

The foregoing derivations of the heat and diffusion equations were based on macroscopic considerations, where observable phenomena were taken into account. In Chapter 1 we derived the diffusion equation by considering the motion of a particle or a group of particles. These microscopic considerations yielded equations that involve observable quantities such as particle densities, drifts, and variances.

Basic Equations of Parabolic, Elliptic, and Hyperbolic Types

In view of our continual use of the equations above and related equations, we define the differential operator L in two or three dimensions as

$$Lu = -\nabla \cdot (p(\mathbf{x}) \nabla u) + q(\mathbf{x}) u, \tag{4.1.6}$$

and in one space dimension as

$$Lu = -\frac{\partial}{\partial x}\left(p(x)\,\frac{\partial u}{\partial x}\right) + q(x)u. \tag{4.1.7}$$

It is assumed that ρ and p are positive functions and that q is a nonnegative function. Then the *parabolic equations* (4.1.4) and (4.1.5) take the form (using higher-dimensional notation)

$$\rho(\mathbf{x})\,\frac{\partial u(\mathbf{x},t)}{\partial t} + Lu(\mathbf{x},t) = \rho(\mathbf{x})F(\mathbf{x},t). \tag{4.1.8}$$

If the problem leading to (4.1.8) is stationary or time independent, (i.e., F and all auxiliary conditions placed on the problem depend on $\mathbf{x}$ only), we may look for a solution in the form $u = u(\mathbf{x})$, and (4.1.8) reduces to the *elliptic equation*

$$Lu(\mathbf{x}) = \rho(\mathbf{x})F(\mathbf{x}). \tag{4.1.9}$$

Further, it can be shown that a large class of *hyperbolic equations* of interest in applications can be expressed as (using higher-dimensional notation)

$$\rho(\mathbf{x})\,\frac{\partial^2 u(\mathbf{x},t)}{\partial t^2} + Lu(\mathbf{x},t) = \rho(\mathbf{x})F(\mathbf{x},t). \tag{4.1.10}$$

For example, in one space dimension the equation governing the *longitudinal vibration of a rod* has the form (4.1.10). We assume the rod has a constant cross-sectional area A and let $u(x,t)$ be the displacement of the section of the rod at the point x as a function of time. By *Hooke's law*, the tension force acting on the section at the point x is given as $p(x)Au_x(x,t)$, where $p(x)$ is *Young's modulus*. A segment of the rod extending from x to $x+\Delta x$, of length Δx, has the momentum $\rho A\,\Delta x\,u_t$, where $\rho(x)$ is the density. Then *Newton's law of motion* applied to this segment of the rod yields $\rho(x)\,\Delta x\,u_{tt}(x,t) = Ap(x+\Delta x)\,u_x(x+\Delta x,t) - Ap(x)\,u_x(x,t)$. On dividing by $A\,\Delta x$ and letting Δx tend to zero, we obtain the one-dimensional form of (4.1.10) with $q = F = 0$.

As has already been remarked in Example 2.4 and will be demonstrated in Example 4.2, the equation of the *vibrating string* has the form (4.1.10) (in one dimension) with $p = T$, the constant tension of the string. Similarly, the *transverse vibration* of a tightly stretched *thin membrane* with density $\rho(x,y)$ is governed by the two-dimensional form of (4.1.10), where p equals T, the constant tension of the membrane. In both cases u is the vertical displacement of a point on the string or on the membrane. If a restoring force proportional to the displacement acts at each point of the string or membrane, in addition to other external forces, the full equation (4.1.10) applies.

The *telegrapher's equation* of Chapter 1 is not of the form (4.1.10) since it contains a $\partial u/\partial t$ term. However, a simple change of the dependent variable reduces it to the form above. In fact, any equation of the form (4.1.10) that contains an additional term $\hat{\rho}\,\partial u/\partial t$ can be transformed into the form (4.1.10) by a change of variable if $\hat{\rho}/\rho =$ constant.

Boundary Conditions

In the following sections we consider initial and/or boundary value problems for the equations (4.1.8)–(4.1.10). Appropriate initial values for the parabolic and hyperbolic equations (4.1.8) and (4.1.10) will be introduced as required. However, the boundary values on ∂G for all three equations (4.1.8)–(4.1.10) are assumed to be of a standard form common in applications.

In the case of two or three space dimensions, we consider boundary conditions of the form

$$\alpha(\mathbf{x})u + \beta(\mathbf{x})\frac{\partial u}{\partial n}\bigg|_{\partial G} = B(\mathbf{x}, t), \tag{4.1.11}$$

where $\alpha(\mathbf{x})$, $\beta(\mathbf{x})$, and $B(\mathbf{x}, t)$ are given functions evaluated on the boundary region ∂G. In the *elliptic case*, G is a region in $\mathbf{x}$-space and $B = B(\mathbf{x})$ is independent of t. In the *parabolic* and *hyperbolic cases*, we consider a region $G_{\mathbf{x}}$ in $\mathbf{x}$-space and construct the region G in $(\mathbf{x}, t)$-space whose points $(\mathbf{x}, t)$ satisfy the condition that $\mathbf{x} \in G$ and that t ranges through a specified set of values, say, $t_0 \leq t \leq t_1$. (The region G can also be characterized as a *direct product* in the form $G = G_{\mathbf{x}} \times [t_0, t_1]$.) The expression $\partial u/\partial n$ denotes the exterior normal derivative on ∂G.

The boundary condition (4.1.11) relates the values of u on ∂G and the flux of u through ∂G. We require that $\alpha(\mathbf{x}) \geq 0$, $\beta(\mathbf{x}) \geq 0$, and $\alpha(\mathbf{x}) + \beta(\mathbf{x}) > 0$ on ∂G. (These conditions may be violated in general.) If (1) $\alpha(\mathbf{x}) \neq 0$, $\beta(\mathbf{x}) = 0$, (2) $\alpha(\mathbf{x}) = 0$, $\beta(\mathbf{x}) \neq 0$, and (3) $\alpha(\mathbf{x}) \neq 0$, $\beta(\mathbf{x}) \neq 0$, (4.1.11) are boundary conditions of the *first, second,* and *third kind*, respectively. Boundary conditions of the first and second kind are generally referred to as *Dirichlet* and *Neumann conditions*, and boundary conditions of the third kind are often called *Robin conditions*. We can also consider boundary conditions of the *mixed kind*. This means that a mixture of boundary conditions of the first, second, and third kinds is specified for u on the boundary.

In the interests of greater precision, we assume that ∂G is subdivided into three disjoint subsets S_1, S_2, and S_3 (i.e., $\partial G = S_1 \bigcup S_2 \bigcup S_3$). On S_1, S_2, and S_3, u satisfies a boundary condition of the *first, second* and *third kind*, respectively. Each of these sets may be an empty set. If at least two of the sets are not empty, the boundary condition is of the *mixed kind*. If only one of the three is not empty, the boundary condition is of the first, second, or third kind according as the set is S_1, S_2, or S_3. To simplify the presentation of formulas and results, this notation is used throughout when dealing with integrals over the boundary and with boundary conditions.

In one dimension, G represents the interval $0 < x < l$ and ∂G the points $x = 0$ and $x = l$. In the *hyperbolic* and *parabolic cases*, (4.1.11) takes the form

$$\alpha_1 u(0, t) - \beta_1 u_x(0, t) = g_1(t), \quad \alpha_2 u(l, t) + \beta_2 u_x(l, t) = g_2(t). \tag{4.1.12}$$

The conditions on (α, β) given above carry over in a natural way to the pairs of constants (α_1, β_1) and (α_2, β_2). The minus sign in the condition at $x = 0$ in (4.1.12) is needed since $\partial u/\partial n = -\partial u/\partial x$ at that point.

When F and B (or g_1 and g_2) in (4.1.8)–(4.1.12) are nonzero, the equations and boundary conditions are of nonhomogeneous type. The separation of variables

method introduced in the following section applies only to the homogeneous cases, when F and B are both zero. However, when this method is applied to boundary value problems for the elliptic equation (4.1.9), we must have $F = 0$ while $B = 0$ on only part of the boundary, as will be seen.

Exercises 4.1

4.1.1. Derive the heat or diffusion equation (4.1.5) from a *balance law* appropriate to the one-dimensional case.

4.1.2. Show that if $\hat{\rho}(\mathbf{x}) = c\rho(\mathbf{x})$, where c is a constant, $\rho(\mathbf{x})u_{tt}(\mathbf{x},t) + \hat{\rho}(\mathbf{x})u_t(\mathbf{x},t) + Lu(\mathbf{x},t) = \rho(\mathbf{x})F(\mathbf{x},t)$ can be brought into the general form (4.1.10) if we set $u(\mathbf{x},t) = \exp[\alpha t]v(\mathbf{x},t)$ and choose α appropriately.

4.1.3. Carry out the transformation of Exercise 4.1.2 for the telegrapher's equation $u_{tt}(x,t) - \gamma^2 u_{xx}(x,t) + 2\hat{\lambda}u_t(x,t) = 0$, where $\hat{\lambda}$ is a positive constant.

4.2 SEPARATION OF VARIABLES

In this section we present the *method of separation of variables* and apply it to the solution of initial and/or boundary value problems for homogeneous versions of the equations and boundary conditions introduced in Section 4.1.

In the *hyperbolic case* we consider the homogeneous PDE

$$\rho(\mathbf{x})\frac{\partial^2 u(\mathbf{x},t)}{\partial t^2} + Lu(\mathbf{x},t) = 0, \qquad \mathbf{x} \in G, \quad t > 0, \tag{4.2.1}$$

where the bounded region G, the coefficient $\rho(\mathbf{x})$, and the operator L are defined as in Section 4.1. The homogeneous boundary conditions are

$$\alpha(\mathbf{x})u(\mathbf{x},t) + \beta(\mathbf{x})\frac{\partial u(\mathbf{x},t)}{\partial n}\bigg|_{\partial G} = 0, \qquad t > 0, \tag{4.2.2}$$

$$\alpha_1 u(0,t) - \beta_1 u_x(0,t) = 0, \quad \alpha_2 u(l,t) + \beta_2 u_x(l,t) = 0, \; t > 0, \tag{4.2.3}$$

in two or three dimensions and in one dimension, respectively. On the boundary ∂G, the coefficients in (4.2.2)–(4.2.3) must satisfy the conditions given in Section 4.1. The initial conditions for the PDE (4.2.1) are given as

$$u(\mathbf{x},0) = f(\mathbf{x}), \qquad u_t(\mathbf{x},0) = g(\mathbf{x}), \qquad \mathbf{x} \in G. \tag{4.2.4}$$

In the *parabolic case*, the equation for $u(\mathbf{x},t)$ is given as

$$\rho(\mathbf{x})\frac{\partial u(\mathbf{x},t)}{\partial t} + Lu(\mathbf{x},t) = 0, \qquad \mathbf{x} \in G, \; t > 0, \tag{4.2.5}$$

with the boundary conditions (4.2.2) or (4.2.3), and the initial condition

$$u(\mathbf{x}, 0) = f(\mathbf{x}), \qquad \mathbf{x} \in G. \tag{4.2.6}$$

In the *elliptic case*, we introduce a function $u = u(\mathbf{x}, y)$, where $\mathbf{x}$ is a point in the region G (which is now one- or two-dimensional, with $\mathbf{x}$ replaced by x in the one-dimensional case) and y is a scalar variable given over the interval $0 < y < \hat{l}$. The PDE for $u(\mathbf{x}, y)$ has the form

$$\rho(\mathbf{x}) \frac{\partial^2 u(\mathbf{x}, y)}{\partial y^2} - Lu(\mathbf{x}, y) = 0, \qquad \mathbf{x} \in G,\ 0 < y < \hat{l}, \tag{4.2.7}$$

with ρ, p, and q as functions of $\mathbf{x}$ only. The boundary conditions in $\mathbf{x}$ or x are of the form (4.2.2) or (4.2.3). These are combined with the following inhomogeneous boundary conditions at $y = 0$ and $y = \hat{l}$:

$$u(\mathbf{x}, 0) = f(\mathbf{x}), \qquad u(\mathbf{x}, \hat{l}) = g(\mathbf{x}), \qquad \mathbf{x} \in G. \tag{4.2.8}$$

Example 4.1. Equations with Constant Coefficients. We consider the one-dimensional case, assume that $\rho(x) = p(x) = 1$ and $q(x) = 0$ in (4.2.1), (4.2.5), and (4.2.7), and formulate the appropriate problems for each of the PDEs above. Figures 4.1 to 4.3 display the regions where the solutions are to be determined and the initial and/or boundary conditions for these problems.

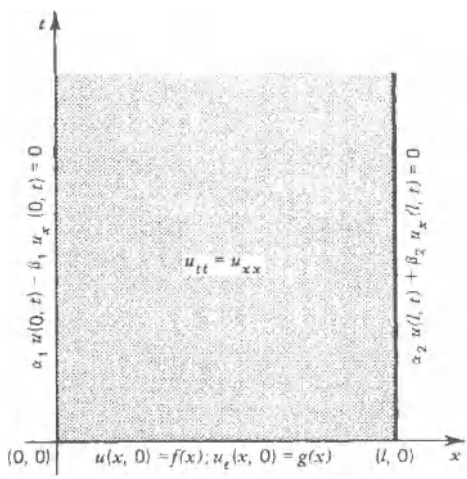

Figure 4.1 Hyperbolic case.

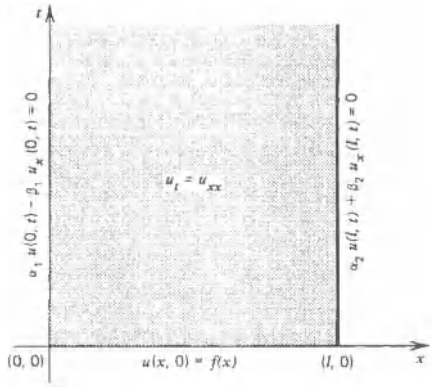

Figure 4.2 Parabolic case.

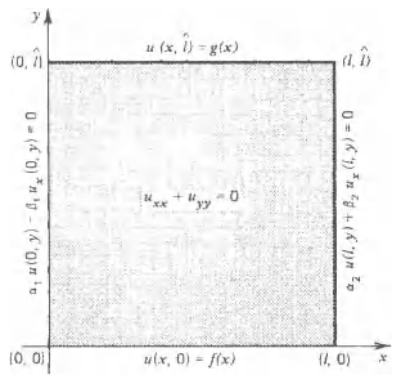

Figure 4.3 Elliptic case.

It should be noted that for the hyperbolic and parabolic cases we are considering initial and boundary value problems, whereas in the elliptic case we have a strict boundary value problem. The equation (4.2.7) for the elliptic problem is written in the form indicated to permit a unified presentation of the separation of variables technique for all three cases.

The method of *separation of variables* asks for a solution of (4.2.1) and (4.2.5) in the form

$$u(\mathbf{x}, t) = M(\mathbf{x})N(t), \tag{4.2.9}$$

and of (4.2.7) in the form

$$u(\mathbf{x}, y) = M(\mathbf{x})N(y), \tag{4.2.10}$$

with the function $M(\mathbf{x})$ required to satisfy the boundary conditions (4.2.2) or (4.2.3), with $\mathbf{x}$ replaced by x as required. Substituting (4.2.9) and (4.2.10) into the appropriate

equations and dividing through by ρMN yields

$$\frac{N''(t)}{N(t)} = -\frac{LM(\mathbf{x})}{\rho(\mathbf{x})M(\mathbf{x})}, \qquad \text{hyperbolic case,} \tag{4.2.11}$$

$$\frac{N'(t)}{N(t)} = -\frac{LM(\mathbf{x})}{\rho(\mathbf{x})M(\mathbf{x})}, \qquad \text{parabolic case,} \tag{4.2.12}$$

$$-\frac{N''(y)}{N(y)} = -\frac{LM(\mathbf{x})}{\rho(\mathbf{x})M(\mathbf{x})}, \qquad \text{elliptic case.} \tag{4.2.13}$$

Since the left and right sides of these equations depend on different variables, they cannot be nonconstant functions of their respective variables. Thus each side of the equations must be a constant. We denote this (separation) constant by $-\lambda$. As a result, we obtain the following equations for M and N:

$$LM(\mathbf{x}) = \lambda\rho(\mathbf{x})M(\mathbf{x}), \tag{4.2.14}$$

$$\begin{cases} N''(t) + \lambda N(t) = 0, & \text{hyperbolic case,} \\ N'(t) + \lambda N(t) = 0, & \text{parabolic case,} \\ N''(y) - \lambda N(y) = 0, & \text{elliptic case.} \end{cases} \tag{4.2.15}$$

In addition to being a solution of (4.2.14), $M(\mathbf{x})$ is required to satisfy the boundary conditions (4.2.2) or (4.2.3). [Then $u(\mathbf{x}, t) = M(\mathbf{x})N(t)$ or $u(\mathbf{x}, y) = M(\mathbf{x})N(y)$ will also satisfy these boundary conditions.] Since both the equation and the boundary conditions for $M(\mathbf{x})$ are homogeneous, $M(\mathbf{x}) = 0$ is a solution of the problem. It must be rejected since it is of no value in solving the given problem for u. Thus, we must find nonzero solutions of the boundary value problem for $M(\mathbf{x})$, referred to as *eigenfunctions*, and such solutions exist only for certain values of the parameter λ in (4.2.14), the corresponding *eigenvalues*. This known as an *eigenvalue problem*.

Self-Adjoint and Positive Operators

When applied to functions that satisfy the boundary conditions (4.2.2) or (4.2.3), the differential operator L [or more precisely, $(1/\rho)L$] in (4.2.14) has the property of being a *self-adjoint* as well as a *positive operator*. We discuss these properties below, following which important consequences of these properties for the *eigenvalue problem* associated with the operator L are examined. Although a large part of our discussion is presented in notation appropriate for the two- and three-dimensional cases, the general results are valid for the one-dimensional case as well.

Let u and w be (smooth) functions defined over the two- or three-dimensional region G that satisfy the boundary condition (4.2.2) on ∂G. Then

$$\iint_G [wLu - uLw]\, dv = \int_{\partial G} p\left[u\frac{\partial w}{\partial n} - w\frac{\partial u}{\partial n}\right] ds = 0, \tag{4.2.16}$$

where $\partial/\partial n$ is the exterior normal derivative, as will be shown. The operator L was determined to be *formally self-adjoint* in Example 3.8, but we verify this here again and go on to show that the boundary value problem is *self-adjoint*.

To prove (4.2.16) we note that

$$w\nabla \cdot (p\nabla u) = \nabla \cdot (pw\nabla u) - p\nabla w \cdot \nabla u. \tag{4.2.17}$$

Interchanging w and u in (4.2.17) and subtracting one expression from the other gives

$$w\nabla \cdot (p\nabla u) - u\nabla \cdot (p\nabla w) = \nabla \cdot (pw\nabla u) - \nabla \cdot (pu\nabla w). \tag{4.2.18}$$

In view of (4.1.6) (i.e., the expression for L), we obtain

$$\iint_G [wLu - uLw]dv = -\int_{\partial G} p[w\nabla u - u\nabla w] \cdot \mathbf{n}\, ds = \int_{\partial G} p\left[u\frac{\partial w}{\partial n} - w\frac{\partial u}{\partial n}\right]ds, \tag{4.2.19}$$

on applying the divergence theorem. Both u and w satisfy (4.2.2), so that

$$\alpha(\mathbf{x})u + \beta(\mathbf{x})\frac{\partial u}{\partial n}\bigg|_{\partial G} = 0, \quad \alpha(\mathbf{x})w + \beta(\mathbf{x})\frac{\partial w}{\partial n}\bigg|_{\partial G} = 0. \tag{4.2.20}$$

We may think of (4.2.20) as a simultaneous system of homogeneous equations for α and β at each point of ∂G. Since $\alpha + \beta > 0$ on ∂G by assumption, the system must have a nonzero solution. This can occur only if the determinant of the coefficients of the system (i.e., $u\,\partial w/\partial n - w\,\partial u/\partial n$) vanishes on ∂G. Consequently, the surface integral in (4.2.19) vanishes, as was to be shown.

Given two functions $f(\mathbf{x})$ and $g(\mathbf{x})$ defined and integrable over the region G, the *inner product* of these two functions with *weight* $\rho(\mathbf{x}) > 0$ is defined as

$$(f(\mathbf{x}), g(\mathbf{x})) = \iint_G \rho(\mathbf{x}) f(\mathbf{x}) g(\mathbf{x})\, dv. \tag{4.2.21}$$

These functions are assumed to be real-valued and need not satisfy the boundary conditions (4.2.2). Since $(f(\mathbf{x}), g(\mathbf{x})) = (g(\mathbf{x}), f(\mathbf{x}))$, the inner product is *symmetric*. (If the functions are complex valued, a different inner product, introduced in the next section, is defined.) If $(f(\mathbf{x}), g(\mathbf{x})) = 0$, the functions $f(\mathbf{x})$ and $g(\mathbf{x})$ are said to be *orthogonal*. In terms of the inner product (4.2.21), we define a *norm* of the function $f(\mathbf{x})$ as

$$||f(\mathbf{x})|| = \sqrt{(f(\mathbf{x}), f(\mathbf{x}))} = \sqrt{\iint_G \rho(\mathbf{x}) f^2(\mathbf{x})\, dv}. \tag{4.2.22}$$

The norm $||f(\mathbf{x})||$ is nonnegative for real valued $f(\mathbf{x})$, and is a measure of the magnitude of $f(\mathbf{x})$. If $f(\mathbf{x})$ is continuous in G, then $||f(\mathbf{x})|| = 0$ if and only if $f(\mathbf{x}) = 0$. Most often one defines the inner product (4.2.21) with $\rho(\mathbf{x}) = 1$, but for our discussion we require the more general definition given above.

In terms of the inner product (4.2.21), (4.2.16) can be expressed as

$$\left(w, \frac{1}{\rho} Lu\right) = \left(\frac{1}{\rho} Lw, u\right). \tag{4.2.23}$$

An operator L with the property that $(w, Lu) = (Lw, u)$ is said to be *self-adjoint*. If $\rho(\mathbf{x}) \neq 1$, the operator $(1/\rho(\mathbf{x}))L$ satisfies the self-adjointness condition in terms of the inner product (4.2.21), while if $\rho(\mathbf{x}) = 1$, the operator L is self-adjoint. Thus, the self-adjointness of $(1/\rho(\mathbf{x}))L$ or L is determined not only by the definition of L but also by the boundary conditions (4.2.2) that u and w must satisfy. (The discussion of adjoint differential operators given in Section 3.6 involved only the concept of formal adjointness.)

Next, we suppose that u satisfies the boundary condition (4.2.2). We have

$$\iint_G uLu\, dv = -\iint_G [\nabla \cdot (pu\nabla u) - p(\nabla u)^2]\, dv + \iint_G qu^2\, dv$$

$$= \iint_G [p(\nabla u)^2 + qu^2]\, dv - \int_{\partial G} pu \frac{\partial u}{\partial n}\, ds, \tag{4.2.24}$$

where (4.2.17) and the divergence theorem have been used. On the subset S_3 of ∂G, (4.2.2) yields $\partial u/\partial n = -(\alpha/\beta)\, u$, while on S_1 and S_2, u and $\partial u/\partial n$ vanish, respectively. Consequently, (4.2.24) can be expressed as

$$\iint_G uLu\, dv = \iint_G [p(\nabla u)^2 + qu^2]\, dv + \int_{S_3} p \frac{\alpha}{\beta} u^2\, ds. \tag{4.2.25}$$

The assumptions on p, q, α, and β given previously imply that the right side of (4.2.25) is nonnegative if u is real valued.

In terms of the inner product (4.2.21), (4.2.25) implies that if u is real valued,

$$\iint_G uLu\, dv = \left(\frac{1}{\rho} Lu, u\right) \geq 0. \tag{4.2.26}$$

A *self-adjoint operator* $\hat{L}$ with the property that $(\hat{L}u, u) \geq 0$ is said to be a *positive operator*. Thus if $\rho \neq 1$, we see that $(1/\rho)L$ is a positive operator, whereas if $\rho = 1$, the operator L is itself positive. Again the positivity of $(1/\rho)L$ or L is based not only on its definition but also on the boundary conditions (4.2.2) that u must satisfy. If we relax the requirement that q in (4.1.4) and α and β in (4.1.11) must be nonnegative, the operator $(1/\rho)L$ is still self-adjoint as long as α and β are not both zero simultaneously. However, for $(1/\rho)L$ to be a positive operator, these conditions cannot be relaxed.

Eigenvalues, Eigenfunctions, and Eigenfunction Expansions

The problem of determining a nonzero solution $M(\mathbf{x})$ of (4.2.14), with boundary conditions of the form (4.2.2) and (4.2.3), is known as an *eigenvalue problem*. The values of λ for which nonzero $M(\mathbf{x})$ exist are known as *eigenvalues*. The corresponding solutions of the eigenvalue problem for $M(\mathbf{x})$ are called *eigenfunctions*. In the one-dimensional case, the eigenvalue problem is known as the *Sturm–Liouville problem*, and it is discussed in the following section.

We now apply the results of the foregoing discussion to the eigenvalue problem (4.2.14). Let $M_k(\mathbf{x})$ and $M_j(\mathbf{x})$ be eigenfunctions corresponding to the distinct eigenvalues λ_k and λ_j, respectively. Since $M_k(\mathbf{x})$ and $M_j(\mathbf{x})$ both satisfy the boundary condition (4.2.2), the result (4.2.16) implies that

$$\iint_G (M_k L M_j - M_j L M_k)\, dv = 0. \tag{4.2.27}$$

But $LM_j = \lambda_j \rho M_j$ and $LM_k = \lambda_k \rho M_k$, so that (4.2.27) yields

$$\iint_G (\lambda_j \rho M_k M_j - \lambda_k \rho M_j M_k)\, dv = (\lambda_j - \lambda_k)(M_k, M_j). \tag{4.2.28}$$

By assumption, $\lambda_j \neq \lambda_k$, so that (4.2.28) implies that

$$(M_k, M_j) = 0, \qquad j \neq k. \tag{4.2.29}$$

Consequently, eigenfunctions corresponding to different eigenvalues must be *orthogonal* in terms of the inner product (4.2.21).

Using (4.2.29) and the fact that the coefficients in the differential operator L are assumed to be real valued, it is easily shown that the eigenvalues for our problem must be real and that the corresponding eigenfunctions may be chosen to be real valued. (This is proven for the one-dimensional eigenvalue problem in Section 4.3, and the method of proof carries over immediately to the higher-dimensional problem.)

Let $M_k(\mathbf{x})$ be a real-valued eigenfunction that corresponds to the real eigenvalue λ_k. Since $LM_k = \lambda_k \rho M_k$, the positivity property (4.2.26) implies that

$$\iint_G M_k L M_k\, dv = \lambda_k \iint_G \rho M_k^2\, dv = \lambda_k ||M_k||^2 \geq 0. \tag{4.2.30}$$

Because $M_k(\mathbf{x})$ is real valued and nonzero by assumption, we conclude that the eigenvalue $\lambda_k \geq 0$.

Thus the self-adjointness property of the operator $(1/\rho)L$ implies that the eigenvalues are real and that eigenfunctions corresponding to different eigenvalues are orthogonal. The positivity property implies that the eigenvalues are nonnegative. It can be shown (see Section 8.1) that there are a countably infinite number of eigenvalues λ_k $(k = 1, 2, \ldots)$ whose only limit point is at infinity (i.e., $\lambda_k \to \infty$ as $k \to \infty$). We denote the corresponding set of (real valued) eigenfunctions by $M_k(\mathbf{x})$. Further properties of eigenvalues and eigenfunctions are presented later, and the eigenvalue problem is reexamined from a variational point of view in Section 8.1.

For each eigenvalue λ_k we obtain an equation for N_k in (4.2.15). Assuming for simplicity that $\lambda_k > 0$ for all k, we have for the N_k

$$\begin{cases} N_k(t) = a_k \cos(\sqrt{\lambda_k}\, t) + b_k \sin(\sqrt{\lambda_k}\, t), & \text{hyperbolic case,} \\ N_k(t) = a_k \exp(-\lambda_k t), & \text{parabolic case,} \\ N_k(y) = a_k \exp(\sqrt{\lambda_k}\, y) + b_k \exp(-\sqrt{\lambda_k}\, y), & \text{elliptic case.} \end{cases} \tag{4.2.31}$$

The a_k and b_k are arbitrary constants that must be determined. Forming the product $M_k N_k$, we obtain $u_k = M_k N_k$, which for each value of k satisfies the appropriate equation (4.2.1), (4.2.5), or (4.2.7), together with the boundary condition (4.2.2) or (4.2.3).

To satisfy the additional conditions on $u(\mathbf{x}, t)$ or $u(\mathbf{x}, y)$ given above, we consider the formal *superposition* of the solutions u_k and construct the series

$$u = \sum_{k=1}^{\infty} u_k = \sum_{k=1}^{\infty} M_k N_k. \tag{4.2.32}$$

The a_k and b_k in the terms N_k are chosen to satisfy the additional conditions placed on u that are as yet unaccounted for by our choice of the $M_k(\mathbf{x})$.

In the *hyperbolic case* the initial conditions (4.2.4) must be satisfied and we obtain formally

$$\begin{cases} u(\mathbf{x}, 0) = \sum_{k=1}^{\infty} M_k(\mathbf{x}) N_k(0) = \sum_{k=1}^{\infty} a_k M_k(\mathbf{x}) = f(\mathbf{x}), \\ u_t(\mathbf{x}, 0) = \sum_{k=1}^{\infty} M_k(\mathbf{x}) N_k'(0) = \sum_{k=1}^{\infty} \sqrt{\lambda_k} b_k M_k(\mathbf{x}) = g(\mathbf{x}). \end{cases} \tag{4.2.33}$$

The expressions (4.2.33) represent *eigenfunction expansions* of $f(\mathbf{x})$ and $g(\mathbf{x})$. The validity of this representation is characterized by the concept of completeness of the set of eigenfunctions $M_k(\mathbf{x})$. If the set is *complete*, then under certain conditions on the functions $f(\mathbf{x})$ and $g(\mathbf{x})$, the expansions (4.2.33) with appropriately chosen a_k and b_k will converge to these functions. The more stringent the conditions, the better the convergence. Since the $M_k(\mathbf{x})$ are eigenfunctions for a self-adjoint operator as we have shown, even under very mild conditions on $f(\mathbf{x})$ and $g(\mathbf{x})$ the series of eigenfunctions will in some sense converge to these functions. This is discussed more fully for the one-dimensional case in Section 4.3 and for the higher-dimensional problem in Chapter 8.

For the *hyperbolic case* considered above and the further discussion in this section, the $M_k(\mathbf{x})$ are assumed to be an orthogonal set; that is, $(M_k, M_j) = 0$ for $k \neq j$. As shown above, this assumption is valid if each $M_k(\mathbf{x})$ corresponds to a different eigenvalue, that is, there is exactly one linearly independent eigenfunction for each eigenvalue. In the case of a multiple eigenvalue for which there are a finite number of linearly independent eigenfunctions, the *Gram-Schmidt orthogonalization process* (see Exercise 8.2.1) can be used to orthogonalize the finite set of eigenfunctions. (The eigenvalue problem we consider here cannot have more than a finite number of linearly independent eigenfunctions for each eigenvalue.)

To determine the a_k in (4.2.33), we multiply the series by $\rho(\mathbf{x}) M_j(\mathbf{x})$ and integrate over G. This gives

$$(f(\mathbf{x}), M_j(\mathbf{x})) = \sum_{k=1}^{\infty} a_k (M_k(\mathbf{x}), M_j(\mathbf{x})) = a_j (M_j(\mathbf{x}), M_j(\mathbf{x})), \tag{4.2.34}$$

on using (4.2.29) and assuming that summation and integration can be interchanged in the series. Proceeding similarly for the b_k, we find that

$$a_k = \frac{(f(\mathbf{x}), M_k(\mathbf{x}))}{(M_k(\mathbf{x}), M_k(\mathbf{x}))}, \qquad b_k = \frac{(g(\mathbf{x}), M_k(\mathbf{x}))}{\sqrt{\lambda_k}(M_k(\mathbf{x}), M_k(\mathbf{x}))}. \tag{4.2.35}$$

The formal solution of the initial and boundary value problem for the hyperbolic equation (4.2.1) is then given as

$$u(\mathbf{x},t)=\sum_{k=1}^{\infty}\left(a_k\cos(\sqrt{\lambda_k}\,t)+b_k\sin(\sqrt{\lambda_k}\,t)\right)M_k(\mathbf{x}). \tag{4.2.36}$$

For the *parabolic case*, the initial condition (4.2.6) must be met, so that

$$u(\mathbf{x},0)=\sum_{k=1}^{\infty}M_k(\mathbf{x})N_k(0)=\sum_{k=1}^{\infty}a_kM_k(\mathbf{x})=f(\mathbf{x}). \tag{4.2.37}$$

Thus the a_k are specified as in (4.2.35) and the formal solution of the initial and boundary value problem for the parabolic equation (4.2.5) is

$$u(\mathbf{x},t)=\sum_{k=1}^{\infty}a_ke^{-\lambda_kt}M_k(\mathbf{x}). \tag{4.2.38}$$

In the *elliptic case*, the boundary conditions (4.2.8) must be satisfied. This yields the eigenfunction expansions

$$\begin{cases} u(\mathbf{x},0)=\sum_{k=1}^{\infty}(a_k+b_k)M_k(\mathbf{x})=f(\mathbf{x}), \\ u(\mathbf{x},\hat{l})=\sum_{k=1}^{\infty}\left(a_k\exp(\sqrt{\lambda_k}\,\hat{l})+b_k\exp(-\sqrt{\lambda_k}\,\hat{l})\right)M_k(\mathbf{x})=g(\mathbf{x}). \end{cases} \tag{4.2.39}$$

Applying the technique used in the hyperbolic case yields

$$\begin{cases} a_k+b_k=(f(\mathbf{x}),M_k(\mathbf{x}))/(M_k(\mathbf{x}),M_k(\mathbf{x})), \\ a_k\exp(\sqrt{\lambda_k}\,\hat{l})+b_k\exp(-\sqrt{\lambda_k}\,\hat{l})=(g(\mathbf{x}),M_k(\mathbf{x}))/(M_k(\mathbf{x}),M_k(\mathbf{x})). \end{cases} \tag{4.2.40}$$

Unique solutions for the a_k and b_k can be determined from the foregoing system. Then the formal solution of the boundary value problem for the elliptic equation (4.2.7) is

$$u(\mathbf{x},y)=\sum_{k=1}^{\infty}\left(a_k\exp(\sqrt{\lambda_k}\,y)+b_k\exp(-\sqrt{\lambda_k}\,y)\right)M_k(\mathbf{x}). \tag{4.2.41}$$

The solutions above are often termed *classical solutions* of the given equations if the formal operations carried out are valid and the series expansions of u can be differentiated term by term as often as required by the equations and the initial and boundary data. However, even if term-by-term differentiability is not valid, the sum u may still be characterized as a *generalized solution* if certain conditions specified later are met.

We have shown that the separation of variables method reduces each of the initial and boundary value problems given in this section to the study of an eigenvalue problem. Unless the eigenvalues λ_k and the eigenfunctions $M_k(\mathbf{x})$ can be determined, the formal series solutions obtained remain unspecified. For many of the problems

considered in this book the eigenvalues and eigenfunctions can be determined exactly, as we demonstrate in this and later chapters. When exact results are not available, approximation methods such as the Rayleigh-Ritz method discussed in Section 8.2 and the finite element method presented in Section 12.7 yield effective approximations to the leading and most significant eigenvalues and eigenfunctions.

Exercises 4.2

4.2.1. Verify the result (4.2.16) in the two-dimensional case if $L = -\nabla^2$, $u = 1 - r^2$, $w = (1 - r)^2$, where $r^2 = x^2 + y^2$, if the region G is the interior of the circle $r = 1$. Note that both u and w vanish on the boundary $r = 1$.

4.2.2. Let $u = (1 - r)^2$ and $w = \cos(\pi r)$, so that $\partial u/\partial n$ and $\partial w/\partial n$ both vanish on the circle $r = 1$. With $L = -\nabla^2$ and G given as the region $r < 1$, verify that the integral over the region G in (4.2.16) vanishes.

4.2.3. With $R^2 = x^2 + y^2 + z^2$, let $u = \exp(1 - R)$ and $w = (R - 1)^2$, so that $\partial u/\partial n + u$ and $\partial w/\partial n + w$ both vanish on the sphere $R = 1$. If G is the interior of the unit sphere (i.e., the region $R < 1$) and $L = -\nabla^2 + 10$, show that (4.2.16) is valid for the functions above.

4.2.4. If the region G is the unit square $0 < x < 1$, $0 < y < 1$, and $u(x, y) = xy(1 - x)(1 - y)$ so that $u(x, y)$ vanishes on ∂G, show that (4.2.26) is satisfied if $L = -\nabla^2$.

4.2.5. Show that $\lambda = 0$ is an eigenvalue and $M(\mathbf{x}) = c =$ constant is an eigenfunction for the eigenvalue problem (4.2.14) if and only if $q = 0$ in the operator L and $\alpha = 0$ in the boundary condition. *Hint*: Use (4.2.25).

4.2.6. Show that if $\lambda = 0$ occurs as an eigenvalue when the elliptic problem (4.2.7) is solved by separation of variables, with the boundary condition (4.2.8) replaced by $u_y(x, 0) = f(x)$ and $u_y(x, \hat{l}) = g(x)$, the problem has no solution unless $f(x)$ and $g(x)$ satisfy compatibility conditions. (Use the results of Exercise 4.2.5 to show this.) Conclude that the solution is not unique if the compatibility conditions are met and verify that these results are consistent with those obtained in Exercise 3.4.9.

4.2.7. Show that $M_j(x, y) = 1$ and $M_k(x, y) = \cos(\pi x)\cos(\pi y)$ are both eigenfunctions for the problem $-(M_{xx}(x, y) + M_{yy}(x, y)) = \lambda M(x, y)$, $0 < x < 1$, $0 < y < 1$, with the boundary condition $\partial M(x, y)/\partial n = 0$ on the unit square. Determine the corresponding eigenvalues and verify directly that the orthogonality condition $(M_j(x, y), M_k(x, y)) = 0$ is satisfied.

4.2.8. Show that $M(\mathbf{x}) = 1$ is an eigenfunction for $-\nabla \cdot (p(\mathbf{x})\nabla M(\mathbf{x})) - M(\mathbf{x}) = \lambda M(\mathbf{x})$ in the region G with $\partial M(\mathbf{x})/\partial n = 0$ on ∂G. Determine the corresponding eigenvalue and explain why the occurrence of a negative eigenvalue does not contradict (4.2.30).

4.2.9. Assuming that the eigenfunctions $M_k(\mathbf{x})$ are uniformly bounded, interpret the fact that $u(\mathbf{x}, t)$ as given in (4.2.38) vanishes as $t \to \infty$ if the λ_k are all positive, in

terms of the physical significance of the heat conduction problem and its boundary conditions. Consider also the special case referred to in Exercise 4.2.5, which yields $\lambda = 0$ as an eigenvalue.

4.2.10. Consider the two-dimensional eigenvalue problem in polar coordinates r and θ, $LM(r,\theta) = -\nabla^2 M(r,\theta) = \lambda M(r,\theta)$ in the region $r < 1$ with the boundary condition (a) $M(l,0) = 0$, (b) $M_r(l,\theta) = 0$, or (c) $M_r(l,\theta) + hM(l,\theta) = 0$, $(h > 0)$. (i) By expressing the Laplacian in polar coordinates and assuming that $M = M(r)$ (i.e., it is independent of θ), show that $M(r)$ satisfies Bessel's equation of order zero. (ii) In the general case, with $M = M(r,\theta)$, set $M(r,\theta) = M_1(r)M_2(\theta)$ and use separation of variables to show that if $M(r,\theta + 2\pi) = M(r,\theta)$, the function $M_1(r)$ must satisfy a Bessel equation of integral order. For each of these eigenvalue problems we require that the solutions be bounded at $r = 0$ (see Section 4.3 for further discussion).

4.2.11. Express Laplace's equation $\nabla^2 u = 0$ in spherical coordinates (r,θ,ϕ) with $r \geq 0$, $0 \leq \theta \leq 2\pi$, and $0 \leq \phi \leq \pi$ (see Example 8.3). Assume that $u = u(r,\phi)$ (i.e., u is independent of θ) and set $u(r,\phi) = N(r)M(\phi)$. Use separation of variables to show that $M(\phi)$ satisfies $-d\left[\sin(\phi)dM(\phi)/d\phi\right]/d\phi = \lambda\sin(\phi)M(\phi)$, with λ as the separation constant. Let $x = \cos(\phi)$ and conclude that $\hat{M}(x) = M(\phi)$ satisfies the Legendre equation. If we require that $\hat{M}(x)$ be bounded at $x = -1$ and $x = +1$, we obtain the eigenvalue problem discussed in Section 4.3.

4.2.12. Given any two functions f and g for which the inner product and norm are defined as in (4.2.21)–(4.2.22), show that they satisfy the *Cauchy-Schwarz inequality* $|(f,g)| \leq ||f||\,||g||$. *Hint*: Expand out the non-negative integral $\iint_G \rho(f + \alpha g)^2\, dv$, where α is a constant, as a quadratic expression in α and show that the discriminant must be negative or zero.

4.2.13. Suppose that $q(\mathbf{x})$ in the operator L [defined in (4.1.6)] can have negative values but that $q(\mathbf{x}) \geq \hat{q}$, with $\hat{q}$ as a negative constant. Also, let $\rho(\mathbf{x}) \geq \hat{\rho}$, with $\hat{\rho}$ as a positive constant. Put $a = \hat{q}/\hat{\rho}$, and define the operator $\hat{L}u(\mathbf{x}) = Lu(\mathbf{x}) - a\rho(\mathbf{x})u(\mathbf{x})$. For functions $u(\mathbf{x})$ that satisfy (4.2.2), show that $(1/\rho(\mathbf{x}))\hat{L}$ is a positive operator. Since the eigenvalue problem for $\hat{L}$ must yield only nonnegative eigenvalues, conclude from the fact that the eigenvalues have no finite limit points, that the eigenvalue problem for L, with $q(\mathbf{x})$ given as above, can have only a finite number of negative eigenvalues.

4.2.14. Consider the operator $\tilde{L}$ defined as $\tilde{L}u(\mathbf{x}) = Lu(\mathbf{x}) + \mathbf{b}(\mathbf{x}) \cdot \nabla u(\mathbf{x})$, where $Lu(\mathbf{x})$ has the form (4.1.6) and $\mathbf{b}(\mathbf{x})$ is a given vector function. Determine the formal adjoint $\tilde{L}^*$ of the operator $\tilde{L}$. Also, determine the boundary conditions that $w(\mathbf{x})$ must satisfy so that the integral of $w(\mathbf{x})\tilde{L}u(\mathbf{x}) - u(\mathbf{x})\tilde{L}^* w(\mathbf{x})$ over the region G vanishes if $u(\mathbf{x})$ satisfies a homogeneous boundary condition of the (a) first kind, (b) second kind, or (c) third kind on ∂G.

4.2.15. Consider the eigenvalue problem $\tilde{L}M(\mathbf{x}) = \lambda\rho(\mathbf{x})M(\mathbf{x})$ for each of the three boundary conditions determined for $\tilde{L}$ in Exercise 4.2.14. Proceed as in the text and use the conclusion of Exercise 4.2.14 to show that eigenfunctions for the eigenvalue

problem for $\tilde{L}$ and the adjoint eigenvalue problem for $\tilde{L}^*$ are orthogonal if they correspond to different eigenvalues. These two sets of eigenfunctions are referred to as a *biorthogonal set.*

4.3 THE STURM-LIOUVILLE PROBLEM AND FOURIER SERIES

Sturm-Liouville Problem

The one-dimensional version of the eigenvalue problem of Section 4.2,

$$L[v(x)] = -\frac{d}{dx}\left[p(x)\frac{dv(x)}{dx}\right] + q(x)v(x) = \lambda\rho(x)v(x), \tag{4.3.1}$$

where $0 < x < l$ and $v(x)$ satisfies the boundary conditions (4.2.3),

$$\alpha_1 v(0) - \beta_1 v'(0) = 0, \qquad \alpha_2 v(l) + \beta_2 v'(l) = 0, \tag{4.3.2}$$

is known as the *Sturm-Liouville problem.* We require that $p(x) > 0$, $\rho(x) > 0$, and $q(x) \geq 0$, and that $p(x)$, $\rho(x)$, $q(x)$, and $p'(x)$ be continuous in the closed interval $0 \leq x \leq l$. Also, we must have $\alpha_i \geq 0$, $\beta_i \geq 0$, and $\alpha_i + \beta_i > 0$ for $i = 1, 2$. Then we have a *regular Sturm-Liouville problem.* However, if one or more of the conditions on the coefficients in (4.3.1) are relaxed, say, if $p(x)$ or $\rho(x)$ or both are permitted to vanish at either or both of the endpoints $x = 0$ and $x = l$, we have a *singular Sturm-Liouville problem.* Both cases are of interest in applications and examples of each type are considered. We note that in discussions of Sturm-Liouville problems in the literature it is generally not required that $q(x)$ and the coefficients in (4.3.2) be nonnegative.

Before discussing and deriving some of the important properties of the eigenvalues and eigenfunctions of the Sturm-Liouville problem, we introduce several definitions and concepts. Some of these were given in Section 4.2 but are repeated here in their one-dimensional form. Unless otherwise specified, the functions considered in the following discussion are real valued.

The *inner product* of two functions $\phi(x)$ and $\psi(x)$ (bounded and integrable over the interval $0 \leq x \leq l$) is defined as

$$(\phi(x),\ \psi(x)) = \int_0^l \rho(x)\phi(x)\psi(x)\,dx, \tag{4.3.3}$$

with the weight function $\rho(x) > 0$ in $0 < x < l$. [Most often one uses a unit weight function $\rho(x) = 1$. For our purposes, the definition (4.3.3) is more appropriate.] The inner product is clearly *symmetric*, that is, $(\phi(x),\ \psi(x)) = (\psi(x),\ \phi(x))$. The *norm* of a function $\phi(x)$ defined in terms of the inner product (4.3.3) (or induced by it) is

$$||\phi(x)|| = \sqrt{(\phi(x), \phi(x))} = \sqrt{\int_0^l \rho(x)\phi^2(x)\,dx}. \tag{4.3.4}$$

The norm, which is clearly nonnegative, is a measure of the magnitude of the function $\phi(x)$ over the interval $0 \leq x \leq l$. If $\phi(x)$ is continuous in that interval, then $||\phi(x)|| = 0$ if and only if $\phi(x) = 0$. Any function $\phi(x)$ with a finite norm [i.e., $||\phi(x)|| < \infty$] is said to be *square integrable* over the interval $0 < x < l$. If $\phi(x)$ is such that $||\phi(x)|| = 1$, then $\phi(x)$ is said to be *normalized* to unity. Any square integrable function can be normalized by defining a new function $\hat{\phi}(x) = \phi(x)/||\phi(x)||$ for which we see that $\hat{\phi}(x) = 1$. [We require that all square integrable functions be integrable (see Exercises 4.3.1 and 4.3.2).]

Two functions $\phi(x)$ and $\psi(x)$ for which

$$(\phi(x),\ \psi(x)) = 0 \tag{4.3.5}$$

are said to be *orthogonal* over the interval $0 < x < l$. A set of functions $\{\phi_k(x)\}$, $k = 1, 2, \ldots$, for which

$$(\phi_k(x),\ \phi_j(x)) = 0, \qquad k \neq j, \tag{4.3.6}$$

is said to be an *orthogonal set*. If, in addition, $||\phi_k(x)|| = 1$ for all $\phi_k(x)$, the set is said to be *orthonormal*. Any orthogonal set can be orthonormalized by normalizing each of the $\phi_k(x)$ in the manner shown above.

If the functions $\phi(x)$, $\psi(x)$, or $\phi_k(x)$ can assume complex values, we replace the inner product (4.3.3) by the (weighted) *Hermitian inner product*

$$(\phi(x),\ \psi(x)) = \int_0^l \rho(x)\phi(x)\overline{\psi(x)}\, dx, \tag{4.3.7}$$

where the overbar denotes complex conjugation. The inner product has *Hermitian symmetry* since $(\phi(x),\ \psi(x)) = \overline{(\psi(x),\ \phi(x))}$. The induced norm is

$$||\phi(x)|| = \sqrt{(\phi(x),\ \phi(x))} = \sqrt{\int_0^l \rho(x)|\phi(x)|^2\, dx}, \tag{4.3.8}$$

so that $||\phi(x)|| \geq 0$.

We now return to a consideration of real-valued functions. Given the orthonormal set of square integrable functions $\{\phi_k(x)\}$, $k = 1, 2, \ldots$, and the square integrable function $\phi(x)$ over the interval $0 < x < l$, the set of numbers $(\phi,\ \phi_k)$ are called the *Fourier coefficients* of $\phi(x)$. The formal series

$$\phi(x) = \sum_{k=1}^{\infty} (\phi(x), \phi_k)\phi_k(x) \tag{4.3.9}$$

is called the *Fourier series* of $\phi(x)$. Even though we have equated $\phi(x)$ to its Fourier series in (4.3.9), we have yet to specify in what sense the Fourier series converges to the function $\phi(x)$, if it converges at all. Further, although we use the term *Fourier series* to denote any expansion of a function in terms of a set of orthonormal (or more generally, orthogonal) functions, in our later discussions, as well as in the literature,

this term is more commonly used to denote an expansion in a series of *trigonometric functions.*

To discuss the convergence properties of the series (4.3.9), we note that for the finite sum $\sum_{k=1}^{N} (\phi(x), \phi_k)\phi_k(x)$, we have

$$\left|\left|\phi(x) - \sum_{k=1}^{N} (\phi, \phi_k)\phi_k(x)\right|\right|^2 = (\phi, \phi) - 2\sum_{k=1}^{N} (\phi, \phi_k)^2 + \sum_{k=1}^{N} (\phi, \phi_k)^2$$

$$= ||\phi||^2 - \sum_{k=1}^{N} (\phi, \phi_k)^2 \geq 0, \tag{4.3.10}$$

where we have used elementary properties of the inner product, the orthonormality of the set $\{\phi_k(x)\}$, and the nonnegativity of the norm. Since $||\phi(x)|| < \infty$ and (4.3.10) is valid for all N, we obtain in the limit as $N \to \infty$

$$\sum_{k=1}^{\infty} (\phi(x), \phi_k)^2 \leq ||\phi(x)||^2, \tag{4.3.11}$$

which is known as *Bessel's inequality.* This shows that the sum of squares of the Fourier coefficients of any square integrable function $\phi(x)$ converges, and implies that $(\phi(x), \phi_k)$ tends to zero as $k \to \infty$.

A sequence of (square integrable) functions $\{\Phi_N(x)\}$, $N = 1, 2, ...$, is said to converge to a function $\phi(x)$ in the mean if

$$\lim_{N\to\infty} ||\phi(x) - \Phi_N(x)|| = 0. \tag{4.3.12}$$

This differs from and does not imply pointwise convergence, and is called *mean square convergence.* Denoting the partial sums of (4.3.9) by

$$\Phi_N(x) = \sum_{k=1}^{N} (\phi(x), \phi_k)\phi_k(x), \tag{4.3.13}$$

we see that if *Parseval's equality,*

$$\sum_{k=1}^{\infty} (\phi(x), \phi_k)^2 = ||\phi(x)||^2, \tag{4.3.14}$$

is satisfied, (4.3.10) implies that

$$\lim_{N\to\infty} ||\phi(x) - \Phi_N(x)||^2 = \lim_{N\to\infty} \left\{ ||\phi(x)||^2 - \sum_{k=1}^{N} (\phi(x), \phi_k)^2 \right\} = 0. \tag{4.3.15}$$

Thus the Fourier series (4.3.9) converges in the mean square sense to the function $\phi(x)$ in the interval $0 < x < l$ if Parseval's equality is satisfied.

A set of square integrable functions $\{\phi_k(x)\}$, $k = 1, 2, \ldots$, is said to be *complete* (with respect to mean square convergence) if for any square integrable function $\phi(x)$, its Fourier series (4.3.9) converges to it in the mean. It follows from (4.3.15) that if Parseval's equation (4.3.14) is satisfied for all square integrable functions $\phi(x)$, the set of functions $\{\phi_k(x)\}$, $k = 1, 2, \ldots$, is complete. Although few restrictions need to be placed on a function $\phi(x)$ in order to achieve mean square convergence for its Fourier series, in applications of Fourier series to the solution of boundary value problems, stronger forms of convergence such as uniform convergence are generally required.

Properties of Eigenvalues and Eigenfunctions

We return to the Sturm-Liouville problem, list the basic properties of the eigenvalues and eigenfunctions, and derive the simplest of these properties.

1. Eigenfunctions corresponding to different eigenvalues are orthogonal. Let λ_i and λ_j be two distinct eigenvalues and $v_i(x)$ and $v_j(x)$ two corresponding eigenfunctions for the problem (4.3.1)–(4.3.2). Then

$$\int_0^l [v_i L v_j - v_j L v_i]\, dx = \int_0^l \frac{d}{dx}\left(p v_j v_i' - p v_i v_j'\right) dx = 0, \tag{4.3.16}$$

as is easily verified on using the boundary conditions (4.3.2) for v_i and v_j. But we also have

$$\int_0^l [v_i L v_j - v_j L v_i]\, dx = (\lambda_i - \lambda_j) \int_0^l \rho v_i v_j\, dx = (\lambda_i - \lambda_j)(v_i, v_j). \tag{4.3.17}$$

Since $\lambda_i \neq \lambda_j$, by assumption, we conclude that

$$(v_i(x), v_j(x)) = 0, \tag{4.3.18}$$

which implies orthogonality. [This result should be compared with (4.2.27)–(4.2.29) and follows from the self-adjointness of the operator $(1/\rho)L$.]

2. The eigenvalues are real and nonnegative and the eigenfunctions may be chosen to be real valued. Suppose that an eigenvalue λ_i is complex valued. Then $\lambda_j = \overline{\lambda_i}$ represents a second, distinct eigenvalue. Because the coefficients in the operator L are real valued, eigenfunctions corresponding to λ_i and $\lambda_j = \overline{\lambda_i}$ are $v_i(x)$ and $v_j(x) = \overline{v_i(x)}$, where $v_j(x)$ is obtained by complex conjugation in the equation for $v_i(x)$. Then since $\lambda_i \neq \lambda_j$, (4.3.18) shows that

$$(v_i(x), v_j(x)) = (v_i(x), \overline{v_i(x)}) = \int_0^l \rho(x) |v_i(x)|^2\, dx = 0. \tag{4.3.19}$$

But $v_i(x) \not\equiv 0$, so that the integral in (4.3.19) must be nonzero, and we have a contradiction. Consequently, we must have $\lambda_j = \lambda_i$ so that λ_i is real valued. Since the coefficients and the eigenvalues in the equation (4.3.1) are now real valued, the

real and imaginary parts of any solution (i.e., eigenfunction) must also satisfy (4.3.1). Therefore, the eigenfunctions can be chosen to be real valued and we assume them to be such.

The nonnegativity of the eigenvalues can be shown as follows. Let $v(x)$ be an eigenfunction. Then

$$\left(v, \frac{1}{\rho}Lv\right) = -pvv'|_0^l + \int_0^l pv'^2\,dx + \int_0^l qv^2\,dx \geq 0, \tag{4.3.20}$$

since, say, at $x = l$ we have

$$-p(l)v(l)v'(l) = \begin{cases} (\alpha_2/\beta_2)p(l)v^2(l) \geq 0, & \beta_2 > 0, \\ 0, & \beta_2 = 0, \end{cases} \tag{4.3.21}$$

with a similar result valid at $x = 0$. [The last two integrals in (4.3.20) are clearly nonnegative.] But we also have

$$\left(v, \frac{1}{\rho}Lv\right) = (v, \lambda v) = \lambda(v, v) = \lambda||v||^2, \tag{4.3.22}$$

with $||v|| > 0$, by assumption. Combining (4.3.20) and (4.3.22) gives

$$\lambda = \frac{(v, (Lv)/\rho)}{||v||^2} \geq 0, \tag{4.3.23}$$

which proves that the eigenvalue λ is nonnegative. [Note that the positivity of the operator $(1/\rho)L$ has been used to prove this result. This may be compared with the discussion in the preceding section.]

The foregoing discussion may be used to show that if $q(x) > 0$ in the interval $0 \leq x \leq l$, $\lambda = 0$ cannot be an eigenvalue of the Sturm-Liouville problem. If $q(x) \equiv 0, \lambda = 0$ is an eigenvalue if and only if $\alpha_1 = \alpha_2 = 0$. This is shown in the exercises.

3. Each eigenvalue is simple. The second order ODE (4.3.1) can have at most two linearly independent eigenfunctions for each eigenvalue λ. But, the Sturm-Liouville problem, (see the exercises), has only one linearly independent eigenfunction for each eigenvalue. That is, each eigenvalue is simple.

4. There is a countable infinity of eigenvalues having a limit point at infinity. The set of eigenvalues can be arranged as follows: $0 \leq \lambda_1 < \lambda_2 < \lambda_3 < \ldots$ with $\lambda_k \to \infty$ as $k \to \infty$. The set of eigenvalues is called the *spectrum* of the operator L, and we see that the spectrum is discrete, nonnegative, and has a limit point at infinity.

5. The set of eigenfunctions $\{v_k(x)\}$, $k = 1, 2, \ldots$ *forms a complete orthonormal set of square integrable functions on the interval* $0 < x < l$. As a result, if the function $v(x)$ is square integrable over the interval $0 < x < l$, the Fourier series or, equivalently, the eigenfunction expansion

$$v(x) = \sum_{k=1}^{\infty} (v(x), v_k)v_k(x) \tag{4.3.24}$$

converges to $v(x)$ in the mean. It can be shown that if $v(x)$ is continuous and, in addition, has a piecewise continuous first derivative in $0 \leq x \leq l$ and $v(x)$ satisfies

the boundary conditions (4.3.2), the Fourier series (4.3.24) converges absolutely and uniformly to $v(x)$ in the given interval. If $v(x)$ has a jump discontinuity at some interior point x_0 in the interval, the Fourier series converges to $(v(x_0-)+v(x_0+))/2$ at that point, where $v(x_0-)$ and $v(x_0+)$ are one-sided limits of $v(x)$ as x approaches x_0 from the left and right, respectively.

Proofs of properties 4 and 5 are not given here. However, the discussion in Section 8.1, which deals with higher-dimensional eigenvalue problems can be adapted to yield a proof of these properties for the Sturm-Liouville problem.

Determination of Eigenvalues and Eigenfunctions

The eigenvalues and eigenfunctions of the Sturm-Liouville problem can be obtained formally as follows. Let $V(x;\lambda)$ and $W(x;\lambda)$ be solutions of the initial value problems for the equation (4.3.1) with the initial conditions

$$V(0;\lambda)=1, \quad V'(0;\lambda)=0, \qquad W(0;\lambda)=0, \quad W'(0;\lambda)=1, \tag{4.3.25}$$

respectively. Then, the function

$$v(x;\lambda)=\beta_1 V(x;\lambda)+\alpha_1 W(x;\lambda) \tag{4.3.26}$$

is a solution of (4.3.1) that satisfies the boundary condition (4.3.2) at $x=0$. To satisfy the condition at $x=l$ we must have

$$\alpha_2\beta_1 V(l;\lambda)+\alpha_2\alpha_1 W(l;\lambda)+\beta_2\beta_1 V'(l;\lambda)+\beta_2\alpha_1 W'(l;\lambda)=0, \tag{4.3.27}$$

which results from the substitution of (4.3.26) into (4.3.2). The eigenvalues $\lambda=\lambda_k$ $(k=1,2,\ldots)$ of the Sturm-Liouville problem are determined as the roots of the equation (4.3.27). The corresponding eigenfunctions are given as

$$v_k(x)\equiv v(x;\lambda_k)=\beta_1 V(x;\lambda_k)+\alpha_1 W(x;\lambda_k), \qquad k=1,2,\ldots. \tag{4.3.28}$$

In the following, we consider a number of regular and singular Sturm-Liouville problems that lead to trigonometric, Bessel, and Legendre eigenfunctions.

Trigonometric Eigenfunctions

We begin with a fairly general case and then specialize the results to obtain the basic trigonometric eigenfunctions. We set $p(x)=\rho(x)=1$ and $q(x)=0$ in (4.3.1) while retaining the general boundary conditions (4.3.2). The resulting equation for the eigenfunction $v(x)$ is

$$-v''(x)=\lambda v(x). \tag{4.3.29}$$

The solutions of (4.3.29) that satisfy (4.3.25) are found to be

$$V(x;\lambda)=\cos(\sqrt{\lambda}\,x), \qquad W(x;\lambda)=\frac{\sin(\sqrt{\lambda}\,x)}{\sqrt{\lambda}}, \tag{4.3.30}$$

assuming that $\lambda > 0$. Exercise 4.3.10 deals with $\lambda = 0$. Then

$$v(x;\lambda) = \beta_1 \cos(\sqrt{\lambda}\, x) + \alpha_1 \frac{\sin(\sqrt{\lambda}\, x)}{\sqrt{\lambda}}, \tag{4.3.31}$$

and the eigenvalues are determined from the transcendental equation

$$\sqrt{\lambda}(\alpha_1\beta_2 + \beta_1\alpha_2)\cos(\sqrt{\lambda}\, l) + (\alpha_2\alpha_1 - \lambda\beta_2\beta_1)\sin(\sqrt{\lambda}\, l) = 0, \tag{4.3.32}$$

some of whose properties are considered in Exercise 4.3.9. Once the eigenvalues λ_k $(k = 1, 2, \ldots)$ are determined from (4.3.32), we obtain the eigenfunctions

$$v_k(x) = \beta_1 \cos(\sqrt{\lambda_k}\, x) + \alpha_1 \frac{\sin(\sqrt{\lambda_k}\, x)}{\sqrt{\lambda_k}}, \qquad k = 1, 2, \ldots. \tag{4.3.33}$$

Fourier Sine Series

We reconsider the eigenvalue equation (4.3.29) but simplify the boundary conditions (4.3.2) and assume that

$$\alpha_1 = \alpha_2 = 1, \qquad \beta_1 = \beta_2 = 0. \tag{4.3.34}$$

Then $\lambda = 0$ is not an eigenvalue (see Exercise 4.3.10) and the eigenvalue equation (4.3.32) reduces to

$$\sin(\sqrt{\lambda}\, l) = 0, \tag{4.3.35}$$

and this yields the eigenvalues

$$\lambda_k = \left(\frac{\pi k}{l}\right)^2, \qquad k = 1, 2, \ldots. \tag{4.3.36}$$

The normalized eigenfunctions are given as

$$v_k(x) = \sqrt{\frac{2}{l}} \sin\left(\frac{\pi k}{l} x\right), \qquad k = 1, 2, \ldots. \tag{4.3.37}$$

as is easily verified.

Based on our exact determination of the eigenvalues and eigenfunctions for the foregoing Sturm-Liouville problem, the first four properties for the general eigenvalue problem listed above can be verified directly. The expansion of a function in a series of eigenfunctions (4.3.37) as given in (4.3.24) is known as a *Fourier sine series.*

Fourier Cosine Series

We reconsider (4.3.29) but with boundary conditions (4.3.2) for which we set

$$\alpha_1 = \alpha_2 = 0, \qquad \beta_1 = \beta_2 = 1. \tag{4.3.38}$$

The eigenvalue equation for $\lambda > 0$ becomes

$$\sin(\sqrt{\lambda}\, l) = 0, \tag{4.3.39}$$

and since $\lambda = 0$ is an eigenvalue (see Exercise 4.3.10) we obtain the eigenvalues

$$\lambda_k = \left(\frac{\pi k}{l}\right)^2, \qquad k = 0, 1, 2, \ldots. \tag{4.3.40}$$

The corresponding normalized eigenfunctions are found to be

$$v_0(x) = \frac{1}{\sqrt{l}}, \qquad v_k(x) = \sqrt{\frac{2}{l}} \cos\left(\frac{\pi k}{l} x\right), \; k = 1, 2, \ldots. \tag{4.3.41}$$

Again, the first four properties for the general Sturm-Liouville problem can be verified. The series expansion of a function in terms of the eigenfunctions (4.3.41) in the form of (4.3.24) is known as a *Fourier cosine series*.

Fourier Series

Finally, we reconsider the equation (4.3.29) but extend the interval from $0 \leq x \leq l$ to $-l \leq x \leq l$. In addition, (4.3.2) are replaced by what are known as *periodic boundary conditions*. Specifically, we study the eigenvalue problem

$$-v''(x) = \lambda v(x), \qquad -l < x < l, \tag{4.3.42}$$

with the (periodic) boundary conditions

$$v(-l) = v(l), \qquad v'(-l) = v'(l). \tag{4.3.43}$$

The eigenvalue problem (4.3.42)–(4.3.43) is not of the Sturm-Liouville type, as it was defined above. Nevertheless, most of the properties listed for the Sturm-Liouville problem remain valid in this case as we shall see.

Using the general solution

$$v(x; \lambda) = a \cos(\sqrt{\lambda}\, x) + b \sin(\sqrt{\lambda}\, x) \tag{4.3.44}$$

of (4.3.42) (where we have tacitly assumed that $\lambda > 0$), we easily conclude that the eigenvalues are given as

$$\lambda_k = \left(\frac{\pi k}{l}\right)^2, \qquad k = 0, 1, 2, \ldots, \tag{4.3.45}$$

on applying (4.3.43). By considering the solution of (4.3.42)–(4.3.43) for $\lambda < 0$ it can be shown that $v(x) = 0$ is the only possible result, so that there are no negative eigenvalues. However, $\lambda_0 = 0$ is a eigenvalue for this problem.

For each λ_k except $\lambda_0 = 0$, $\cos(\sqrt{\lambda_k}\, x)$ and $\sin(\sqrt{\lambda_k}\, x)$ are linearly independent eigenfunctions, so that each $\lambda_k > 0$ is a *double eigenvalue*. Thus property 3 given for the Sturm-Liouville problem is not valid in this case.

The full set of normalized eigenfunctions in terms of the (new) inner product (defined over $-l < x < l$)

$$(\phi(x), \psi(x)) = \int_{-l}^{l} \phi(x)\psi(x)\, dx, \tag{4.3.46}$$

and the induced norm

$$||\phi(x)|| = \sqrt{\int_{-l}^{l} \phi^2(x)\, dx}, \tag{4.3.47}$$

is given as

$$\hat{v}_0(x) = \frac{1}{\sqrt{2l}}, \qquad \hat{v}_k(x) = \frac{1}{\sqrt{l}} \cos\left(\frac{\pi k}{l} x\right), \qquad k = 1, 2, \ldots, \tag{4.3.48}$$

$$v_k(x) = \frac{1}{\sqrt{l}} \sin\left(\frac{\pi k}{l} x\right), \qquad k = 1, 2, \ldots, \tag{4.3.49}$$

For $k \geq 1$, $\hat{v}_k(x)$ and $v_k(x)$ are linearly independent normalized eigenfunctions that correspond to the eigenvalues λ_k. It can be verified directly that

$$(\hat{v}_k(x), \hat{v}_j(x)) = (v_k(x), v_j(x)) = 0, \qquad k \neq j, \tag{4.3.50}$$

$$(\hat{v}_k(x), v_j(x)) = 0, \qquad k,\ j = 0, 1, 2, \ldots. \tag{4.3.51}$$

Thus the set of eigenfunctions $\{\hat{v}_k(x), v_j(x)\}$, $k = 0, 1, 2, \ldots$, $j = 1, 2, \ldots$ is an orthonormal set with respect to the inner product (4.3.46). Further, it can be shown that the eigenfunctions form a complete set with respect to square integrable functions $v(x)$ over the interval $-l < x < l$. Such functions have a *Fourier series* expansion

$$v(x) = (v(x), \hat{v}_0)\hat{v}_0(x) + \sum_{k=1}^{\infty} [(v(x), \hat{v}_k)\hat{v}_k(x) + (v(x), v_k)v_k(x)] \tag{4.3.52}$$

that converges to $v(x)$ in the mean.

If $v(x)$ is continuous and, in addition, has a piecewise continuous first derivative in the interval $-l \leq x \leq l$ and $v(x)$ satisfies the boundary conditions (4.3.43), the series (4.3.52) converges absolutely and uniformly to $v(x)$. It may be noted that all essential properties of the Sturm-Liouville eigenvalues and eigenfunctions carry over to the present case.

Properties of Trigonometric Fourier Series

Trigonometric Fourier series are of sufficient interest to be studied without relating them to an eigenvalue problem. Their basic common feature is that they all involve

expansions in the functions $\{\cos[(\pi k/l)x]\}$, $k = 0, 1, 2, \ldots$ and $\{\sin[(\pi k/l)x]\}$, $k = 1, 2, \ldots$ These functions are all periodic of period $2l$, the cosine functions are even functions of x, and the sine functions are odd functions of x. [A function is periodic of period P if $\phi(x + P) = \phi(x)$ for all x. It is an even function if $\phi(-x) = \phi(x)$ for all x, and it is an odd function if $\phi(-x) = -\phi(x)$ for all x.] Thus, although the functions $v(x)$ that are expanded in trigonometric series are defined over the interval $0 \leq x \leq l$ or $-l \leq x \leq l$, the Fourier series themselves are defined for all x.

As a result, assuming that the trigonometric (Fourier) series converge, they may be considered to provide an extension of the definition of the function $v(x)$ from its given interval to the entire real line. Thus in the case of *Fourier series*, the given function $v(x)$ is extended as the function $V(x)$ as follows:

$$\begin{cases} V(x) = v(x), & -l \leq x \leq l, \\ V(x + 2l) = V(x), & -\infty < x < \infty. \end{cases} \tag{4.3.53}$$

The *Fourier sine series* of $v(x)$, defined over the interval $0 \leq x \leq l$, is extended as the function $V_o(x)$ which is given as

$$\begin{cases} V_o(x) = v(x), & 0 \leq x \leq l, \\ V_o(x) = -v(-x), & -l < x < 0, \\ V_o(x + 2l) = V_o(x), & -\infty < x < \infty. \end{cases} \tag{4.3.54}$$

Thus $V_o(x)$ first extends $v(x)$ as an odd function into the interval $-l < x < 0$ and then as an (odd) periodic function of period $2l$ over the entire x-axis. The odd extension is relevant here since all the sine functions are odd.

For the *Fourier cosine series* we extend $v(x)$, which is defined over the interval $0 \leq x \leq l$, to the entire x-axis via the function $V_e(x)$ as follows:

$$\begin{cases} V_e(x) = v(x), & 0 \leq x \leq l, \\ V_e(x) = v(-x), & -l < x < 0, \\ V_e(x + 2l) = V_e(x), & -\infty < x < \infty. \end{cases} \tag{4.3.55}$$

Thus $V_e(x)$ is an even periodic extension of period $2l$ of the function $v(x)$.

We have shown that the Fourier series, sine series, and cosine series not only represent $v(x)$ over the appropriate interval (i.e., either $0 \leq x \leq l$ or $-l \leq x \leq l$) but also the functions $V(x)$, $V_o(x)$, and $V_e(x)$, respectively, over the entire real line. Consequently, if $V(x)$ is an odd periodic function, the Fourier series must take the form of a sine series. Similarly, if $V(x)$ is periodic and even, the Fourier series must be a cosine series. Also, the graphs of the extended functions $V(x)$, $V_o(x)$, and $V_e(x)$ provide some insight into the pointwise convergence properties of the appropriate Fourier series.

As an example, we consider the function $v(x) = x$ over the interval $0 \leq x \leq l$. The odd and even periodic extensions of $v(x)$ are shown in Figures 4.4 and 4.5. If $v(x) = x$ in $-l \leq x \leq l$, the periodic extension of $v(x)$ to $V(x)$ with $V(x + 2l) = V(x)$, is identical to the function $V_o(x)$ graphed in Figure 4.4. In either case the extended function $V_o(x) = V(x)$ has jump discontinuities at the points

$x = \pm(2n+1)l$ $(n = 0, 1, 2, \ldots)$. These discontinuities cause the pointwise convergence of the Fourier series to be slow away from these points. At these points the Fourier series converges to zero. In the case of the sine series, $v(x) = x$ does not satisfy the appropriate boundary condition (4.3.2) at $x = l$, while for the Fourier series the function $v(x) = x$ is not periodic over the interval $-l \leq x \leq l$ since $v(-l) \neq v(l)$. For the even periodic extension, $V_e(x)$ is continuous for all x but is not differentiable at the points $x = nl$ $(n = 0, \pm 1, \pm 2, \ldots)$. Pointwise convergence is more rapid in this case than for the odd or periodic extension considered above. Yet it is still fairly slow because the curve is not smooth. In fact, with $v(x) = x$ we have $v'(0) = v'(l) = 1$, so that the boundary condition (4.3.2) appropriate for the cosine eigenfunctions is not satisfied.

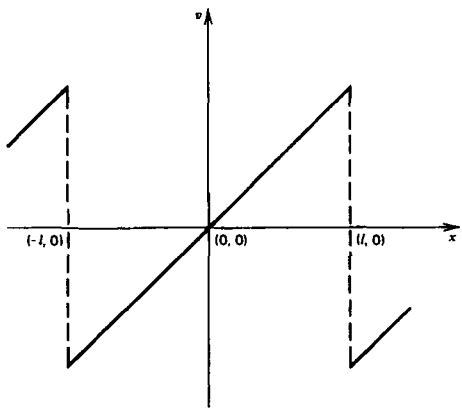

Figure 4.4 Odd extension of $v(x) = x$.

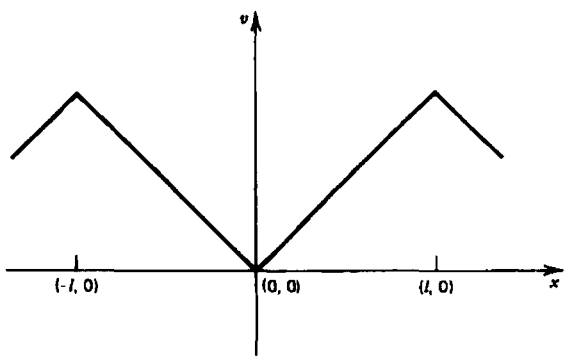

Figure 4.5 Even extension of $v(x) = x$.

If the boundary conditions for the relevant eigenvalue problem are satisfied by the given function $v(x)$, the smoother $v(x)$ is, the better the convergence. This is

evidenced by the smoothness of the relevant extended functions $V_o(x)$, $V_e(x)$, and $V(x)$. In general, the greater the number of continuous derivatives the extended functions have, the more rapid the pointwise convergence of the Fourier series. This matter is examined further in the exercises.

Bessel Eigenfunctions and Their Series

We consider a *singular Sturm-Liouville problem* that arises from the eigenvalue problem for *Bessel's equation* of order n, with n as an integer or zero,

$$L[v(x)] = -\frac{d}{dx}\left[x\frac{dv(x)}{dx}\right] + \frac{n^2}{x}v(x) = \lambda x v(x), \tag{4.3.56}$$

on the interval $0 < x < l$. [The ODE (4.3.56) has the self-adjoint form of (4.3.1). On replacing x by $\sqrt{\lambda}x$, (4.3.56) becomes Bessel's equation of order n.] Since $p(x) = \rho(x) = x$ and $q(x) = n^2/x$, we see that these functions are positive for $0 < x < l$, but $p(x)$ and $\rho(x)$ vanish at $x = 0$ and $q(x)$ is singular there. Nevertheless, $x = 0$ is a *regular singular point* for Bessel's equation, and bounded solutions of (4.3.56) do exist.

To obtain these solutions we impose the boundary conditions

$$v(x) \quad \text{bounded at} \quad x = 0, \qquad v(l) = 0. \tag{4.3.57}$$

(The condition at $x = l$ can be replaced by a boundary condition of the second or third kind. The eigenvalues and eigenfunctions for these problems are considered in the exercises.) In general, the boundedness condition at the singular point is supplemented by the condition $\lim_{x\to 0} p(x)v'(x) = 0$. This guarantees that not only are the eigenfunctions bounded, but that eigenfunctions corresponding to different eigenvalues are orthogonal. To see this, we recall from (4.3.16)–(4.3.17) that with $\lambda_i \neq \lambda_j$,

$$(\lambda_i - \lambda_j)(v_i(x), v_j(x)) = [p(x)v_j(x)v_i'(x) - p(x)v_i(x)v_j'(x)]\Big|_{x=0}^{x=l}. \tag{4.3.58}$$

For the right side to vanish so that orthogonality results, not only must v_i and v_j be bounded at $x = 0$ and $x = l$ but also, $\lim p(x)v_i'(x) = \lim p(x)v_j'(x) = 0$ as $x \to 0$ and $x \to l$ if a singularity exists at $x = 0$ and $x = l$.

For our eigenvalue problem (4.3.56)–(4.3.57), the singular solution of Bessel's equation is rejected in view of (4.3.57), and the eigenfunctions must be determined from the Bessel functions of order n, that is, $J_n(\sqrt{\lambda}x)$. Since $J_0(\sqrt{\lambda}x) = 1 + O(x^2)$ and $J_n(\sqrt{\lambda}x) = O(x^n)$ as $x \to 0$, we find that these functions are bounded at $x = 0$ and that $xJ_0'(\sqrt{\lambda}x) = O(x^2)$ and $xJ_n'(\sqrt{\lambda}x) = O(x^n)$ there. Also, $p(x) = x$, so that $\lim_{x\to 0} p(x)v'(x) = 0$ is satisfied.

The eigenvalues are determined from the boundary condition $v(l) = 0$. We denote the positive zeros of the Bessel function $J_n(x)$ by α_{kn}, $k = 1, 2, \ldots$, so that $J_n(\alpha_{kn}) = 0$. [There are infinitely many real positive zeros of the Bessel functions

$J_n(x)$ with a limit point at infinity, and they are tabulated or can be found using Maple, for example.] Then $v(l) = 0$ implies that $J_n(\sqrt{\lambda}\, l) = 0$ and the eigenvalues for each n are given by

$$\lambda_{kn} = \left(\frac{\alpha_{kn}}{l}\right)^2, \qquad k = 1, 2, ... \tag{4.3.59}$$

Using the properties of the Bessel function, it is shown in the exercises that the square of the norm of the eigenfunctions $\hat{v}_{kn}(x) = J_n(\sqrt{\lambda_{kn}}\, l)$ is

$$||J_n(\sqrt{\lambda_{kn}}\, x)||^2 = \int_0^l x J_n^2(\sqrt{\lambda_{kn}}\, x)\, dx = \frac{l^2}{2} J_{n+1}^2(\sqrt{\lambda_{kn}}\, l). \tag{4.3.60}$$

In view of the orthogonality property demonstrated above, the set

$$v_{kn}(x) = \frac{\sqrt{2}}{l} \frac{J_n(\sqrt{\lambda_{kn}}\, x)}{|J_{n+1}(\sqrt{\lambda_{kn}}\, l)|}, \qquad k = 1, 2, \ldots, \tag{4.3.61}$$

is an orthonormal set of eigenfunctions for the Sturm-Liouville problem (4.3.56) and (4.3.57) for each n. These eigenfunctions form a complete set, and any smooth function $v(x)$ defined over the interval $0 \le x \le l$ that satisfies the boundary condition $v(l) = 0$ can be expanded in a convergent series

$$v(x) = \sum_{k=1}^{\infty} (v(x), v_{kn}(x)) v_{kn}(x). \tag{4.3.62}$$

Legendre Polynomial Eigenfunctions and Their Series

The eigenvalue problem for the self-adjoint form of the *Legendre equation* is

$$L[v(x)] = -\frac{d}{dx}\left[(1 - x^2)\frac{dv(x)}{dx}\right] = \lambda v(x), \tag{4.3.63}$$

where $v(x)$ is defined over the interval $-1 < x < 1$ and satisfies the boundary conditions

$$v(x) \quad \text{bounded at} \quad x = \pm 1. \tag{4.3.64}$$

As $p(x) = 1 - x^2 > 0$ for $-1 < x < 1$ and vanishes at $x = \pm 1$, (4.3.63)–(4.3.64) is a *singular Sturm-Liouville problem.* The points $x = \pm 1$ are *regular singular points* for Legendre's equation, and bounded solutions exist.

The *Frobenius theory* of power series solutions for (4.3.63) shows that the only bounded solutions of the Legendre equation (4.3.63) over $-1 \le x \le 1$ are polynomial solutions, and these occur only if λ assumes the eigenvalues

$$\lambda_k = k(k + 1), \qquad k = 0, 1, 2, ... \tag{4.3.65}$$

The eigenfunctions that correspond to the eigenvalues λ_k are usually normalized to yield the *Legendre polynomials* of degree k. They are given as

$$P_k(x) = \left(\frac{1}{2^k k!}\right) \frac{d^k}{dx^k}(x^2 - 1)^k = \sum_{n=0}^{k} \frac{(n+k)!}{(k-n)!\, 2^n (n!)^2}(x-1)^n, \qquad (4.3.66)$$

with the normalization chosen such that $P_k(1) = 1$ for all k. Since we have $\lim_{x=\pm 1}(1 - x^2)P_k'(x) = 0$, the supplementary condition that $p(x)v'(x)$ have a finite limit at the singular points is satisfied [see the discussion leading to (4.3.58)]. Consequently, the Legendre polynomials form an orthogonal set over the interval $-1 < x < 1$. Further, it can be shown that

$$||P_k(x)||^2 = \int_{-1}^{1} P_k^2(x)\, dx = \frac{2}{2k+1}, \qquad k = 0, 1, 2, \ldots. \qquad (4.3.67)$$

Thus the orthonormal set of eigenfunctions is given as

$$v_k(x) = \sqrt{\frac{2k+1}{2}} P_k(x), \qquad k = 0, 1, 2, \ldots. \qquad (4.3.68)$$

The set $\{v_k(x)\}$ is complete and any smooth function $v(x)$ defined over the interval $-1 \le x \le 1$ can be expanded in a convergent series of eigenfunctions

$$v(x) = \sum_{k=0}^{\infty} \left(v(x), \sqrt{\frac{2k+1}{2}} P_k(x)\right) \sqrt{\frac{2k+1}{2}} P_k(x). \qquad (4.3.69)$$

Exercises 4.3

4.3.1. Let $\rho(x) = 1$ in the norm (4.3.4). Show that the function $\phi(x) = 1/\sqrt{x}$ is integrable (as an improper integral) over the interval $0 < x < 1$, but that it is not square integrable.

4.3.2. Let $\rho(x) = 1$ in the norm (4.3.4). Show that the function $q(x)$, defined to equal 1 when x is a rational number and to equal -1 when x is irrational, is not integrable over any finite interval but is square integrable.

4.3.3. Show that the square integrable functions $\hat{\phi}_k(x) = \sin(\pi k \log x/\log 2), k \ge 1$, are orthogonal over the interval $1 < x < 2$ with respect to the inner product with weight function $\rho(x) = 1/x$. Obtain the norms of these functions and construct an orthonormal set.

4.3.4. Show that the set of complex-valued functions $\hat{\phi}_k(x) = \exp(ikx)$, $k = 0, \pm 1, \pm 2, \ldots$, is orthogonal with respect to the Hermitian inner product (4.3.7) over the interval $-\pi < x < \pi$ if the weight function $\rho(x) = 1$. Determine the norms of the $\hat{\phi}_k(x)$ and construct an orthonormal set.

4.3.5. Obtain the Fourier coefficients of the function $\phi(x) = x$ with respect to the orthonormal set of functions obtained in Exercise 4.3.3.

4.3.6. Adapt the argument given following (4.2.20) to show that the boundary terms in (4.3.16) vanish.

4.3.7. Conclude from (4.3.22) that if $\lambda = 0$ is an eigenvalue of the (regular) Sturm-Liouville problem, the corresponding eigenfunction $v(x)$ must satisfy $\int_0^l [p(x)v'(x)^2 + q(x)v^2(x)]dx - p(x)v(x)v'(x)\Big|_0^l = 0$, and we must have $q(x)v^2(x) = p(x)v'(x)^2 = 0$, since the boundary contributions are nonnegative. Noting that $v(x)$ cannot vanish identically, show that $q(x) = 0$ and $v(x)$ = constant. This must, in turn, imply that $\alpha_1 = \alpha_2 = 0$, since the boundary conditions could not be satisfied otherwise.

4.3.8. Suppose that there are two linearly independent eigenfunctions $v(x)$ and $w(x)$ that correspond to the eigenvalue λ. Since they both satisfy the same equation and boundary conditions, show that we must have $\alpha_1 v(0) - \beta_1 v'(0) = 0$, $\alpha_1 w(0) - \beta_1 w'(0) = 0$. Noting that α_1 and β_1 cannot both vanish, show that the determinant of the system above, which equals the Wronskian of $v(x)$ and $w(x)$ evaluated at zero, must vanish. Thereby, conclude that $v(x)$ and $w(x)$ are linearly dependent, so that each eigenvalue must be simple.

4.3.9. Consider the transcendental equation (4.3.32) for the determination of the eigenvalues in Section 4.3. Let $\lambda = \rho^2$ and show that the equation can be written as $\tan(\rho l) = \rho(\alpha_1\beta_2 + \alpha_2\beta_1)/(\rho^2\beta_1\beta_2 - \alpha_1\alpha_2)$. Put $x = \tan(\rho l)$ and $x = \rho(\alpha_1\beta_2 + \alpha_2\beta_1)/(\rho^2\beta_1\beta_2 - \alpha_1\alpha_2)$. Graph both functions in the (ρ, x)-plane and show thereby that there are an infinite number of eigenvalues λ_k and that $\lambda_k \to \infty$ as $k \to \infty$. Show that for large k, $\lambda_k \approx (nk/l)^2$.

4.3.10. Obtain the equation that corresponds to (4.3.32) if we put $\lambda = 0$ in (4.3.29). Show that zero is an eigenvalue only if $\alpha_1 = \alpha_2 = 0$ in the boundary conditions (4.3.2) and find the corresponding eigenfunction.

4.3.11. Use (4.3.1) and integrate by parts to show that if $v(x)$ and the eigenfunctions $v_k(x)$ satisfy the same boundary conditions and $v(x)$ is sufficiently smooth, the Fourier coefficients of $v(x)$ are given as $(v(x), v_k(x)) = (1/\lambda_k)\,(v_k(x), Lv(x)/\rho(x))$, $\lambda_k \neq 0$. It can be shown that $\lambda_k \approx c^2k^2$ as $k \to \infty$ for any (regular) Sturm-Liouville problem. (The constant c may be different for each problem.) Conclude thereby that if $|v_k(x)|$ and $|(v_k(x), Lv(x)/\rho(x))|$ are uniformly bounded, the Fourier series (4.3.24) converges uniformly and absolutely. *Hint*: Use the Weierstrass M-test.

4.3.12. Verify the results of Exercise 4.3.11 for the case of the Fourier sine series assuming that $v(x)$ has two continuous derivatives and that $v(0) = v(l) = 0$.

4.3.13. Determine the eigenvalues and eigenfunctions of the problem $v''(x) + \lambda v(x) = 0$, $v'(0) = 0$, $v(l) = 0$. Show that the properties 1–4 in the text are valid for this problem and normalize the set of eigenfunctions.

4.3.14. Determine the eigenvalues (approximately) and the eigenfunctions of the problem $v''(x) + \lambda v(x) = 0$, $v(0) = 0$, $v'(l) + \beta v(l) = 0$, where $\beta > 0$. Show that the properties 1–4 in the text are satisfied and normalize the set of eigenfunctions.

4.3.15. Find the eigenvalues and eigenfunctions of the problem $[(1+x)^2 v'(x)]' + \lambda v(x) = 0,\ v(0) = v(l) = 0$. Verify the properties 1–4 in the text and normalize the eigenfunctions. *Hint*: Put $t = 1 + x$ and transform the given equation into an Euler (homogeneous) equation.

4.3.16. Obtain the Fourier sine and cosine series of the following functions given over the interval $0 < x < l$; (a) $v(x) = x$; (b) $v(x) = x(x-l)$; (c) $v(x) = x^2(x-l)^2$; (d) $v(x) = \cos(5\pi x/l)$. Construct the even and odd extensions of each of the functions above. Compare the rates of convergence for each of the Fourier series based on the smoothness properties of the even and odd extensions.

4.3.17. Obtain the Fourier series expansion of $v(x) = \sin^2(\pi x/l)(x-l)^2$ over the interval $-l < x < l$. Discuss the convergence properties of the series.

4.3.18. With $\lambda = \lambda_{kn}$, $v(x) = J_n[\sqrt{\lambda_{kn}}\, x]$ and $t = \sqrt{\lambda_{kn}}\, x$ in (4.3.56), multiply across by tJ_n' and integrate from 0 to $\sqrt{\lambda_{kn}}\, l$ to obtain

$$-\int_0^{\sqrt{\lambda_{kn}}\, l} \frac{d}{dt}[(tJ_n')^2]\, dt + n^2 \int_0^{\sqrt{\lambda_{kn}}\, l} (J_n^2)'\, dt = \int_0^{\sqrt{\lambda_{kn}}\, l} t^2 (J_n^2)'\, dt$$

and conclude that $||J_n(\sqrt{\lambda_{kn}}x)||^2 = (1/2\lambda_{kn})[(\lambda_{kn}l^2 - n^2)[J_n(\sqrt{\lambda_{kn}}\, l)]^2 + \lambda_{kn}l^2 [J_n'(\sqrt{\lambda_{kn}}\, l)]^2]$. (a) In the case where $v(l) = 0$, determine the normalization constant (4.3.60) by using the formula $J_n'(x) = -J_{n+1}(x) + (n/x)J_n(x)$. (b) Determine the eigenvalues λ_{kn} and the normalization constant for the eigenfunctions if the boundary condition is $v'(l) = 0$. (c) Determine the eigenvalues λ_{kn} and the normalization constant for the eigenfunctions if the boundary condition is $v'(l) + hv(l) = 0$, with positive h. [In parts (b) and (c) the eigenvalues are given in terms of zeros of equations that involve Bessel functions and/or their derivatives.] (d) Show that zero is an eigenvalue only if $n = 0$ and the boundary condition is $v'(l) = 0$. Determine the normalized eigenfunction in that case.

4.3.19. Given the Taylor expansion $J_0(x) = \sum_{j=0}^{\infty} (-1)^j/(j!)^2\, (x/2)^{2j}$, show that the first three (positive) zeros of $J_0(x)$ are (approximately) $\alpha_{10} = 2.41$, $\alpha_{20} = 5.52$, and $\alpha_{30} = 8.65$. [Note that the series for $J_0(x)$ is alternating.]

4.3.20. Use $J_n(\sqrt{\lambda}\, x) \approx \sqrt{2/(\pi\sqrt{\lambda}\, x)} \cos\left(\sqrt{\lambda}x - \pi n/2 - \pi/4\right)$, which is valid as $\sqrt{\lambda}\, x \to \infty$, to show that the large eigenvalues (4.3.59) are given approximately as $\lambda_{kn} \approx (\pi k/l)^2$.

4.3.21. Expand the function $v(x) = 1$ in a series of eigenfunctions $v_{k,0}(x)$ as given in (4.3.61)–(4.3.62). *Hint*: Use the formulas in Exercise 4.3.22.

4.3.22. Expand $v(x) = l^2 - x^2$ in a series of eigenfunctions $v_{k,0}(x)$ as given in (4.3.61)–(4.3.62). Evaluate the Fourier coefficients by using integration by parts and the formulas $[x^n J_n(x)]' = x^n J_{n-1}(x)$ and $J_{-n}(x) = (-1)^n J_n(x)$.

4.3.23. Obtain the first five Legendre polynomials $P_0(x), \ldots, P_4(x)$ and show that they are mutually orthogonal.

4.3.24. Use the derivative definition of the Legendre polynomials $P_k(x)$ given in (4.3.66) (this is known as *Rodrigues' formula*) to obtain the normalization constants (4.3.67). *Hint*: Integrate by parts as often as necessary.

4.3.25. Expand the following functions in a series of Legendre polynomials: (a) $v(x) = x$; (b) $v(x) = 5 - 4x + 10x^3$; (c) $v(x) = x^4$.

4.3.26. Determine the eigenvalues and eigenfunctions for the problem $y''(x) + \lambda y(x) = 0, \quad 0 < x < l, \; y'(0) + y(0) = 0, \; y(l) = 0$. Note that the boundary condition at $x = 0$ does not have the required form (4.2.3) and show that there is one negative eigenvalue. (Show by graphical means how the eigenvalues are distributed.)

4.3.27 Determine the eigenvalues and eigenfunctions for the eigenvalue problem $y''(x) + 2y'(x) + (1 + \lambda)y(x) = 0, \; 0 < x < \pi, \; y'(0) = y'(\pi) = 0$. (Note that this is not a Sturm-Liouville problem since the equation is not self-adjoint. It can be reduced to self-adjoint form via a transformation of the dependent variable.)

4.3.28. Show that the eigenvalue problem $y''(x) - 2y'(x) + (1 + \lambda)y(x) = 0, \; 0 < x < \pi, \; y'(0) = 2y(0), \; y'(\pi) = 2y(\pi)$ is the adjoint problem for that of in Exercise 4.3.27, in the sense of Exercise 4.2.14. Show that the eigenvalues for both problems are the same and that the eigenfunctions for the original and the adjoint problems are orthogonal if they correspond to different eigenvalues.

4.4 SERIES SOLUTIONS OF BOUNDARY AND INITIAL AND BOUNDARY VALUE PROBLEMS

In this section the *method of separation of variables* is applied to specific boundary and initial and boundary value problems for PDEs. Examples dealing with initial and boundary value problems for the *wave* and the *heat equations*, and a boundary value problem for *Laplace's equation*, are presented. The solution of each of the problems is discussed in some detail.

Example 4.2. The Wave Equation: Vibrations of a Fixed String. Let $u(x,t)$ represent the transverse displacement of a tightly stretched string of length l. Under various simplifying assumptions it is found that $u(x,t)$ satisfies the *wave equation*

$$u_{tt}(x,t) - c^2 u_{xx}(x,t) = 0, \qquad 0 < x < l, \; t > 0, \tag{4.4.1}$$

with $c = \sqrt{T/\rho}$, where T is the tension and ρ is the density of the string, both of which are assumed to be constant.

Now $u = u(x,t)$ is the graph or shape of the string at the time t, and $u_x(x,t)$ is the slope at a point x on the string. We assume that $|u_x| \ll 1$, so that the curvature of the string approximately equals $u_{xx}(x,t)$. The mass per unit length of the string equals the density ρ of the string. Then *Newton's law of motion* states that ρu_{tt} at a point x on the string equals the transverse force acting on it. All forces acting on the string are neglected compared to the force due to the string's tension. This force

is proportional to the curvature u_{xx}, with a constant of proportionality equal to the tension T. The greater the curvature, the greater this (tension) force, whose effect is to straighten out the string. As a result, $u(x,t)$ satisfies the wave equation (4.4.1).

The ends of the string are assumed to be fixed for all time, and this yields the boundary conditions

$$u(0,t) = 0, \qquad u(l,t) = 0, \qquad t > 0. \tag{4.4.2}$$

The prescribed displacements and velocities at $t = 0$ yield the initial conditions

$$u(x,0) = f(x), \qquad u_t(x,0) = g(x), \qquad 0 < x < l. \tag{4.4.3}$$

To solve the initial and boundary value problem (4.4.1)–(4.4.3), we apply the separation of variables method of Section 4.2 and set $u(x,t) = M(x)N(t)$. Then $M(x)$ satisfies the following eigenvalue problem:

$$-c^2 M''(x) = \lambda M(x), \qquad M(0) = M(l) = 0. \tag{4.4.4}$$

An equivalent Sturm-Liouville problem was considered in Section 4.3 in connection with the Fourier sine series. [We must replace λ by λ/c^2 in (4.3.29).] Then, the eigenvalues and the normalized eigenfunctions for (4.4.4) are

$$\lambda_k = \left(\frac{\pi k c}{l}\right)^2, \quad M_k(x) = \sqrt{\frac{2}{l}} \sin\left(\frac{\pi k}{l} x\right), \quad k = 1, 2, \ldots. \tag{4.4.5}$$

From (4.2.31) we obtain

$$N_k(t) = a_k \cos\left(\frac{\pi k c}{l} t\right) + b_k \sin\left(\frac{\pi k c}{l} t\right), \quad k = 1, 2, \ldots, \tag{4.4.6}$$

and the corresponding solutions $u_k(x,t)$, which are known as *normal modes*,

$$u_k(x,t) = \left[a_k \cos\left(\frac{\pi k c}{l} t\right) + b_k \sin\left(\frac{\pi k c}{l} t\right)\right] \sqrt{\frac{2}{l}} \sin\left(\frac{\pi k}{l} x\right), \quad k = 1, 2, \ldots. \tag{4.4.7}$$

Each $u_k(x,t)$ satisfies (4.4.1) and the boundary conditions (4.4.2).

To satisfy the initial conditions (4.4.3), we form the sum (4.2.32), set

$$u(x,t) = \sum_{k=1}^{\infty} u_k(x,t), \tag{4.4.8}$$

and apply the initial conditions to $u(x,t)$. This gives us

$$u(x,0)=\sum_{k=1}^{\infty}u_k(x,0)=\sum_{k=1}^{\infty}a_k\sqrt{\frac{2}{l}}\sin\left(\frac{\pi k}{l}x\right)=f(x), \tag{4.4.9}$$

$$u_t(x,0)=\sum_{k=1}^{\infty}\frac{\partial u_k(x,0)}{\partial t}=\sum_{k=1}^{\infty}\left(\frac{\pi kc}{l}\right)b_k\sqrt{\frac{2}{l}}\sin\left(\frac{\pi k}{l}x\right)=g(x). \tag{4.4.10}$$

These are Fourier sine series with the eigenfunctions $M_k=\sqrt{2/l}\sin[(\pi k/l)x]$ (they are normalized) and (4.2.35) yields

$$a_k=(f(x),M_k(x))=\sqrt{\frac{2}{l}}\int_0^l f(x)\,\sin\left(\frac{\pi k}{l}x\right)dx,\ k=1,2,\ldots, \tag{4.4.11}$$

$$b_k=\left(\frac{g(x)}{\sqrt{\lambda_k}},M_k(x)\right)=\frac{\sqrt{2l}}{\pi kc}\int_0^l g(x)\,\sin\left(\frac{\pi k}{l}x\right)dx,\ k=1,2,\ldots. \tag{4.4.12}$$

This completes the formal determination of the solution.

To assure that the formal series solution (4.4.8) represents a classical solution of the wave equation and satisfies the initial conditions, we must place conditions on the data $f(x)$ and $g(x)$. The function $f(x)$ is required to have two continuous derivatives and a third piecewise continuous derivative, and in addition we must have $f(0)=f(l)=0$ and $f''(0)=f''(l)=0$. The function $g(x)$ must be continuously differentiable and have a piecewise continuous second derivative as well as satisfying the conditions $g(0)=g(l)=0$. Under these conditions the series (4.4.8) can be differentiated term by term twice and the resulting series converge uniformly. Then $u(x,t)$ is a solution of the wave equation since each $u_k(x,t)$ satisfies the equation. The conditions on $f(x)$ and $g(x)$ guarantee that the sine series as well as the series of its derivatives converges, since it is required that $f(x)$ and $g(x)$ satisfy the boundary conditions for the appropriate Sturm-Liouville problems to assure the uniform convergence of the Fourier series, as indicated in the preceding section (see Exercises 4.3.11 and 4.3.12).

We now present a brief interpretation of the solution (4.4.8). Each of the normal modes $u_k(x,t)$ [i.e., (4.4.7)] can be expressed as

$$u_k(x,t)=\alpha_k\cos\left(\frac{\pi kc}{l}(t+\delta_k)\right)\sin\left(\frac{\pi k}{l}x\right),\qquad k=1,2,\ldots, \tag{4.4.13}$$

$$\alpha_k=\sqrt{\frac{2}{l}}\sqrt{a_k^2+b_k^2},\quad \delta_k=-\frac{l}{\pi kc}\arctan\left(\frac{b_k}{a_k}\right),\quad k=1,2,\ldots. \tag{4.4.14}$$

For each solution $u_k(x,t)$, a fixed point x_0 executes *harmonic vibrations* with amplitude

$$A_k=\alpha_k\left|\sin\left(\frac{\pi k}{l}x_0\right)\right|. \tag{4.4.15}$$

Solutions of the wave equation are called *waves* and we refer to $u_k(x,t)$ as a *standing wave*, since each point x on the string oscillates in place as t varies. The points $x_m=ml/k$ $(m=1,2,\ldots,k-1)$ at which $\sin[(\pi k/l)x_m]=0$ remain fixed during the vibration and are called the *nodes* of the standing wave. The standing wave has

maximum amplitudes at the points $x_n = (2n+1)l/2k$ $(n = 0, 1, 2, \ldots, k-1)$ at which $\sin[(\pi k/l)x_n] = \pm 1$, and they are called *antinodes*. The standing wave $u_k(x,t)$ is also known as the kth *harmonic*. Thus,

$$\omega_k = \frac{\pi k c}{l} = \sqrt{\lambda_k}, \qquad k = 1, 2, 3, \ldots, \tag{4.4.16}$$

is the *frequency of vibration* of the kth harmonic $u_k(x,t)$.

The *energy of the vibrating string* at the time t is given by the integral

$$E(t) = \frac{1}{2}\int_0^l \left\{ \rho \left(\frac{\partial u(x,t)}{\partial t} \right)^2 + T \left(\frac{\partial u(x,t)}{\partial x} \right)^2 \right\} dx. \tag{4.4.17}$$

It represents a sum of the *kinetic energy* of motion and the *potential energy* due to the tension. The total energy is a sum of the energies distributed among the harmonics $u_k(x,t)$. The energy of the kth harmonic $u_k(x,t)$ is

$$E_k(t) = \frac{1}{2}\int_0^l \left\{ \rho \left(\frac{\partial u_k(x,t)}{\partial t} \right)^2 + T \left(\frac{\partial u_k(x,t)}{\partial x} \right)^2 \right\} dx = \frac{\omega_k^2 m (a_k^2 + b_k^2)}{2l}, \tag{4.4.18}$$

where $m = \rho l$ is the mass of the string. We observe that $E_k(t)$ is independent of t, so that the total energy

$$E(t) = \sum_{k=1}^{\infty} E_k(t) = \frac{m}{2l} \sum_{k=1}^{\infty} \omega_k^2 (a_k^2 + b_k^2) \tag{4.4.19}$$

is constant in time, and $E(t) = E(0)$. The initial energy is conserved since we have neglected dissipative effects, consistent with our characterization of the wave equation in Example 3.7 as being of conservative type.

It was shown in Example 2.4 that the solution of the initial value problem for the wave equation can be written as the sum of two *propagating* or *traveling waves*; that is, it has the form

$$u(x,t) = \hat{F}(x - ct) + \hat{G}(x + ct), \tag{4.4.20}$$

where $\hat{F}$ and $\hat{G}$ are determined from the initial conditions. The normal modes $u_k(x,t)$ can also be expressed as the sum of propagating waves as

$$\begin{aligned} u_k(x,t) = {} & \frac{1}{\sqrt{2l}} \left(a_k \sin\left(\frac{\pi k}{l}(x + ct) \right) + a_k \sin\left(\frac{\pi k}{l}(x - ct) \right) \right) \\ & + \frac{1}{\sqrt{2l}} \left(b_k \cos\left(\frac{\pi k}{l}(x - ct) \right) - b_k \cos\left(\frac{\pi k}{l}(x + ct) \right) \right). \end{aligned} \tag{4.4.21}$$

Although there are four *propagating waves* in (4.4.21), they combine to form a *standing wave* through interference. In this representation, $u(x,t)$ equals

$$u(x,t) = \frac{1}{\sqrt{2l}} \left[\sum_{k=1}^{\infty} a_k \sin\left(\frac{\pi k}{l}(x + ct) \right) + \sum_{k=1}^{\infty} a_k \sin\left(\frac{\pi k}{l}(x - ct) \right) \right], \tag{4.4.22}$$

where we have assumed for simplicity that $g(x) = 0$, so that $b_k = 0$ for all k.

Apart from a factor $\frac{1}{2}$, these series are identical in form to the Fourier sine series (4.4.9) for $f(x)$, if the argument x is replaced by $x \pm ct$. Now the initial value $f(x)$ is not defined outside the interval $[0, l]$. Proceeding as in Section 4.3, we define the function $F(x)$ as

$$\begin{cases} F(x) = f(x), & 0 \le x \le l, \\ F(x) = -f(-x), & -l < x < 0, \\ F(x + 2l) = F(x), & -\infty < x < \infty, \end{cases} \tag{4.4.23}$$

so that $F(x)$ is an odd periodic extension of $f(x)$ of period $2l$ over the entire x-axis. Then the Fourier series expansion of $F(x)$ is identical to the sine series (4.4.9) for $f(x)$, since $F(x)$ is an odd periodic function. Since $F(x)$ is defined for all x, the Fourier series for the periodic function $F(x \pm ct)$ can be constructed. Its form is precisely that of (4.4.9) with x replaced by $x \pm ct$. Consequently, the solution (4.4.22) can be expressed as

$$u(x,t) = \frac{1}{2}F(x + ct) + \frac{1}{2}F(x - ct), \qquad 0 < x < l, t > 0. \tag{4.4.24}$$

The conditions under which (4.4.24) represents a solution of the initial and boundary value problem for the wave equation [with $g(x) = 0$], based on its construction as an odd periodic extension of $f(x)$, can be shown to be slightly weaker than those given for the standing wave representation of the solution. The *traveling wave* representation (4.4.24) of the solution of the wave equation can be interpreted in terms of waves being continually reflected off the ends of the string. As a result of interference, only standing wave motion seems to occur when the vibration of a string is observed over a length of time.

As indicated previously, the Fourier series representation of $F(x)$ does not smooth out the discontinuities of $F(x)$. Thus unless $F(x)$ is a smooth function, the solution $u(x,t)$ may have discontinuities or discontinuous derivatives along the curves $x \pm ct$ = constant, in view of (4.4.24). This shows that discontinuities or discontinuous derivaties of solutions of the wave equation can occur only along the characteristic curves, as shown in Section 3.2. Later in the book we show how the concept of solution may be generalized to include functions that do not have the required number of derivatives.

Example 4.3 The Heat Equation: Heat Conduction in a Finite Rod. Let $u(x,t)$ represent the temperature in a homogeneous, laterally insulated rod of length l at the point x $(0 < x < l)$ and at the time t. Since no heat can escape through the sides of the rod, the problem is effectively one-dimensional and in the absence of heat sources, $u(x,t)$ satisfies the *heat equation*

$$u_t(x,t) - c^2 u_{xx}(x,t) = 0, \qquad 0 < x < l, \ t > 0, \tag{4.4.25}$$

where c^2 is the coefficient of heat conduction. As in Section 4.1, c^2 is the thermal conductivity divided by the product of the specific heat and the density of the medium,

all of which are taken to be constant. The ends of the rod at $x = 0$ and $x = l$ are assumed to be kept at zero temperature, so that

$$u(0,t) = 0, \qquad u(l,t) = 0, \qquad t > 0, \tag{4.4.26}$$

whereas initially, at $t = 0$, there is a given temperature distribution

$$u(x,0) = f(x), \qquad 0 < x < l. \tag{4.4.27}$$

We apply the *separation of variables* technique to solve this problem and proceed as in Example 4.2. With $u_k(x,t) = M_k(x)N_k(t)$, the eigenvalues λ_k and the eigenfunctions M_k are given as in (4.4.5) and $N_k(t)$ is found to be [see (4.2.31)]

$$N_k(t) = a_k \exp\left[-\left(\frac{\pi kc}{l}\right)^2 t\right], \qquad k = 1, 2, \ldots. \tag{4.4.28}$$

The solution $u(x,t)$ becomes

$$u(x,t) = \sum_{k=1}^{\infty} M_k(x)N_k(t) = \sqrt{\frac{2}{l}} \sum_{k=1}^{\infty} a_k e^{-(\pi kc/l)^2 t} \sin\left(\frac{\pi k}{l}x\right). \tag{4.4.29}$$

Initially, we have

$$u(x,0) = \sum_{k=1}^{\infty} a_k \sqrt{\frac{2}{l}} \sin\left(\frac{\pi k}{l}x\right) = f(x), \tag{4.4.30}$$

so that the a_k are defined as in (4.4.11). Thus the formal solution of (4.4.25)–(4.4.27) is specified completely.

To examine the validity of the expansion (4.4.29), we note that for $t > 0$ the terms in the series decay exponentially. If we assume that the initial temperature $f(x)$ is bounded over the interval $[0, l]$, the Fourier coefficients a_k are bounded. As a result, the series for $u(x,t)$ with $t > 0$ can be differentiated term by term as often as required. Since each term $u_k(x,t)$ satisfies the heat equation, so does $u(x,t)$ for $t > 0$ and we also have $u(0,t) = u(l,t) = 0$. To assure that the initial condition is satisfied, the Fourier series (4.4.30) must converge uniformly. To achieve this, we require that $f(x)$ be continuous and piecewise continuously differentiable in $0 \le x \le l$ and that $f(0) = f(l) = 0$. The uniform convergence of the Fourier sine series (4.4.30) implies that the series (4.4.29) for $u(x,t)$ converges uniformly for $t \ge 0$ and that the solution is continuous for $0 \le x \le l$ and $t \ge 0$.

In the vibrating string problem of Example 4.2 (which involved a hyperbolic equation) the solution as given in (4.4.22) essentially has the form of a Fourier series. This has the effect that any discontinuities in the data are preserved and transmitted along the characteristic lines $x \pm ct$ = constant in (x,t)-space. For the heat equation, however, even if the initial temperature $f(x)$ is discontinuous but bounded, the solution $u(x,t)$ is continuously differentiable as often as is required. The more derivatives

a function has, the smoother it is, so that heat conduction is seen to be a smoothing process. Even if the heat is initially concentrated near one point in $0 < x < l$, it is distributed instantaneously, according to our solution, in an even and smooth fashion throughout the interval. Further, as $t \to \infty$, the temperature $u(x, t) \to 0$ since there are no heat sources in this problem and the rod is continually being cooled at its endpoints.

As indicated in our discussion of the diffusion equation in Chapter 1, disturbances move at infinite speeds for the heat equation, and this fact is borne out by the properties of the solution obtained. It may also be noted that the characteristics of the heat equation are the lines t = constant. Since characteristics are carriers of discontinuities or rapid variations of the solution, there is no mechanism whereby these effects can be transmitted into the region $t > 0$ from the initial line $t = 0$, which is itself a characteristic curve.

Example 4.4. Laplace's Equation in a Rectangle. We consider the boundary value problem for *Laplace's equation*

$$u_{xx}(x, y) + u_{yy}(x, y) = 0 \tag{4.4.31}$$

in the rectangle $0 < x < l$, $0 < y < \hat{l}$, with the boundary conditions

$$u(x, 0) = f(x),\ u(x, \hat{l}) = g(x),\ u(0, y) = 0,\ u(l, y) = 0. \tag{4.4.32}$$

The function $u(x, y)$ may represent the steady-state displacement of a stretched membrane whose rectangular boundaries are fixed according to (4.4.32). [Proceeding similarly, we can solve another problem for (4.4.31) with nonzero boundary conditions on $x = 0$ and $x = l$ and zero boundary conditions on $y = 0$ and $y = \hat{l}$. Adding the solutions of both problems yields a solution of Laplace's equation with arbitrary boundary conditions on a rectangle.]

Applying the *separation of variables method* of Section 4.2 and proceeding as in Example 4.2, we obtain the eigenfunctions $M_k(x)$ as in (4.4.5) and the eigenvalues $\lambda_k = (\pi k/l)^2$. With $u_k(x, y) = M_k(x)N_k(y)$, the $N_k(y)$ are given as [see (4.2.31)]

$$N_k(y) = \hat{a}_k \exp\left(\frac{\pi k}{l}\, y\right) + \hat{b}_k \exp\left(-\frac{\pi k}{l}\, y\right), \qquad k = 1, 2, \ldots. \tag{4.4.33}$$

The functions $N_k(y)$ can be expressed as linear combinations of hyperbolic sine functions and we write $u_k(x, y)$ as

$$u_k(x, y) = \left[a_k \sinh\left(\frac{\pi k}{l}\, y\right) + b_k \sinh\left(\frac{\pi k}{l}(y - \hat{l})\right)\right] \sqrt{\frac{2}{l}} \sin\left(\frac{\pi k}{l} x\right), \tag{4.4.34}$$

with the a_k and b_k as yet arbitrary. A formal solution of Laplace's equation obtained by superposition is

$$u(x, y) = \sum_{k=1}^{\infty} u_k(x, y). \tag{4.4.35}$$

The boundary conditions at $y = 0$ and $y = \hat{l}$ imply that

$$u(x,0) = \sum_{k=1}^{\infty} b_k \sinh\left(-\frac{\pi k}{l}\hat{l}\right)\sqrt{\frac{2}{l}}\sin\left(\frac{\pi k}{l}x\right) = f(x), \tag{4.4.36}$$

$$u(x,\hat{l}) = \sum_{k=1}^{\infty} a_k \sinh\left(\frac{\pi k}{l}\hat{l}\right)\sqrt{\frac{2}{l}}\sin\left(\frac{\pi k}{l}x\right) = g(x). \tag{4.4.37}$$

The Fourier coefficients of these sine series are [see (4.4.11)–(4.4.12)]

$$\begin{cases} b_k = (f(x), M_k(x))/\sinh(-\pi k\hat{l}/l), & k = 1,2,\ldots, \\ a_k = (g(x), M_k(x))/\sinh(\pi k\hat{l}/l), & k = 1,2,\ldots. \end{cases} \tag{4.4.38}$$

This completes the formal solution of the boundary value problem.

Assuming that the integrals $\int_0^l |f(x)|\,dx$ and $\int_0^l |g(x)|\,dx$ are bounded by m, and noting that $\sinh x = (e^x - e^{-x})/2 \le (1/2)e^x$ for $x \ge 0$, we obtain

$$\left|b_k \sinh\left[\frac{\pi k}{l}(y-\hat{l})\right]\right| \le \frac{m\sqrt{2/l}\exp[(\pi k/l)(\hat{l}-y)]}{\exp[\pi k\hat{l}/l][1-\exp[-2\pi k\hat{l}/l]]} \le \frac{m\sqrt{2/l}\exp[-\pi k y/l]}{1-\exp[-2\pi\hat{l}/l]} \tag{4.4.39}$$

with a similar bound valid for $|a_k \sinh[(\pi k/l)y]|$. Thus the terms in the series $u(x,y) = \sum_{k=1}^{\infty} u_k(x,y)$ are bounded by exponentially decaying terms for large k in the open interval $0 < y < \hat{l}$. (The exponential bound for the b_k and a_k terms breaks down at $y = 0$ and $y = \hat{l}$, respectively.) Consequently, the series can be differentiated term by term as often as desired in $0 < y < \hat{l}$ and $u(x,y)$ satisfies Laplace's equation and the boundary conditions at $x = 0$ and $x = l$, since each of the $u_k(x,y)$ does so. To assure that $u(x,y)$ is continuous up to $y = 0$ and $y = \hat{l}$ and assumes the boundary values there, the Fourier series (4.4.36)–(4.4.37) must be uniformly convergent in $0 \le x \le l$. This is the case if $f(0) = f(l) = g(0) = g(l) = 0$, if $f(x)$ and $g(x)$ are continuous, and if $f'(x)$ and $g'(x)$ are piecewise continuous in that interval.

For $u(x,y)$ to satisfy Laplace's equation inside the rectangle we merely require that $f(x)$ and $g(x)$ be bounded or that $\int_0^l |f(x)|\,dx$ and $\int_0^l |g(x)|\,dx$ exist [i.e., $f(x)$ and $g(x)$ may even be singular]. Then $u(x,y)$ is defined by the series to be an infinitely differentiable function within the rectangle. However, the boundary values at $y = 0$ and $y = \hat{l}$ need not be assumed continuously at all points unless the additional conditions on $f(x)$ and $g(x)$ given previously are met. Again, this contrasts with the results obtained for the wave equation, where discontinuities are seen to spread into interior regions from the boundary. Since Laplace's equation has no real characteristics, boundary data discontinuities must be confined to the boundary and the interior solution is smooth. Physically, since Laplace's equation characterizes steady-state or equilibrium situations, we may expect that effects due to discontinuities in the data have smoothed themselves out.

This concludes our discussion of the separation of variables method for homogeneous problems. Further examples involving other eigenvalue problems are considered in the exercises and in Section 4.6.

Exercises 4.4

4.4.1. Determine the motion of a *plucked string*. That is, find a solution $u(x,t)$ of the wave equation (4.4.1) with $u(0,t) = u(l,t) = 0$ and the initial conditions

$$u(x,0) = f(x) = \begin{cases} hx/a, & 0 < x < a, \\ h(l-x)/(l-a), & a \le x < l, \end{cases} \qquad u_t(x,0) = g(x) = 0.$$

4.4.2. Solve the initial and boundary value problem for the wave equation (4.4.1) with $u(0,t) = u(l,t) = 0$ and the initial conditions $f(x) = 0$, $g(x) = \delta(x-a)$, $0 < a < l$, where $\delta(x)$ is the Dirac delta function. This yields the motion of a string due to a point impulse at $x = a$ administered at the time $t = 0$.

4.4.3. Solve (4.4.1)–(4.4.3) if $f(x) = 0$, $g(x) = x(x-l)$.

4.4.4. Solve (4.4.1)–(4.4.3) if $f(x) = \sin^2(\pi x/l)$, $g(x) = 0$.

4.4.5. Express the solution of Exercise 4.4.1 in the form (4.4.24) and use this result to describe the displacement $u(x,t)$ of the string at various times $t > 0$.

4.4.6. Use separation of variables to obtain the general form of the (series) solution of the problem $u_{tt}(x,t) - c^2 u_{xx}(x,t) = 0$, $0 < x < l$, $t > 0$, with the initial conditions $u(x,0) = f(x)$, $u_t(x,0) = g(x)$, and the boundary conditions $u_x(0,t) = u_x(l,t) = 0$.

4.4.7. Obtain the solution of the problem in Exercise 4.4.6 if $f(x) = 0$ and $g(x) = 1$. Interpret the result.

4.4.8. Solve the problem of Exercise 4.4.6 if $f(x) = x^2(x-l)^2$, $g(x) = 0$.

4.4.9. Using separation of variables, solve the following initial and boundary value problem $u_{tt}(x,t) - c^2 u_{xx}(x,t) = 0$, $0 < x < l$, $t > 0, u(x,0) = f(x)$, $u_t(x,0) = g(x)$, $u(0,t) = 0$, $u_x(l,t) = 0$.

4.4.10. Solve the problem in Exercise 4.4.9 if $f(x) = x(x-l)^2$, $g(x) = 0$.

4.4.11. Apply the method of separation of variables to solve the wave equation (4.4.1) with the initial data (4.4.3) and the boundary data $u(0,t) = 0$, $u_x(l,t) + \beta u(l,t) = 0$, $\beta > 0$.

4.4.12. Obtain the solution of the problem in Exercise 4.4.11 if the initial data are $f(x) = \sin^2(\pi x/l)$, $g(x) = 0$.

4.4.13. Multiply the wave equation (4.4.1) by $u(x,t)$ and make use of the identity $u_{xx}(x,t)u_t(x,t) = (u_t(x,t)u_x(x,t))_x - \frac{1}{2}(u_x^2(x,t))_t$, to obtain $\frac{1}{2}\frac{\partial}{\partial t}[u_t^2(x,t) + c^2 u_x^2(x,t)] - c^2\frac{\partial}{\partial x}[u_t(x,t)u_x(x,t)] = 0$. Integrate over the interval $0 < x < l$ and show that if either $u(x,t)$ or $u_x(x,t)$ vanishes at $x = 0$ and $x = l$, the energy integral (4.4.17) is a constant (recall that $T = c^2\rho$). Obtain an appropriate energy integral if $u_x(0,t) - \beta_1 u(0,t) = 0$ and $u_x(l,t) + \beta_2 u(l,t) = 0$ with β_1, $\beta_2 > 0$.

4.4.14. Verify the result (4.4.19).

4.4.15. Use separation of variables to solve the following problem for the telegrapher's equation $v_{tt}(x,t) - \gamma^2 v_{xx}(x,t) + 2\hat{\lambda} v_t(x,t) = 0,\ 0 < x < l,\ t > 0$, where $\hat{\lambda} > 0$, with $v(x,0) = f(x),\ v_t(x,0) = g(x),\ v(0,t) = v(l,t) = 0$. Show that the solution $v(x,t)$ tends to zero as $t \to \infty$.

4.4.16. Solve the initial and boundary value problem for the Klein-Gordon equation $u_{tt}(x,t) - \gamma^2 u_{xx}(x,t) + c^2 u_t(x,t) = 0,\ 0 < x < l,\ t > 0,\ u(x,0) = f(x),\ u_t(x,0) = g(x),\ u(0,t) = u(l,t) = 0$. Use separation of variables.

4.4.17. Solve the heat equation (4.4.25) with the boundary data (4.4.26) and the initial data $u(x,0) = f(x) = x(l-x)$.

4.4.18. Solve the problem (4.4.25)–(4.4.27) if $f(x) = \delta(x-a),\ 0 < a < l$, where $\delta(x)$ is the Dirac delta function.

4.4.19. Apply separation of variables to solve $u_t(x,t) - c^2 u_{xx}(x,t) = 0,\ 0 < x < l,\ t > 0,\ u(x,0) = f(x),\ u_x(0,t) = u_x(l,t) = 0$. The boundary conditions imply that no heat escapes through the ends of the rod. Show that as $t \to \infty$ we have $\lim_{t\to\infty} u(x,t) = (1/l)\int_0^l f(x)\,dx$. This represents the average of the initial temperature distribution.

4.4.20. Solve the problem in Exercise 4.4.19 if (a) $f(x) = x$; (b) $f(x) = \sin^2(\pi x/l)$.

4.4.21. By integrating the heat equation (4.4.25) over the interval $0 < x < l$, show that if $u_x(0,t) = u_x(l,t) = 0$, we have $\int_0^l u(x,t)\,dx =$ constant. Explain why this is consistent with the result in Exercise 4.4.19 regarding the limit of $u(x,t)$ as $t \to \infty$.

4.4.22. Use separation of variables to solve the following problem for the heat equation: $u_t(x,t) - c^2 u_{xx}(x,t) = 0,\ 0 < x < l,\ t > 0,\ u(x,0) = f(x),\ u(0,t) = 0,\ u_x(l,t) + \beta u(l,t) = 0,\ \beta > 0$. The boundary condition of the third kind at $x = l$ results from *Newton's law of cooling* if there is convective heat exchange between the rod and a medium adjacent to the rod at $x = l$ that is kept at zero temperature.

4.4.23. Solve the problem in Exercise 4.4.22 if $f(x) = x$.

4.4.24. If the lateral portion of a rod undergoes convective heat exchange with a medium kept at zero temperature, the temperature $u(x,t)$ in the rod satisfies the equation $u_t(x,t) - c^2 u_{xx}(x,t) + a^2 u(x,t) = 0,\ 0 < x < l,\ t > 0$. Assume that $u(0,t) = u(l,t) = 0,\ u(x,0) = f(x)$, and that a is a constant. Find the temperature $u(x,t)$ by the method of separation of variables. Obtain the limit as $t \to \infty$ of the temperature $u(x,t)$.

4.4.25. Use separation of variables to solve Laplace's equation $\nabla^2 u(x,y) = 0$ in the rectangle $0 < x < l,\ 0 < y < \hat{l}$, with the boundary conditions of the second kind $u_x(0,y) = u_x(l,y) = 0,\ 0 < y < \hat{l},\ u_y(x,0) = f(x),\ u_y(x,\hat{l}) = g(x),\ 0 < x < l$. Determine conditions on $f(x)$ and $g(x)$ for a solution to exist. Discuss the rate of convergence of the solution.

4.4.26. Solve the following boundary value problem for Laplace's equation $\nabla^2 u(x,y) = 0,\ 0 < x < l,\ 0 < y < \hat{l},\ u(0,y) = u(l,y) = 0,\ 0 < y < \hat{l},\ u_y(x,0) = f(x),\ u_y(x,\hat{l}) = g(x),\ 0 < x < l$.

4.4.27. Solve Laplace's equation $\nabla^2 u(x,y) = 0$ in the rectangle $0 < x < l$, $0 < y < \hat{l}$ with the boundary conditions $u(0,y) = 0$, $u_x(l,y) + \beta u(l,y) = 0$, $\beta > 0$, $0 < y < \hat{l}$, $u(x,0) = f(x)$, $u(x,\hat{l}) = 0$, $0 < x < l$.

4.4.28. Let $f(x) = x(x-l)$ and $g(x) = 0$. Solve the following boundary value problems: (a) Example 4.4; (b) Exercise 4.4.25; (c) Exercise 4.4.26.

4.4.29. Consider Laplace's equation $\nabla^2 u(x,y) = 0$ in the unbounded region $0 < x < l$, $y > 0$. Let $u(0,y) = u(l,y) = 0$, $y > 0$, $u(x,0) = f(x)$, $0 < x < l$, as well as requiring that $u(x,y)$ be (uniformly) bounded as $y \to \infty$ for $0 < x < l$. Show that the solution of this problem can be obtained by separation of variables. Find the limit of $u(x,y)$ as $y \to \infty$.

4.4.30. Apply separation of variables to solve the boundary value problem for the elliptic equation $u_{xx}(x,y) + u_{yy}(x,y) - c^2 u(x,y) = 0$, $0 < x < l$, $0 < y < \hat{l}$ in the given rectangle, with the boundary conditions $u(0,y) = u(l,y) = 0$, $u(x,0) = f(x)$, $u(x,\hat{l}) = g(x)$.

4.4.31. Use separation of variables to show that the reduced wave equation $u_{xx}(x,y) + u_{yy}(x,y) + k^2 u(x,y) = 0$, $0 < x < l$, $0 < y < \hat{l}$ with the (homogeneous) boundary condition $u(x,y) = 0$ on $x = 0$, $x = l$, $y = 0$, $y = \hat{l}$ can have nonzero solutions for certain values of k. Determine these values of $k = k_n$. (They correspond to the eigenvalues for Laplace's equation in a rectangle.) By applying Green's theorem [see (4.2.19)], conclude that if $k = k_n$, the Dirichlet problem for the reduced wave equation has no solution unless the boundary values satisfy a compatibility or orthogonality condition. [Let $w(x,y) = w_n(x,y)$ be a solution of $\nabla^2 w_n(x,y) + k_n^2 w_n(x,y) = 0$, with $w_n(x,y) = 0$ on the boundary of the rectangle, in the application of (4.2.19).] Show that if a solution does exist for $k = k_n$, it is not unique.

4.4.32. Solve the initial and boundary value problem for the two-dimensional wave equation in a disk, $u_{tt}(x,y,t) - c^2[u_{xx}(x,y,t) + u_{yy}(x,y,t)] = 0$, $x^2 + y^2 < l^2$, $t > 0$, with the initial conditions $u(x,y,0) = f(x,y)$, $u_t(x,y,0) = g(x,y)$ and the boundary condition $u(x,y,t) = 0$, $x^2 + y^2 = l^2$, $t > 0$. Use separation of variables and the results of Exercise 4.2.10 and Section 4.3. (The solution of this problem describes the *vibration of a circular membrane* with a fixed edge if $c^2 = T/\rho$, where T is the tension and ρ is the density.)

4.4.33. Solve the initial and boundary value problem for the heat equation in a disk, $u_t(x,y,t) - c^2[u_{xx}(x,y,t) + u_{yy}(x,y,t)] = 0$, $x^2 + y^2 < l^2$, $t > 0$, with the initial condition $u(x,y,0) = f(x,y)$, and the boundary condition $u(x,y,t) = 0$, $x^2 + y^2 = l^2$, $t > 0$. Use separation of variables and see Exercise 4.4.32.

4.4.34. Express the Laplacian in cylindrical coordinates (r,θ,z) and solve the Dirichlet problem for $\nabla^2 u(r,\theta,z) = 0$ in the finite cylinder $0 \le r < R$, $0 < \theta < 2\pi$, $0 < z < l$ with $u(R,\theta,z) = f(z)$ and $u(r,\theta,0) = u(r,\theta,l) = 0$, by looking for a solution in the form $u = u(r,z)$. *Hint*: The solution must be bounded at $r = 0$ and involves the modified Bessel function $I_0(x)$.

4.4.35. Solve the problem of Exercise 4.4.34 if the boundary conditions are replaced by $u(r,\theta,0) = 0$, $u(r,\theta,l) = g(r)$ and $u(R,\theta,z) = 0$, by looking for a solution in the form $u = u(r,z)$.

4.4.36. Express $\nabla^2 u = 0$ in spherical coordinates (r,θ,ϕ) with $r \geq 0$, $0 \leq \theta \leq 2\pi$, and $0 \leq \phi \leq \pi$. Consider the boundary value problem for $\nabla^2 u(r,\theta,\phi) = 0$ within the sphere $r < l$ for the function $u = u(r,\theta,\phi)$ with the boundary condition $u(l,\theta,\phi) = f(\phi)$, $0 \leq \phi \leq \pi$ (i.e., the boundary values are independent of θ). Look for a solution in the form $u = U(r,\phi)$ and use separation of variables to construct an eigenfunction expansion of the solution in terms of Legendre polynomials (see Exercise 4.2.11, Section 4.3, and Example 8.3).

4.5 INHOMOGENEOUS EQUATIONS: DUHAMEL'S PRINCIPLE

The method of separation of variables was applied in the foregoing to obtain solutions of initial and boundary value problems for homogeneous PDEs. Here we construct solutions of *inhomogeneous equations* using a technique known as *Duhamel's principle*, which effectively relates the problem to one involving a homogeneous equation. The method is valid for initial value problems and initial and boundary value problems for hyperbolic and parabolic equations.

We consider equations of the form

$$\begin{cases} \rho(\mathbf{x})u_{tt}(\mathbf{x},t) + L[u(\mathbf{x},t)] = g(\mathbf{x},t), & \text{hyperbolic case,} \\ \rho(\mathbf{x})u_t(\mathbf{x},t) + L[u(\mathbf{x},t)] = g(\mathbf{x},t), & \text{parabolic case,} \end{cases} \tag{4.5.1}$$

with $L[u]$ defined as in (4.1.6) or (4.1.7), $\rho(\mathbf{x}) > 0$, and $g(\mathbf{x},t)$ a given forcing or source term. (While $\mathbf{x}$ is treated as a vector variable, the results are valid for one space dimension as well). If we consider the initial value problem for (4.5.1), we assume $u(\mathbf{x},t)$ satisfies homogeneous initial conditions at $t = 0$. For the initial and boundary value problem for (4.5.1) in a bounded region G, we again assume that homogeneous initial conditions for $u(\mathbf{x},t)$ at $t = 0$ in addition to the homogeneous boundary conditions (4.2.2) or (4.2.3). Thus

$$\begin{cases} u(\mathbf{x},0) = u_t(\mathbf{x},0) = 0, & \text{hyperbolic case,} \\ u(\mathbf{x},0) = 0, & \text{parabolic case.} \end{cases} \tag{4.5.2}$$

Duhamel's principle proceeds as follows. Consider a homogeneous version of (4.5.1), that is,

$$\begin{cases} \rho(\mathbf{x})v_{tt}(\mathbf{x},t) + L[v(\mathbf{x},t)] = 0, & \text{hyperbolic case,} \\ \rho(\mathbf{x})v_t(\mathbf{x},t) + L[v(\mathbf{x},t)] = 0, & \text{parabolic case,} \end{cases} \tag{4.5.3}$$

for the function $v(\mathbf{x}, t)$, which is assumed to satisfy the same boundary conditions (if any are given) as $u(\mathbf{x}, t)$. Let $v(\mathbf{x}, t)$ satisfy the following initial conditions given at $t = \tau$, where $\tau \geq 0$:

$$\begin{cases} v(\mathbf{x}, \tau) = 0, \ v_t(\mathbf{x}, \tau) = g(\mathbf{x}, \tau)/\rho(\mathbf{x}), & \text{hyperbolic case,} \\ v(\mathbf{x}, \tau) = g(\mathbf{x}, \tau)/\rho(\mathbf{x}), & \text{parabolic case,} \end{cases} \tag{4.5.4}$$

with $\rho(\mathbf{x})$ and $g(\mathbf{x}, t)$ given as above.

We assume that the problem above for $v(\mathbf{x}, t)$ can be solved by separation of variables for the initial and boundary value problem or by other means for the initial value problem. The solution depends on the parameter τ (i.e., the initial time), so we write it as $v = v(\mathbf{x}, t; \tau)$. Then Duhamel's principle states that the solution $u(\mathbf{x}, t)$ of the given inhomogeneous problem is

$$u(\mathbf{x}, t) = \int_0^t v(\mathbf{x}, t; \tau) \, d\tau. \tag{4.5.5}$$

[A motivation for the method is obtained by noting that the effect of the term $g(\mathbf{x}, t)$ can be characterized as resulting from a superposition of impulses at times $t = \tau$ over the time span $0 \leq \tau \leq t$.]

To verify that $u(\mathbf{x}, t)$ as given in (4.5.5) is a solution of the problem, we note that $u_t(\mathbf{x}, t) = v(\mathbf{x}, t; t) + \int_0^t v_t(\mathbf{x}, t; \tau) \, d\tau$, $u_{tt}(\mathbf{x}, t) = \partial[v(\mathbf{x}, t; t)]/\partial t + v_t(\mathbf{x}, t; t) + \int_0^t v_{tt}(\mathbf{x}, t; \tau) \, d\tau$, and $L[u(\mathbf{x}, t)] = \int_0^t L[v(\mathbf{x}, t; \tau)] \, d\tau$. In view of (4.5.4), we have in the *parabolic case*,

$$v(\mathbf{x}, t; t) = g(\mathbf{x}, t)/\rho(\mathbf{x}), \tag{4.5.6}$$

and in the *hyperbolic case*,

$$v(\mathbf{x}, t; t) = 0, \qquad v_t(\mathbf{x}, t; t) = g(\mathbf{x}, t)/\rho(\mathbf{x}). \tag{4.5.7}$$

Therefore,

$$\rho u_t(\mathbf{x}, t) + L[u(\mathbf{x}, t)] = g(\mathbf{x}, t) + \int_0^t [\rho v_t(\mathbf{x}, t; \tau) + L(v(\mathbf{x}, t; \tau))] \, d\tau = g(\mathbf{x}, t), \tag{4.5.8}$$

$$\rho u_{tt}(\mathbf{x}, t) + L[u(\mathbf{x}, t)] = g(\mathbf{x}, t) + \int_0^t [\rho v_{tt}(\mathbf{x}, t; \tau) + L(v(\mathbf{x}, t; \tau))] \, d\tau = g(\mathbf{x}, t). \tag{4.5.9}$$

Also, $u(\mathbf{x}, 0) = 0$ and $u_t(\mathbf{x}, 0) = v(\mathbf{x}, 0; 0) = 0$, so that $u(\mathbf{x}, t)$ as defined by (4.5.5) satisfies all the conditions of the problem.

Examples

Next we consider two examples where we apply Duhamel's principle to an initial value problem for the inhomogeneous wave equation and an initial and boundary value problem for the inhomogeneous heat or diffusion equation.

Example 4.5. The Inhomogeneous One-Dimensional Wave Equation. We consider the wave equation with a given forcing term $g(x,t)$,

$$u_{tt}(x,t) - c^2 u_{xx}(x,t) = g(x,t), \qquad -\infty < x < \infty,\ t > 0, \tag{4.5.10}$$

and the homogeneous initial conditions

$$u(x,0) = u_t(x,0) = 0, \qquad -\infty < x < \infty. \tag{4.5.11}$$

[We can add a solution of the homogeneous wave equation with arbitrary data to $u(x,t)$, enabling us to solve the general initial value problem for (4.5.10).]

Applying *Duhamel's principle*, we consider the function $v(x,t;\tau)$ that satisfies the equation

$$v_{tt}(x,t;\tau) - c^2 v_{xx}(x,t;\tau) = 0, \qquad -\infty < x < \infty,\ t > \tau, \tag{4.5.12}$$

and the initial conditions at $t = \tau$,

$$v(x,\tau;\tau) = 0, \qquad v_t(x,\tau;\tau) = g(x,\tau), \qquad -\infty < x < \infty. \tag{4.5.13}$$

From *d'Alembert's solution* (see Example 2.4) we easily obtain

$$v(x,t;\tau) = \frac{1}{2c}\int_{x-c(t-\tau)}^{x+c(t-\tau)} g(\sigma,\tau)\, d\sigma, \tag{4.5.14}$$

and from (4.5.5), $u(x,t)$ takes the form

$$u(x,t) = \frac{1}{2c}\int_0^t \int_{x-c(t-\tau)}^{x+c(t-\tau)} g(\sigma,\tau)\, d\sigma\, d\tau. \tag{4.5.15}$$

If we replace the initial data (4.5.11) by the arbitrary data

$$u(x,0) = F(x), \qquad u_t(x,0) = G(x), \tag{4.5.16}$$

the solution of the initial value problem (4.5.10) and (4.5.16) is

$$u(x,t) = \frac{1}{2}[F(x+ct) + F(x-ct)] + \frac{1}{2c}\int_{x-ct}^{x+ct} G(\sigma)\, d\sigma$$
$$+ \frac{1}{2c}\int_0^t \int_{x-c(t-\tau)}^{x+c(t-\tau)} g(\sigma,\tau)\, d\sigma\, d\tau. \tag{4.5.17}$$

At the arbitrary point (x_0, t_0) with $t_0 > 0$, $u(x_0,t_0)$ depends only on values x and t within the *characteristic triangle* pictured in Figure 4.6. This follows on inspection of the arguments of $F(x_0 \pm ct_0)$ and the domains of integration for $G(\sigma)$ and $g(\sigma,\tau)$ in (4.5.17). As a result, the characteristic triangle is called the *domain of dependence* of the solution at the point x_0 and at the time t_0. Similarly, the sector pictured in

Figure 4.7 that is bounded by the characteristic lines issuing from the point $(x_0, 0)$ is called the *domain* or *region of influence* of the initial point $(x_0, 0)$. All points (x, t) in this sector have the point $(x_0, 0)$ within their domain of dependence. In particular, if $g(x, t) = 0$ and $F(x)$ and $G(x)$ are concentrated at the point x_0, we see that the solution $u(x, t)$ vanishes identically outside the domain of influence of the point x_0. Within the sector, $u(x, t)$ may or may not be zero. The two characteristic lines $x \pm ct = x_0 \pm ct_0$ represent wave fronts for the solution, since $u(x, t) = 0$ ahead of the wave fronts (which move right and left) and $u(x, t) \neq 0$, in general, at points (x, t) that lie within the sector behind the wave fronts.

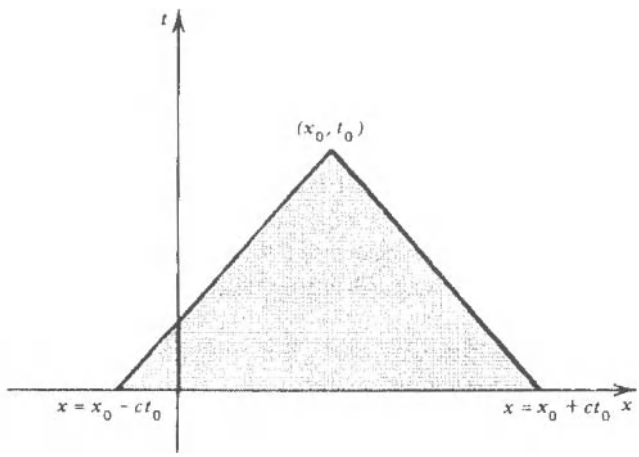

Figure 4.6 Domain of dependence.

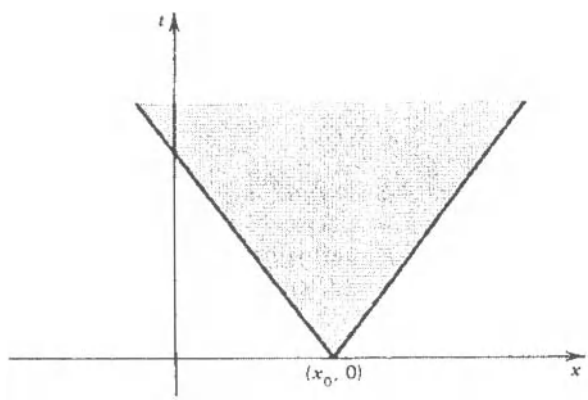

Figure 4.7 Domain of influence.

Example 4.6. The Inhomogeneous One-Dimensional Heat Equation. The heat equation in a finite interval $0 < x < l$, with a heat source, has the form

$$u_t(x,t) - c^2 u_{xx}(x,t) = g(x,t), \qquad 0 < x < l,\ t > 0. \tag{4.5.18}$$

We assume a homogeneous initial condition $u(x,0) = 0$ and the boundary conditions $u(0,t) = u(l,t) = 0$ at $x = 0$ and $x = l$.

On applying *Duhamel's principle* and using the solution obtained in (4.4.29) for the homogeneous heat equation with t replaced by $t - \tau$, we obtain the result

$$v(x,t;\tau) = \sqrt{\frac{2}{l}} \sum_{k=1}^{\infty} a_k(\tau) \exp\left[-\left(\frac{\pi k c}{l}\right)^2 (t-\tau)\right] \sin\left(\frac{\pi k}{l} x\right), \tag{4.5.19}$$

with the $a_k(\tau)$ determined from the initial condition

$$v(x,\tau;\tau) = g(x,\tau) = \sqrt{\frac{2}{l}} \sum_{k=1}^{\infty} a_k(\tau) \sin\left(\frac{\pi k}{l} x\right). \tag{4.5.20}$$

Note that the Fourier coefficients a_k are functions of the parameter τ [i.e., for each value of τ we obtain a Fourier sine series of $g(x,\tau)$]. As a result, the solution of the initial and boundary value problem for (4.5.18) is given in the form

$$u(x,t) = \sqrt{\frac{2}{l}} \sum_{k=1}^{\infty} \left\{ \int_0^t a_k(\tau) \exp\left[-\left(\frac{\pi k c}{l}\right)^2 (t-\tau)\right] d\tau \right\} \sin\left(\frac{\pi k}{l} x\right), \tag{4.5.21}$$

where summation and integration have been interchanged. This is a valid procedure since the series converges uniformly, say, if we assume that $g(x,t)$ is a bounded function. If $u(x,0) = f(x)$ and is nonzero, we may add the solution (4.4.29) to (4.5.21). Thereby we obtain the solution of the modified problem.

Next, we specialize the foregoing result to a specific problem. We suppose that $g(x,t) = \sin(\pi x/l)$. Then it can be seen by inspection that $a_1 = \sqrt{l/2}$ and $a_k = 0,\ k > 1$. Consequently, the solution (4.5.21) takes on the simple form

$$u(x,t) = \left(\frac{l}{\pi c}\right)^2 \left\{1 - \exp\left[-\left(\frac{\pi c}{l}\right)^2 t\right]\right\} \sin\left(\frac{\pi}{l} x\right). \tag{4.5.22}$$

Note that we are again assuming that the initial condition is given as $u(x,0) = 0$.

As $t \to \infty$ we see that

$$\lim_{t\to\infty} u(x,t) = \left(\frac{l}{\pi c}\right)^2 \sin\left(\frac{\pi}{l} x\right) \equiv v(x). \tag{4.5.23}$$

The limiting function $v(x)$ in (4.5.23) is a solution of the *steady-state problem* for the heat equation with $g(x,t) = \sin(\pi x/l)$. It is given as

$$-c^2 v''(x) = \sin\left(\frac{\pi}{l} x\right), \qquad v(0) = v(l) = 0. \tag{4.5.24}$$

This problem results when we assume that $u(x,t) = v(x)$ [i.e., $u(x,t)$ is independent of t] and the initial condition is dropped.

More generally, if the inhomogeneous term $g(x,t)$ is independent of time, that is, $g(x,t) = G(x)$, with $G(x)$ arbitrary, we obtain from (4.5.21)

$$u(x,t) = \sqrt{\frac{2}{l}} \sum_{k=1}^{\infty} \left\{ \left(\frac{l}{\pi k c} \right)^2 a_k \left(1 - \exp\left[-\left(\frac{\pi k c}{l} \right)^2 t \right] \right) \sin\left(\frac{\pi k}{l} x \right) \right\}, \tag{4.5.25}$$

where we have used the fact that a_k is independent of τ to carry out the τ integration. As $t \to \infty$ in (4.5.25) we have

$$\lim_{t \to \infty} u(x,t) = \sqrt{\frac{2}{l}} \sum_{k=1}^{\infty} \left(\frac{l}{\pi k c} \right)^2 a_k \sin\left(\frac{\pi k}{l} x \right) \equiv v(x). \tag{4.5.26}$$

We easily check that $v(x)$ solves the steady-state problem for the heat equation

$$-c^2 v''(x) = G(x), \qquad v(0) = v(l) = 0. \tag{4.5.27}$$

Now even if $u(x,0) = f(x) \neq 0$, we have already seen in Example 4.10 that the effect of the nonzero initial heat distribution dies out as $t \to \infty$, so that when the heat source term $g(x,t) = G(x)$ is independent of time, the temperature distribution (when the ends of the rod are fixed at zero temperature) is well approximated by the solution of (4.5.27) for large t.

This completes our discussion of Duhamel's principle. The method of finite Fourier transforms and other techniques for handling inhomogeneous problems are discussed in the following section.

Exercises 4.5

4.5.1. Solve the initial value problem $u_{tt}(x,t) - c^2 u_{xx}(x,t) = g(x,t)$, $-\infty < x < \infty$, $t > 0$, $u(x,0) = u_t(x,0) = 0$ if $g(x,t)$ has the following form: (a) $g(x,t) = e^{-t}\sin(x)$; (b) $g(x,t) = t$; (c) $g(x,t) = \delta(x)\cos(t)$; (d) $g(x,t) = \delta(x-a)\delta(t-b)$, a, b = constant, $b > 0$. In parts (c) and (d), $\delta(z)$ is the Dirac delta function.

4.5.2. Use Duhamel's principle to solve $u_{tt}(x,t) - c^2 u_{xx}(x,t) = g(x,t)$, $0 < x < l$, $t > 0$, $u(x,0) = u_t(x,0) = 0$, $u(0,t) = u(l,t) = 0$.

4.5.3. Solve the problem in Exercise 4.5.2 if $g(x,t)$ assumes the following forms: (a) $g(x,t) = x(x-l)$; (b) $g(x,t) = \sin(\omega x)\sin(\omega c t)$, $\omega \neq \pi k/l$; (c) $g(x,t) = \sin(\pi x/l)\sin(\pi c t/l)$.

4.5.4. Apply Duhamel's principle to obtain the solution of the initial and boundary value problem $u_t(x,t) - c^2 u_{xx}(x,t) = g(x,t), \quad 0 < x < l, \ t > 0, \quad u(x,0) = 0, \ u(0,t) = u(l,t) = 0.$

4.5.5. Consider the special case in which $g = g(x)$ (i.e., it is independent of t) in Exercise 4.5.4 and examine the limit of the solution $u(x,t)$ as $t \to \infty$. What condition must $g(x)$ satisfy so that this limit exists?

4.5.6. Let $g(x,t) = \cos(\pi x/l)$ in Exercise 4.5.4 and obtain the solution of the initial and boundary value problem in that case.

4.5.7. Use Duhamel's principle to solve the following problem for the inhomogeneous telegrapher's equation $u_{tt}(x,t) - c^2 u_{xx}(x,t) + 2\hat{\lambda} u_t(x,t) = g(x)$, $0 < x < l$, $t > 0$, where $\hat{\lambda} > 0$, with the initial and boundary conditions $u(x,0) = 0$, $u_t(x,0) = 0$, $u(0,t) = u(l,t) = 0$. Find the limit of the solution as $t \to \infty$ and show that it satisfies an appropriate steady-state problem. Verify this result if $g(x) = \sin(\pi x/l)$.

4.5.8. Use Duhamel's principle to obtain the solution of the inhomogeneous Klein-Gordon equation $u_{tt}(x,t) - u_{xx}(x,t) + c^2 u(x,t) = g(x,t)$, $0 < x < l$, $t > 0$, with the initial and boundary conditions $u(x,0) = 0$, $u_t(x,0) = 0$, $u(0,t) = u(l,t) = 0$. Use the results of Exercise 4.4.16.

4.6 EIGENFUNCTION EXPANSIONS: FINITE FOURIER TRANSFORMS

The method of separation of variables presented in Section 4.2 led to the representation of the solution of the given problem in a series of eigenfunctions of the differential operator L defined in (4.1.6)–(4.1.7). To generate the appropriate eigenvalue problem, it was necessary that both the equation and the boundary conditions be homogeneous. For the self-adjoint operators $(1/\rho)L$ under consideration, the eigenfunctions have the property of completeness, so that under certain conditions a function $f(x)$ can be expanded in a (convergent) series of eigenfunctions. The orthogonality of the eigenfunctions renders the determination of the coefficients in the series to be quite simple.

In this section we reverse the foregoing approach and begin by directly expressing the solution of the problem as a *series of eigenfunctions*. The coefficients in the series are arbitrary and must be determined from the differential equation and the data for the problem. Rather than substituting the series directly into the equation, we convert the PDE into a hierarchy of ODEs for the determination of the unknown coefficients, termed *Fourier coefficients*. The equation is thereby transformed into a collection of equations for these Fourier coefficients, and this technique is often called the *method of finite Fourier transforms*. More specifically, if the method involves the use of Fourier sine, Fourier cosine, or Bessel function expansions of the form considered in Section 4.3, it is known as the *method of finite sine, cosine*, or *Hankel transforms*, respectively. The reason for this will become clearer when we discuss (infinite) Fourier transforms in Chapter 5. This method applies to inhomogeneous equations

with inhomogeneous initial and boundary data. Technically, therefore, it enables us to solve the most general problems formulated in Section 4.1.

PDEs with General Inhomogeneous Terms and Data

We consider PDEs of the form

$$\rho(\mathbf{x})Ku + Lu = \rho(\mathbf{x})F, \tag{4.6.1}$$

where L is defined as in (4.1.6) or (4.1.7), and $\rho(\mathbf{x}) > 0$ and F are given functions. The differential operator K can be given as $K = \partial/\partial t$, $K = \partial^2/\partial t^2$, or $K = -\partial^2/\partial y^2$, so that (4.6.1) has the form of (4.1.8), (4.1.10), or (4.2.7), respectively. Or $K = 0$ and (4.6.1) reduces to (4.1.9). In fact, L may have the form associated with any of the eigenvalue problems considered in Section 4.3. All that is required of K is that it does not depend on any of the variables that occur explicitly in the eigenvalue problem associated with L. Equations of parabolic, hyperbolic, or elliptic type can be treated by this method. The inhomogeneous boundary conditions for the problem are as given in (4.1.11) or (4.1.12) and initial conditions are assigned in the hyperbolic and parabolic cases as in (4.2.4) or (4.2.6). Conditions of the form (4.2.8) for the elliptic case can also be given.

The solution u of each of the foregoing problems is expanded in a series of eigenfunctions $M_k(\mathbf{x})$ determined from the eigenvalue problem

$$LM_k(\mathbf{x}) = \lambda_k \rho(\mathbf{x}) M_k(\mathbf{x}), \qquad k = 1, 2, \ldots \tag{4.6.2}$$

with homogeneous boundary conditions of the form (4.2.2) or (4.2.3). For this discussion it is assumed that the $M_k(\mathbf{x})$ form a complete orthonormal set. Thus $(M_k(\mathbf{x}),\ M_j(\mathbf{x})) = \delta_{kj}$, the Kronecker delta. We expand u as

$$u = \sum_{k=1}^{\infty} N_k M_k(\mathbf{x}), \tag{4.6.3}$$

where the Fourier coefficients N_k (which may depend on t or on y) are to be specified. In terms of the inner product (4.2.21), the N_k are formally given as

$$N_k = (u, M_k(\mathbf{x})), \tag{4.6.4}$$

since the $M_k(\mathbf{x})$ are orthonormalized, but the solution u is not yet known. To specify the N_k we do not substitute (4.6.3) directly into the equation (4.6.1) since the series may not be differentiable term by term. Instead, we transform (4.6.1) into a system of equations for the Fourier coefficients N_k. For each k, N_k is a *finite Fourier transform* of u. Or, more precisely, the full set of the N_k represents the finite Fourier transform of u. Once the full set of the N_k is known, u can be reproduced from the series (4.6.3). In this sense (4.6.3) inverts the transforms N_k and the series representation of u is called the *inverse* (finite) *Fourier transform* of the N_k.

To proceed, we multiply (4.6.1) by $M_k(\mathbf{x})$ and integrate over the underlying region G to obtain

$$\iint_G \rho(\mathbf{x}) M_k(\mathbf{x}) K u \, dv = K \iint_G \rho(\mathbf{x}) M_k(\mathbf{x}) u \, dv = K(u, M_k(\mathbf{x}))$$

$$= -\iint_G M_k(\mathbf{x}) L u \, dv + \iint_G \rho(\mathbf{x}) F M_k(\mathbf{x}) \, dv, \tag{4.6.5}$$

where we have pulled the differential operator K out of the integral. Now from (4.2.19) we have

$$\iint_G M_k(\mathbf{x}) L u \, dv = \iint_G u L M_k(\mathbf{x}) \, dv - \int_{\partial G} p(\mathbf{x}) \left[M_k(\mathbf{x}) \frac{\partial u}{\partial n} - u \frac{\partial M_k(\mathbf{x})}{\partial n} \right] ds. \tag{4.6.6}$$

Then, using (4.6.2) and the boundary condition satisfied by u and $M_k(\mathbf{x})$ in (4.6.6) we readily obtain

$$-\iint_G M_k(\mathbf{x}) L u \, dv = -\lambda_k(u, M_k(\mathbf{x}))$$

$$- \int_{S_1} \frac{p(\mathbf{x})}{\alpha(\mathbf{x})} \frac{\partial M_k(\mathbf{x})}{\partial n} B \, ds + \int_{S_2 \cup S_3} \frac{p(\mathbf{x})}{\beta(\mathbf{x})} M_k(\mathbf{x}) B \, ds. \tag{4.6.7}$$

We denote the boundary integral terms in (4.6.7) by B_k. (We assume in our discussion that zero is not an eigenvalue for any of the problems we consider.)

Combining (4.6.7) and (4.6.5), we obtain the *hierarchy of ordinary differential equations* for N_k,

$$K N_k + \lambda_k N_k = F_k + B_k, \qquad k = 1, 2, \ldots, \tag{4.6.8}$$

where F_k is the Fourier coefficient of F [i.e., $F_k = (F, M_k(\mathbf{x}))$].

In the *elliptic case* that corresponds to (4.1.9) and for which $K = 0$ in (4.6.1), the equations (4.6.8) are, in fact, algebraic. Since $\lambda_k > 0$ for all k by assumption, the N_k are uniquely determined. For the *elliptic case* that corresponds to (4.2.7) with $K = -\partial^2/\partial y^2$, (4.6.8) is a second order ordinary differential equation in y and the $N_k(y)$ must satisfy the boundary conditions obtained by setting $u(x, y) = f(x)$ and $u(x, y) = g(x)$ in (4.6.4) at $y = 0$ and $y = \hat{l}$, respectively, in view of (4.2.8).

For the *parabolic case*, $K = \partial/\partial t$ and we have for $N_k(t)$,

$$\frac{dN_k(t)}{dt} + \lambda_k N_k(t) = F_k(t) + B_k(t), \; N_k(0) = (f(\mathbf{x}), M_k(\mathbf{x})). \tag{4.6.9}$$

The initial condition is obtained from the eigenfunction expansion of $u(\mathbf{x}, 0) = f(\mathbf{x})$. The solution of this problem is

$$N_k(t) = (f(\mathbf{x}), M_k(\mathbf{x})) \exp(-\lambda_k t) + \int_0^t [F_k(\tau) + B_k(\tau)] \exp[-\lambda_k(t - \tau)] \, d\tau. \tag{4.6.10}$$

In the *hyperbolic case*, with $K = \partial^2/\partial t^2$ we obtain the initial conditions from

$$u(\mathbf{x},0) = f(\mathbf{x}) = \sum_{k=1}^{\infty} N_k(0)M_k(\mathbf{x}),\ u_t(\mathbf{x},0) = g(\mathbf{x}) = \sum_{k=1}^{\infty} N_k'(0)M_k(\mathbf{x}), \tag{4.6.11}$$

so that at $t = 0$ we have for $N_k(t)$,

$$N_k(0) = (f(\mathbf{x}), M_k(\mathbf{x})), \qquad N_k'(0) = (g(\mathbf{x}), M_k(\mathbf{x})). \tag{4.6.12}$$

The solution of the equation

$$\frac{d^2N_k(t)}{dt^2} + \lambda_k N_k(t) = F_k(t) + B_k(t), \tag{4.6.13}$$

which satisfies (4.6.12), is

$$N_k(t) = (f(\mathbf{x}), M_k(\mathbf{x}))\cos(\sqrt{\lambda_k}\,t) + \frac{1}{\sqrt{\lambda_k}}(g(\mathbf{x}), M_k(\mathbf{x}))\sin(\sqrt{\lambda_k}\,t)$$
$$+\frac{1}{\sqrt{\lambda_k}}\int_0^t [F_k(\tau) + B_k(\tau)]\sin[\sqrt{\lambda_k}(t-\tau)]\,d\tau. \tag{4.6.14}$$

The foregoing results for the parabolic and hyperbolic cases reduce to those obtained in Section 4.2 by means of separation of variables if $F_k = B_k = 0$.

The solutions u obtained as a series of eigenfunctions cannot, in general, be expected to be *classical solutions* that satisfy the differential equation (4.6.1) and assume the initial and boundary values pointwise. In fact, each of the eigenfunctions $M_k(\mathbf{x})$ satisfies a homogeneous boundary condition, whereas the solution u given as a sum of these eigenfunctions may satisfy an inhomogeneous boundary condition. Nevertheless, these boundary conditions are certainly accounted for in the equations (4.6.8) for the N_k. Thus convergence up to the boundary may have to be interpreted in a *generalized sense*, perhaps using mean square convergence. Additionally, although there may be pointwise convergence in interior regions, the lack of convergence up to the boundary generally slows the rate of convergence for the series everywhere. Consequently, we present a method (following Example 4.8) that often enables us to circumvent the foregoing difficulty by converting the problem to one with homogeneous boundary conditions.

Examples

Example 4.7. Hyperbolic Equations: Resonance. We consider the inhomogeneous hyperbolic equation

$$\rho(\mathbf{x})u_{tt}(\mathbf{x},t) + Lu(\mathbf{x},t) = \rho(\mathbf{x})M_i(\mathbf{x})\sin(\omega t), \tag{4.6.15}$$

where $M_i(\mathbf{x})$ is one of the eigenfunctions determined from (4.6.2) and $F(\mathbf{x},t) = M_i(\mathbf{x})\sin(\omega t)$ is a periodic forcing term in t with frequency of vibration ω = constant. The initial conditions as well as the boundary conditions are assumed to be homogeneous; that is, $f(\mathbf{x}) = g(\mathbf{x}) = B(\mathbf{x},t) = 0$.

Now $F_k(t) = (F(\mathbf{x},t), M_k(\mathbf{x})) = \sin(\omega t)(M_i(\mathbf{x}), M_k(\mathbf{x})) = \delta_{ik}\sin(\omega t)$, where the Kronecker delta $\delta_{ik} = 0$ for $i \neq k$ and $\delta_{ii} = 1$ for all i. Since $(f(\mathbf{x}), M_k(\mathbf{x})) = (g(\mathbf{x}), M_k(\mathbf{x})) = B_k(t) = 0$ for all k in this problem, we have, from (4.6.14),

$$N_i(t) = \frac{1}{\sqrt{\lambda_i}}\int_0^t \sin(\omega t)\sin[\sqrt{\lambda_i}(t-\tau)]\,d\tau = \frac{\omega\sqrt{\lambda_i}\sin(\sqrt{\lambda_i}\,t) - \sin(\omega t)}{\sqrt{\lambda_i}(\omega^2 - \lambda_i)} \tag{4.6.16}$$

and $N_k(t) = 0$ for $k \neq i$. Thus the series (4.6.3) reduces to the single term

$$u(\mathbf{x},t) = \frac{1}{\omega^2 - \lambda_i}\left[\frac{\omega}{\sqrt{\lambda_i}}\sin(\sqrt{\lambda_i}\,t) - \sin(\omega t)\right] M_i(\mathbf{x}), \tag{4.6.17}$$

which is, in fact, the exact solution of the problem.

Clearly, (4.6.17) is not valid if $\omega = \sqrt{\lambda_i}$. We may obtain the solution in that case by going to the limit as $\omega \to \sqrt{\lambda_i}$ in (4.6.17). l'Hospital's rule gives

$$u(\mathbf{x},t) = \frac{1}{2\sqrt{\lambda_i}}\left[\frac{\sin(\sqrt{\lambda_i}\,t)}{\sqrt{\lambda_i}} - t\cos(\sqrt{\lambda_i}\,t)\right] M_i(\mathbf{x}). \tag{4.6.18}$$

To interpret these results, we observe that the numbers $\omega_k = \sqrt{\lambda_k}$ $(k = 1, 2, \ldots)$ represent the *natural frequencies of vibration* of the solution of the homogeneous initial and boundary value problem for (4.6.15) [i.e., with $F(\mathbf{x},t) = 0$] in view of (4.2.31). Thus when $\omega \neq \omega_i$ (one of the natural frequencies), the solution (4.6.17) oscillates with the imposed external frequency ω, as well as the natural frequency $\omega_i = \sqrt{\lambda_i}$. The external energy fed into the system is distributed between these two frequencies. However, if $\omega = \omega_i = \sqrt{\lambda_i}$, we find that the amplitude of the oscillatory solution increases unboundedly with t as $t \to \infty$. The effect is known as *resonance*. It results because the entire external energy is concentrated within a single natural frequency of vibration, and the continuous input builds up its amplitude.

Example 4.8. Poisson's Equation in a Circle. We consider the boundary value problem for *Poisson's equation*,

$$\nabla^2 u(x,y) = -F(x,y), \tag{4.6.19}$$

within a disk of radius R with center at the origin and an inhomogeneous boundary condition assigned on the boundary circle. Expressing the problem in *polar coordinates* (r,θ), we obtain for $u(r,\theta)$ the equation

$$\frac{\partial^2 u(r,\theta)}{\partial r^2} + \frac{1}{r}\frac{\partial u(r,\theta)}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u(r,\theta)}{\partial \theta^2} = -F(r,\theta), \tag{4.6.20}$$

within the circle (i.e., for $r < R$) and the boundary condition

$$u(R,\theta) = f(\theta). \tag{4.6.21}$$

In view of our discussion following equation (4.6.1), this problem can be solved in terms of the eigenfunctions $M_k(r,\theta)$ determined from $\nabla^2 M_k(r,\theta) = -\lambda_k M_k(r,\theta)$ for $r < R$ with $M_k(R,\theta) = 0$. Since this higher-dimensional eigenvalue problem is not considered until Chapter 8, we present an alternative approach to the given boundary value problem that involves an expansion in terms of one-dimensional eigenfunctions.

The solution of (4.6.20)–(4.6.21) is expected to be single valued at any point within the disk, so we must have $u(r,\theta+2\pi) = u(r,\theta)$ with a similar result for $F(r,\theta)$ and $f(\theta)$. That is, $u(r,\theta)$, $F(r,\theta)$, and $f(\theta)$ are periodic of period 2π in θ. This suggests that the eigenvalue problem associated with $\hat{L}u = -\partial^2 u/\partial\theta^2$ with periodic boundary conditions of period 2π is appropriate here. Accordingly, we write (4.6.20) in the form

$$Ku(r,\theta)+\hat{L}u(r,\theta) \equiv -r^2 u_{rr}(r,\theta) - r u_r(r,\theta) - u_{\theta\theta}(r,\theta) = r^2 F(r,\theta), \quad (4.6.22)$$

and expand the solution in a series of eigenfunctions for the operator $\hat{L}$. The foregoing eigenvalue problem was discussed in Example 4.6, where with $l = \pi$, we obtained the eigenfunctions $\{\cos(k\theta)\}$, $k = 0,1,2,\ldots$ and $\{\sin(k\theta)\}$, $k = 1,2,\ldots$, and the eigenvalues $\lambda_k = k^2$, $k = 0,1,2,\ldots$. Using the orthonormal set of eigenfunctions (4.3.48)–(4.3.49) we obtain the *eigenfunction expansion*

$$u(r,\theta) = \frac{1}{\sqrt{2\pi}} a_0(r) + \sum_{k=1}^{\infty} \frac{1}{\sqrt{\pi}} \{a_k(r)\cos(k\theta) + b_k(r)\sin(k\theta)\} \qquad (4.6.23),$$

which corresponds to (4.6.3). Even though this eigenvalue problem is not of Sturm-Liouville type, the finite Fourier transform method can be applied.

To begin, we note that the Fourier coefficients $a_0(r)$, $a_k(r)$ and $b_k(r)$ are given in terms of $u(r,\theta)$ as

$$a_0(r) = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} u(r,\theta)\, d\theta, \; a_k(r) = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} u(r,\theta)\cos(k\theta)\, d\theta, \; k = 1,2,\ldots. \quad (4.6.24)$$

$$b_k(r) = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} u(r,\theta)\sin(k\theta)\, d\theta, \qquad k = 1,2,\ldots. \quad (4.6.25)$$

The equations for the Fourier coefficients $a_0(r)$, $a_k(r)$, and $b_k(r)$ are obtained from (4.6.20) on multiplying across by the eigenfunctions and integrating from $-\pi$ to π as

$$\frac{d^2 a_k(r)}{dr^2} + \frac{1}{r}\frac{da_k(r)}{dr} - \frac{k^2}{r^2} a_k(r) = -A_k(r), \qquad k = 0,1,2,\ldots, \quad (4.6.26)$$

$$\frac{d^2 b_k(r)}{dr^2} + \frac{1}{r}\frac{db_k(r)}{dr} - \frac{k^2}{r^2} b_k(r) = -B_k(r), \qquad k = 1,2,\ldots, \quad (4.6.27)$$

where $A_k(r)$ and $B_k(r)$ are the Fourier coefficients of $F(r,\theta)$. Since $r = 0$ is a singular point in the equations for $a_k(r)$ and $b_k(r)$ and we require the solution $u(r,\theta)$

to be bounded at $r = 0$, we obtain the boundary conditions $a_0(r)$, $a_k(r)$, $b_k(r)$ bounded at $r = 0$, $k = 1, 2, \ldots$. At $r = R$, (4.6.23) must represent the Fourier series expansion of $f(\theta)$. Thus, the boundary condition (4.6.21) implies that $a_0(R) = \hat{f}_0$, $a_k(R) = \hat{f}_k$, $b_k(R) = f_k$, $k \geq 1$, where $\hat{f}_0$, $\hat{f}_k$, and f_k are the Fourier coefficients of $f(\theta)$.

Noting that (4.6.26)–(4.6.27) are both inhomogeneous forms of *Euler's equation*

$$C_k''(r) + \frac{1}{r}C_k'(r) - \frac{k^2}{r^2}C_k(r) = 0, \qquad k = 0, 1, 2, \ldots, \tag{4.6.28}$$

for which a fundamental set of solutions is

$$C_0(r) = \begin{cases} \text{constant}, \\ \log(r), \end{cases} \qquad C_k(r) = \begin{cases} r^k, \\ r^{-k}, \end{cases} \qquad k \geq 1, \tag{4.6.29}$$

we obtain, on using *variation of parameters*, the following solutions of the foregoing boundary value problems:

$$a_0(r) = \int_0^r \log\left(\frac{R}{r}\right) A_0(t)t\, dt + \int_r^R \log\left(\frac{R}{t}\right) A_0(t)t\, dt + \hat{f}_0, \tag{4.6.30}$$

$$a_k(r) = \frac{1}{2k}\int_0^r \left[\left(\frac{R}{r}\right)^k - \left(\frac{r}{R}\right)^k\right]\left(\frac{t}{R}\right)^k A_k(t)t\, dt$$

$$+\frac{1}{2k}\int_r^R \left[\left(\frac{R}{t}\right)^k - \left(\frac{t}{R}\right)^k\right]\left(\frac{r}{R}\right)^k A_k(t)t\, dt + \left(\frac{r}{R}\right)^k \hat{f}_k, \qquad k \geq 1, \tag{4.6.31}$$

with the $b_k(r)$ identical in form to the $a_k(r)$ $(k \geq 1)$ except that $A_k(t)$ and $\hat{f}_k$ are replaced by $B_k(t)$ and f_k, respectively.

Thus the Fourier coefficients in the series (4.6.23) for $u(r, \theta)$ are specified completely. Under suitable conditions on $F(r, \theta)$, say $\int_0^R \int_{-\pi}^{\pi} F^2(r, \theta) r\, dr\, d\theta < \infty$, with $f(\theta)$ continuous and piecewise smooth, the Fourier series (4.6.23) converges uniformly. Then $u(r, \theta)$ is continuous and $u(R, \theta) = f(\theta)$.

Time-Dependent PDEs with Stationary Inhomogeneities

An interesting special case for the parabolic and hyperbolic problems discussed arises when the functions $F(\mathbf{x}, t)$ and $B(\mathbf{x}, t)$ are independent of t. Since these terms cause the problems to be inhomogeneous, the inhomogeneities are said to be *stationary* (i.e., time-independent). This assumption has the effect that the terms F_k and B_k are time-independent constants. As a result, the integrals in (4.6.10) and (4.6.14) can be evaluated, and we obtain for $N_k(t)$,

$$N_k(t) = \left[(f, M_k) - \frac{F_k + B_k}{\lambda_k}\right] e^{-\lambda_k t} + \frac{F_k + B_k}{\lambda_k}, \tag{4.6.32}$$

$$N_k(t) = \left[(f, M_k) - \frac{F_k + B_k}{\lambda_k}\right] \cos(\sqrt{\lambda_k}\, t) + \frac{1}{\sqrt{\lambda_k}}(g, M_k) \sin(\sqrt{\lambda_k}\, t) + \frac{F_k + B_k}{\lambda_k}, \tag{4.6.33}$$

in the *parabolic* and *hyperbolic cases*, respectively.

In both cases we can decompose $N_k(t)$ into the form $N_k(t) = \tilde{N}_k(t) + \hat{N}_k$, where $\hat{N}_k = (1/\lambda_k)(F_k + B_k)$ is independent of t. Correspondingly, the solution $u(\mathbf{x}, t)$ can be formally expressed as $u(\mathbf{x}, t) = \sum_{k=1}^{\infty} \tilde{N}_k(t) M_k(\mathbf{x}) + \sum_{k=1}^{\infty} \hat{N}_k M_k(\mathbf{x})$. If we set $F(\mathbf{x}, t) = \hat{F}(\mathbf{x})$ and $B(\mathbf{x}, t) = \hat{B}(\mathbf{x})$ since F and B are independent of t and define

$$v(\mathbf{x}) = \sum_{k=1}^{\infty} \hat{N}_k M_k(\mathbf{x}), \tag{4.6.34}$$

it is readily seen that $v(\mathbf{x})$ is the solution of the boundary value problem for the stationary version of (4.6.1),

$$Lv(\mathbf{x}) = \rho(\mathbf{x})\hat{F}(\mathbf{x}), \qquad \alpha(\mathbf{x})v(\mathbf{x}) + \beta(\mathbf{x})\frac{\partial v(\mathbf{x})}{\partial n}\bigg|_{\partial G} = \hat{B}(\mathbf{x}). \tag{4.6.35}$$

[On solving (4.6.35) by the finite transform method, we obtain (4.6.34).]

Then if we put

$$u(\mathbf{x}, t) = w(\mathbf{x}, t) + v(\mathbf{x}), \tag{4.6.36}$$

with $v(\mathbf{x})$ defined as above, $w(\mathbf{x}, t)$ satisfies the homogeneous equation

$$\rho(\mathbf{x}) K w(\mathbf{x}, t) + Lw(\mathbf{x}, t) = 0, \tag{4.6.37}$$

with the boundary condition

$$\alpha(\mathbf{x})w(\mathbf{x}, t) + \beta(\mathbf{x})\frac{\partial w(\mathbf{x}, t)}{\partial n}\bigg|_{\partial G} = 0, \tag{4.6.38}$$

where $K = \partial/\partial t$ and $K = \partial^2/\partial t^2$ in the parabolic and hyperbolic cases, respectively. The respective initial conditions are

$$w(\mathbf{x}, 0) = f(\mathbf{x}) - v(\mathbf{x}), \qquad w(\mathbf{x}, 0) = f(\mathbf{x}) - v(\mathbf{x}),\ w_t(\mathbf{x}, 0) = g(\mathbf{x}). \tag{4.6.39}$$

Once the possibility of the decomposition (4.6.36) has been recognized, it may be introduced directly when solving initial and boundary value problems with *stationary inhomogeneities*. Since the equation for $w(\mathbf{x}, t)$ and the boundary conditions are homogeneous, the rate of convergence of the formal series solution of the problem for $w(\mathbf{x}, t)$ is accelerated. Additionally, if the problem involves only one space dimension, we obtain an ordinary differential equation for $v(x)$. Then if $v(x)$ can be determined without the use of an eigenfunction expansion, we expect that the solution of the given problem $u(x, t) = w(x, t) + v(x)$ takes a form better suited for numerical evaluation. The same is true for the determination of $v(\mathbf{x})$.

Conversion to Problems with Homogeneous Boundary Data

We have indicated that the nonhomogeneity of the boundary conditions prescribed for the solution of (4.6.1) weakens the convergence rate of the eigenfunction expansion (4.6.3). The following method can be used as a remedy. Given the boundary condition (4.1.11) or (4.1.12), we seek a function $V(\mathbf{x}, t)$ such that

$$\alpha(\mathbf{x})V(\mathbf{x},t) + \beta(\mathbf{x})\frac{\partial V(\mathbf{x},t)}{\partial n}\bigg|_{\partial G} = B(\mathbf{x},t), \tag{4.6.40}$$

or $V(x,t)$ such that $\alpha_1 V(0,t) - \beta_1 V_x(0,t) = g_1(t)$, $\alpha_2 V(l,t) + \beta_2 V_x(l,t) = g_2(t)$. The function $V(\mathbf{x},t)$ or $V(x,t)$ must be differentiable as often as required in the given differential equations. If such a function can be constructed we set (working with the higher-dimensional case) $u(\mathbf{x},t) = W(\mathbf{x},t) + V(\mathbf{x},t)$, and find that $W(\mathbf{x},t)$ satisfies the equation

$$\rho(\mathbf{x})KW(\mathbf{x},t) + LW(\mathbf{x},t) = \rho(\mathbf{x})F(\mathbf{x},t) - \rho(\mathbf{x})KV(\mathbf{x},t) - LV(\mathbf{x},t), \tag{4.6.41}$$

with initial conditions of the form (4.6.39) [where $v(\mathbf{x})$ is replaced by $V(\mathbf{x},0)$ and a term $-V_t(\mathbf{x},0)$ is added to the second equation in (4.6.39), depending on whether the equation is parabolic or hyperbolic]. However, the boundary conditions for $W(\mathbf{x},t)$ are easily found to be homogeneous. *Duhamel's principle* or the *finite transform method* can now be applied to solve for $W(\mathbf{x},t)$.

Example 4.9. The Inhomogeneous Heat Equation. We consider the fully inhomogeneous initial and (Dirichlet) boundary value problem for the one-dimensional heat equation in a finite interval. We have

$$u_t(x,t) - c^2 u_{xx}(x,t) = g(x,t), \qquad 0 < x < l,\ t > 0, \tag{4.6.42}$$

$$u(x,0) = f(x),\ 0 < x < l, \quad u(0,t) = g_1(t),\ u(l,t) = g_2(t),\ t > 0, \tag{4.6.43}$$

where the inhomogeneous term $g(x,t)$ and the data $f(x)$, $g_1(t)$, and $g_2(t)$ are prescribed.

The use of linear interpolation, readily results in the construction of the function

$$V(x,t) = \frac{1}{l}\,[xg_2(t) + (l-x)g_1(t)], \tag{4.6.44}$$

which satisfies the boundary conditions in (4.6.43). Then with $u(x,t) = W(x,t) + V(x,t)$, we find that $W(x,t)$ satisfies the equation

$$W_t(x,t) - c^2 W_{xx}(x,t) = g(x,t) - \frac{1}{l}\,[xg_2'(t) + (l-x)g_1'(t)] \tag{4.6.45}$$

for $0 < x < l$ and $t > 0$, with the initial and homogeneous boundary conditions given as

$$W(x,0) = f(x) - \frac{1}{l}\,[xg_2(0) + (l-x)g_1(0)], \quad W(0,t) = W(l,t) = 0. \tag{4.6.46}$$

It may be noted that if $g(x,t) = 0$ and the functions $g_1(t)$ and $g_2(t)$ are time-independent constants so that the inhomogeneities are stationary, we have $V(x,t) \equiv V(x) = (1/l)\,[xg_2 + (l-x)g_1]$. In that case, $V(x)$ is a solution of the stationary form of the heat equation $-c^2u_{xx}(x,t) = 0$ with the boundary conditions $u(0,t) = g_1$ and $u(l,t) = g_2$. Nevertheless, this does not signify that there is an equivalence between the two decomposition methods given above in all cases in which the given partial differential equation is homogeneous and the boundary conditions are stationary. For example, another possible choice for $V(x,t)$ in the above problem is $V(x,t) = (1/\sinh(l))\,[\sinh(x)g_2(t) + \sinh(l-x)g_1(t)]$. This does not yield a solution of the stationary case if g_1 and g_2 are constant.

Exercises 4.6

4.6.1. Obtain the coefficients N_k in the eigenfunction expansion (4.6.3) for the one-dimensional hyperbolic and parabolic problems.

4.6.2. Let the operator K in (4.6.1) be given as $K = \partial^2/\partial t^2 + 2\hat{\lambda}\partial/\partial t$, where $\hat{\lambda}$ is a positive constant. Given the same initial and boundary conditions as for the hyperbolic problem discussed in the text, determine an expression for the Fourier coefficient N_k analogous to equation (4.6.14).

4.6.3. It was assumed in the text that the eigenvalues λ_k are all positive. Discuss the necessary modifications if $\lambda_0 = 0$ is an eigenvalue and the term $N_0M_0(\mathbf{x})$ occurs in the expansion (4.6.3).

4.6.4. Use the finite sine transform to solve the problem $u_{tt}(x,t) - c^2u_{xx}(x,t) = xe^{-t}$, $0 < x < l, t > 0$, $u(x,0) = 0$, $u_t(x,0) = 0$, $u(0,t) = \sin(t), u(l,t) = 1$.

4.6.5. Apply the finite cosine transform to solve the problem $u_{tt}(x,t) - c^2u_{xx}(x,t) = 0$, $0 < x < l$, $t > 0$, $u(x,0) = 0$, $u_t(x,0) = 0$, $u_x(0,t) = t$, $u_x(l,t) = 0$.

4.6.6. Use the finite sine transform to solve the problem $u_{tt}(x,t) - c^2u_{xx}(x,t) = F(x)\sin(\omega t)$, $0 < x < l$, $t > 0$, where $\omega \neq \pi kc/l$, and the initial and boundary data are $u(x,0) = u_t(x,0) = u(0,t) = u(l,t) = 0$. Consider the limit of the solution $u(x,t)$ as $\omega \to \pi kc/l$ and obtain a resonance effect.

4.6.7. Consider the two-dimensional Laplace's equation $\nabla^2u(r,\theta) = 0$ in the disk $r < R$ with the boundary condition $u(R,\theta) = f(\theta)$. Solve this problem by using finite Fourier transforms.

4.6.8. Use the finite sine transform to solve the Dirichlet problem for Laplace's equation in a rectangle $u_{xx}(x,y) + u_{yy}(x,y) = 0$, $0 < x < l$, $0 < y < \hat{l}$, $u(x,0) = f(x)$, $u(x,\hat{l}) = g(x)$, $u(0,y) = h(y)$, $u(l,y) = r(y)$.

4.6.9. Apply the finite cosine transform to solve the Neumann problem for Poisson's equation in a rectangle $u_{xx}(x,y) + u_{yy}(x,y) = -F(x,y)$, $0 < x < l$, $0 < y < \hat{l}$, $u_y(x,0) = f(x)$, $u_y(x,\hat{l}) = g(x)$, $u_x(0,y) = h(y)$, $u_x(l,y) = r(y)$. Determine conditions on $F(x,y)$ and the boundary values of $u(x,y)$ so that a solution exists.

4.6.10. Solve the Neumann problem for Laplace's equation $\nabla^2 u(r,\theta) = 0$ in the disk $r < R$ if $\partial u(R,\theta)/\partial r = \sin^3(\theta)$.

4.6.11. Consider the equation $Ku(r,\theta,t) - \nabla^2 u(r,\theta,t) = F(r,t)$, $0 < r < R$, $0 < t$, $0 \leq \theta < 2\pi$, where K is either a first or second order partial derivative operator in t with constant coefficients. If $u(R,\theta,t) = B(t)$ and the initial data are functions of r only, we may set $u = u(r,t)$. Show that the equation can then be written in the form $rKu(r,t) + Lu(r,t) = rF(r,t)$, where L is the Bessel operator (with $n = 0$) defined in (4.3.56). Show how the results in the text can be used to develop a finite transform method for solving the given problem. This transform, which involves Bessel functions, is called the *finite Hankel transform. Hint*: See Section 5.5.

4.6.12. Use the finite sine transform to solve the following problem for the heat equation: $u_t(x,t) - c^2 u_{xx}(x,t) = e^{-t}$, $0 < x < l$, $t > 0$, $u(x,0) = 0$, $u(0,t) = \alpha$, $u(l,t) = \beta$, where α and β are constants. Obtain the behavior of the solution as t gets large.

4.6.13. Use the finite cosine transform to solve the following problem: $u_t(x,t) - c^2 u_{xx}(x,t) = 0$, $0 < x < l$, $t > 0$, $u(x,0) = 0$, $u_x(0,t) = 0$, $u_x(l,t) = e^{-t}$.

4.6.14. Consider the problem $u_t(x,t) - c^2 u_{xx}(x,t) = 0$, $0 < x < l$, $t > 0$, $u(x,0) = 0$, $u(0,t) = 0$, $u(l,t) = 1$. (a) Solve the problem by using (4.6.44) to eliminate the inhomogeneous boundary terms. (b) Solve the problem by using the finite sine transform. (c) Compare the rates of convergence of the series solutions obtained in parts (a) and (b).

4.6.15. Given the problem $u_t(x,t) - c^2 u_{xx}(x,t) = 0$, $0 < x < l$, $t > 0$, $u(x,0) = 0$, $u(0,t) = 0$, $u(l,t) = e^{-t}$, determine conditions on c and l such that there is a solution of the homogeneous heat equation in the form $u(x,t) = v(x)e^{-t}$ that satisfies the boundary conditions above. Use this solution to solve the given initial and boundary value problem.

4.6.16. Obtain a solution of the following problem for the heat equation $u_t(x,t) - c^2 u_{xx}(x,t) = 0$, $0 < x < l$, $t > 0$, $u(0,t) = 0$, $u(l,t) = t$, in the form of a polynomial in x and t. Show how this solution can be used to solve the initial and boundary value problem for the heat equation with the boundary data above and $u(x,0) = f(x)$, $0 < x < l$. Discuss the behavior of the solution of the latter problem for large t.

4.6.17. Solve the problem $u_{tt}(x,t) - c^2 u_{xx}(x,t) = 0$, $0 < x < l$, $t > 0$, $u(x,0) = u_t(x,0) = 0$, $u(0,t) = t$, $u(l,t) = 1$, by introducing a change of the dependent variable that renders the boundary conditions homogeneous.

4.7 NONLINEAR STABILITY THEORY: EIGENFUNCTION EXPANSIONS

Nonlinear Heat Equation: Stability Theory

In this section we consider a *nonlinear heat conduction equation* in one dimension and investigate the *stability* of the equilibrium or zero temperature distribution. That is, we wish to determine if small perturbations around the zero temperature distribution at the initial time decay to zero as t increases or perhaps develop into a new stationary solution with increasing t. We are concerned primarily with the problem of heat conduction in an (insulated) rod of finite length with the ends of the rod kept at zero temperature for all time. But to compare results of the present stability analysis with those of Section 3.5, we consider briefly the stability problem for a rod of infinite extent.

Let $u(x,t)$ represent the temperature distribution of an insulated rod and assume that there is a nonlinear heat source of strength $-\hat{\lambda}u(x,t)(1-u^2(x,t))$. Then the (nonlinear) heat equation to be studied has the form

$$u_t(x,t) - u_{xx}(x,t) = \hat{\lambda}u(x,t)(1-u^2(x,t)), \tag{4.7.1}$$

where the parameter $\hat{\lambda}$ depends on the properties of the rod. We are concerned with the growth in time of small initial perturbations around the equilibrium solution $u_0(x,t) = 0$ of (4.7.1), so we introduce the initial condition

$$u(x,0) = \epsilon h(x), \tag{4.7.2}$$

where $h(x)$ is uniformly bounded and $0 < \epsilon \ll 1$.

We are interested in two problems. The basic problem is that in which the rod is of *finite extent*—it is assumed to occupy the interval $0 \le x \le \pi$—and the boundary conditions at the endpoints $x = 0$ and $x = \pi$ are

$$u(0,t) = 0, \qquad u(\pi,t) = 0. \tag{4.7.3}$$

[In this case (4.7.1)–(4.7.2) are valid in the interval $0 < x < \pi$.] The other problem deals with a rod of *infinite extent* so that (4.7.1)–(4.7.2) are given over the interval $-\infty < x < \infty$. Since the data for either problem are small (of order of magnitude ϵ), we look for a solution of each problem in the form

$$u(x,t) = \epsilon w(x,t), \tag{4.7.4}$$

and trace its evolution in time. If this *perturbation* around the solution $u_0(x,t) = 0$ remains small as t increases, the zero solution is *stable*. However, if $w(x,t)$ grows without bound as time increases, the equilibrium solution is *unstable*. Since we are unable to solve either of the foregoing nonlinear problems exactly, we use approximate methods for studying the stability of the equilibrium solutions of the problems.

Inserting (4.7.4) into (4.7.1) and dividing by ϵ gives

$$w_t(x,t) - w_{xx}(x,t) = \hat{\lambda}w(x,t) - \epsilon^2\hat{\lambda}w^3(x,t). \tag{4.7.5}$$

We first carry out a *linear stability analysis*. Assuming that the solution $u(x,t) = \epsilon w(x,t)$ is stable so that it does not grow without bound and noting that $\epsilon^2 \ll 1$, we *linearize* the equation (4.7.5) by dropping the term proportional to $w^3(x,t)$. The resulting linear equation is

$$w_t(x,t) - w_{xx}(x,t) = \hat{\lambda} w(x,t), \tag{4.7.6}$$

and we perform a stability analysis of (4.7.6) with the initial and boundary conditions carried over from the nonlinear problem. If the solutions of the linear problem are stable, our assumption leading to the neglect of the nonlinear term in (4.7.5) is valid for all time. We conclude that the nonlinear problem is stable and is well approximated by the linearized version of the problem.

However, if the linearized problem exhibits instability, the solution $w(x,t)$ grows in time so that, eventually, the term $\epsilon^2 w^3(x,t)$ can attain an order of magnitude equal to that of the term $w(x,t)$. Consequently, the nonlinear term in (4.7.5) cannot be neglected and the full nonlinear equation must be used. Then, although the linear stability analysis predicts that the equilibrium solution $u_0(x,t) = 0$ is unstable and that the perturbation grows without bound, a nonlinear analysis may show that, in fact, the perturbation grows only until it reaches another equilibrium solution. This is demonstrated for the initial and boundary value problem for (4.7.1).

Nonlinear Heat Equation: Cauchy Problem

We begin by considering a *linear stability analysis* of the *Cauchy problem* for (4.7.1). Thus we study the normal mode solutions

$$w(x,t) = a(k) \exp[ikx + \lambda(k)t] \tag{4.7.7}$$

of the linearized equation (4.7.6). Inserting (4.7.7) into (4.7.6) gives $\lambda(k) = \hat{\lambda} - k^2$. Since $\hat{\lambda}$ is a real constant, the stability index Ω defined as in Section 3.5 is given by $\Omega = \hat{\lambda}$. Thus (4.7.6) is strictly stable if $\hat{\lambda} < 0$ and is unstable if $\hat{\lambda} > 0$. For $\hat{\lambda} > 0$, the linear stability analysis predicts unbounded growth for the perturbation ϵw and instability for the equilibrium solution $u_0(x,t) = 0$.

To see that the perturbation $\epsilon w(x,t)$ need not necessarily grow without bound but may reach another (stable) equilibrium solution, we proceed as follows. We first note that $u_1(x,t) = 1$ is a solution of (4.7.1). To check for the stability of this (additional) equilibrium solution, we set $u(x,t) = 1 + \epsilon v(x,t)$, with $0 < \epsilon \ll 1$ and analyze the linearized equation for $v(x,t)$; that is,

$$v_t(x,t) - v_{xx}(x,t) = -2\hat{\lambda} v(x,t). \tag{4.7.8}$$

For the normal mode solutions (4.7.7) [with $w(x,t)$ replaced by $v(x,t)$] we now obtain $\lambda(k) = -2\hat{\lambda} - k^2$, so that the stability index is $\Omega = -2\hat{\lambda}$. Thus for $\hat{\lambda} > 0$, in which case the linear stability analysis showed the zero (equilibrium) solution $u_0(x,t) = 0$ to be unstable, we now have stability for the equilibrium solution

$u_1(x,t) = 1$. This suggests that perturbations of the zero solution do not grow unboundedly but only until they reach the stable equilibrium $u_1(x,t) = 1$. We do not prove this is the case in general, but find an explicit solution of a Cauchy problem for (4.7.1) where this is exactly what happens.

Let $h(x) = 1$ in (4.7.2) and look for a solution of (4.7.1) that is independent of x; that is, $u(x,t) = U(t)$. Then $U(t)$ satisfies the initial value problem for the ordinary differential equation

$$U'(t) = \hat{\lambda}U(t)(1 - U(t)^2), \quad U(0) = \epsilon. \tag{4.7.9}$$

The solution of (4.7.9) obtained by separation of variables is

$$u(x,t) \equiv U(t) = \frac{\epsilon e^{\hat{\lambda}t}}{\sqrt{1 + \epsilon^2(e^{2\hat{\lambda}t} - 1)}}. \tag{4.7.10}$$

For small values of t, we expand this solution in powers of ϵ and find that $u(x,t) \approx \epsilon e^{\hat{\lambda}t} = \epsilon w(x,t)$, where $w(x,t)$ satisfies (4.7.6) and $w(x,0) = 1$. $w(x,t) \to \infty$ as $t \to \infty$ if $\hat{\lambda} > 0$, but (4.7.10) does not tend to infinity but approaches the equilibrium solution $u_1(x,t) = 1$. If $\hat{\lambda} < 0$, the term $e^{2\hat{\lambda}t}$ in the denominator of (4.7.10) tends to zero as $t \to \infty$. Thus since $e^{2\hat{\lambda}t} - 1 \approx 0$ for small t and $\epsilon^2 \ll 1$, we conclude that $u(x,t)$ as given in (4.7.10) is well approximated by $\epsilon w(x,t) = \epsilon e^{\hat{\lambda}t}$ for all time. That is, the linearization procedure is valid for all time if the linearized problem is stable.

Nonlinear Heat Equation: Initial and Boundary Value Problem

Turning now to the consideration of the *initial and boundary value problem* (4.7.1)–(4.7.3) in the interval $0 < x < \pi$, we begin by applying *linear stability theory*. We look for normal mode solutions of the linearized equation (4.7.6) that satisfy the homogeneous boundary conditions (4.7.3). Although (4.7.7) yields solutions of (4.7.6) for all real k, to find solutions that vanish at $x = 0$ and $x = \pi$ we must take linear combinations of the normal modes and restrict the values of k. We obtain

$$w_n(x,t) = h_n \exp[(\hat{\lambda} - n^2)t] \sin(nx), \qquad n = 1, 2, \ldots, \tag{4.7.11}$$

where h_n is a constant and we have set $k = n$ since $w(x,t)$ vanishes at $x = 0$ and $x = \pi$ for these values of k. [The solutions (4.7.11) are precisely what results if separation of variables is applied to (4.7.6) with the boundary conditions (4.7.3).]

The solution of the initial and boundary value problem for the equation (4.7.6) with the conditions (4.7.2)–(4.7.3) is a linear combination of the (normal mode) solutions (4.7.11). Thus the stability properties are determined from the discrete set of solutions $w_n(x,t)$ given in (4.7.11) rather than from the continuous set (4.7.7) with $-\infty < k < \infty$, which is relevant for the Cauchy problem. We see that if $\hat{\lambda} < 1$, the $w_n(x,t)$ in (4.7.11) all decay to zero as $t \to \infty$, whereas if $\hat{\lambda} > 1$, at least one of the $w_n(x,t)$ grows exponentially as $t \to \infty$. Consequently, the linear

stability analysis predicts that the equilibrium solution $u_0 = 0$ is unstable to small perturbations if $\hat{\lambda} > 1$, while if $\hat{\lambda} < 1$, it is stable and all perturbations eventually die out as t increases. Again, if $\hat{\lambda} > 1$, the *linear stability analysis* that predicts instability becomes invalid after a finite time since the nonlinear term in (4.7.5) becomes equally important with the linear term given the growth of $w(x, t)$. Thus, a *nonlinear stability analysis* of the growth of the perturbation term $u = \epsilon w$ is needed.

The values $\hat{\lambda} = 0$ and $\hat{\lambda} = 1$ given for the Cauchy and the initial and boundary value problems for the linearized equation (4.7.6) determine the *threshold of instability* (in the linear theory) for the given problems. These values of $\hat{\lambda}$ are denoted by $\hat{\lambda}_c$, the *critical value* of $\hat{\lambda}$. That is, if $\hat{\lambda} > \hat{\lambda}_c$, we have instability, and if $\hat{\lambda} < \hat{\lambda}_c$, we have stability. It should be noted that the critical value is determined not only from the given (linearized) equation but also by the boundary conditions, if any are given. As shown above, the critical values $\hat{\lambda}_c$ differ in the case where no boundary conditions are assigned and in the case where boundary conditions are given.

The initial and boundary value problem (4.7.1)–(4.7.3) cannot be solved exactly. There are several approximate methods for analyzing this nonlinear problem, but we consider only one approach, which is based on eigenfunction expansions or equivalently, finite Fourier transforms.

The operator $L = -\partial^2/\partial x^2$ in (4.7.1), together with the boundary conditions (4.7.3), is associated with the eigenfunctions $M_k(x) = \sqrt{2/\pi}\sin(kx)$, $k = 1, 2, \ldots$ [i.e., $LM_k(x) = \lambda_k M_k(x)$ with $\lambda_k = k^2$ and $M_k(0) = M_k(\pi) = 0$]. The normalized set $\{M_k(x)\}$ is complete and we represent the function $w(x, t)$ (recall that $u = \epsilon w$) in the form

$$w(x,t) = \sum_{k=1}^{\infty} N_k(t) M_k(x), \qquad N_k(t) = (w(x,t), M_k(x)). \tag{4.7.12}$$

Using the initial condition (4.7.2), we have

$$N_k(0) = (h(x), M_k(x)). \tag{4.7.13}$$

[The inner product for this problem is $(f(x), g(x)) = \int_0^\pi f(x)g(x)\,dx$.]

We obtain an equation for $N_k(t)$ by taking a *finite sine transform* in (4.7.5); that is, we multiply by $M_k(x)$ and integrate between 0 and π. This gives us

$$N_k'(t) + k^2 N_k(t) = \hat{\lambda} N_k(t) - \hat{\lambda}\epsilon^2 \int_0^\pi w^3(x,t) M_k(x)\,dx, \; k = 1, 2, \ldots. \tag{4.7.14}$$

The integral term yields

$$\int_0^\pi w^3(x,t) M_k(x)\,dx = \sum_{i,j,l=1}^{\infty} a_{ijl}^{(k)} N_i(t) N_j(t) N_l(t), \tag{4.7.15}$$

where the coefficients $a_{ijl}^{(k)}$ are obtained by cubing the series for $w(x, t)$ and integrating term by term. This yields the infinite system of coupled equations

$$N_k'(t) + (k^2 - \hat{\lambda})N_k(t) = -\hat{\lambda}\epsilon^2 \sum_{i,j,l=1}^{\infty} a_{ijl}^{(k)} N_i(t)N_j(t)N_l(t), \qquad (4.7.16)$$

with the initial conditions (4.7.13).

Neglecting the terms of order ϵ^2 in the equations (4.7.16) is equivalent to linearizing the problem and has the effect of uncoupling the equations. Then the $N_k(t)M_k(x)$ have the form (4.7.11) with n replaced by k. We observe that for $\hat{\lambda} < \hat{\lambda}_c = 1$ all the $N_k(t)$ in the linearized case tend to zero as $t \to \infty$, whereas for $\hat{\lambda} > \hat{\lambda}_c = 1$, at least one of the $N_k(t)$ grows exponentially in t.

We wish to examine the behavior of the solution if $\hat{\lambda}$ is slightly larger than the critical value $\hat{\lambda}_c$; that is, $\hat{\lambda} \approx 1$ but $\hat{\lambda} > 1$. Then $k^2 - \hat{\lambda} > 0$ for $k \geq 2$ but $1 - \hat{\lambda} < 0$. It is of interest to study the behavior of the solution in the neighborhood of the critical value where according to the linear theory a transition from stability to instability takes place.

Since for $k \geq 2$, $k^2 - \hat{\lambda} > 0$, we take as a first approximation

$$N_k(t) = (h(x), M_k(x)) \exp[(\hat{\lambda} - k^2)t] \qquad k \geq 2; \qquad (4.7.17)$$

that is, we neglect the terms of order ϵ^2 in (4.7.14) with $k \geq 2$. A similar approximation in the equation for $N_1(t)$ leads to an exponentially growing term of the form (4.7.17) with $k = 1$. Inserting these expressions for $N_k(t)$ $(k = 1, 2, ...)$ into the series on the right side of (4.7.16) shows that all the terms in the sum decay exponentially except for the terms $a_{111}^{(k)} N_1^3(t)$.

The linearized form of the equation for $N_1(t)$ implies that it is growing exponentially, while all other $N_k(t)$ decay exponentially, so we retain the term $\hat{\lambda}\epsilon^2 a_{111}^{(1)} N_1^3(t)$ in the equation for $N_1(t)$ and obtain

$$N_1'(t) + (1 - \hat{\lambda})N_1(t) + \hat{\lambda}\epsilon^2 a_{111}^{(1)} N_1^3(t) = 0. \qquad (4.7.18)$$

For the coefficient $a_{111}^{(1)}$ we have from (4.7.15)

$$a_{111}^{(1)} = \int_0^{\pi} (M_1(x))^4 \, dx = \frac{4}{\pi^2} \int_0^{\pi} \sin(x)^4 \, dx = \frac{3}{2\pi}. \qquad (4.7.19)$$

Thus $N_1(t)$ satisfies

$$N_1'(t) + (1 - \hat{\lambda})N_1(t) + \frac{3\hat{\lambda}}{2\pi} \epsilon^2 N_1^3(t) = 0. \qquad (4.7.20)$$

To solve (4.7.20) we multiply across by $N_1(t)$ and obtain

$$\frac{1}{2} \frac{d(N_1^2(t))}{dt} + (1 - \hat{\lambda})N_1^2(t) + \frac{3\hat{\lambda}\epsilon^2}{2\pi} N_1^4(t) = 0, \qquad (4.7.21)$$

which is a *Riccati equation* for $N_1^2(t)$. The solution of (4.7.21) satisfying $N_1(0) = (h, M_1)$ is found to be

$$N_1(t) = \frac{(h(x), M_1(x))e^{(\hat{\lambda}-1)t}}{[1 + [3\hat{\lambda}\epsilon^2(h(x), M_1(x))^2/(2\pi(\hat{\lambda}-1))](e^{2(\hat{\lambda}-1)t} - 1)]^{1/2}}. \tag{4.7.22}$$

For small values of t this expression reduces to $N_1(t) \approx (h(x), M_1(x))e^{(\hat{\lambda}-1)t}$, which is the solution of the linearized equation, as was expected. However, as $t \to \infty$, since $\hat{\lambda} - 1 > 0$ we have

$$\lim_{t\to\infty} N_1(t) = \frac{(h(x), M_1(x))}{|(h(x), M_1(x))|}\sqrt{\frac{2\pi}{3}}\left(\frac{\hat{\lambda}-1}{\hat{\lambda}\,\epsilon^2}\right)^{1/2}, \tag{4.7.23}$$

so that $N_1(t)$ tends to a finite (stationary) value, rather than growing without bound as predicted by the linear theory.

We remark that the above results are self-consistent since we now find that the terms $a_{111}^{(k)}N_1^3(t)$ in the equations (4.7.16) for $k \geq 2$ are uniformly bounded for all t and it is correct to approximate $N_k(t)$ as in (4.7.17) since the right side of (4.7.16) is of order ϵ^2. Thus we obtain from (4.7.12), with $u(x,t) = \epsilon\, w(x,t)$,

$$\epsilon w(x,t) \approx \frac{(h(x), M_1(x))}{|(h(x), M_1(x))|}\frac{2}{\sqrt{3}}\left(\frac{\hat{\lambda}-1}{\hat{\lambda}}\right)^{1/2}\sin(x), \tag{4.7.24}$$

as $t \to \infty$ in view of (4.7.17) and (4.7.23). We assume that $(h(x), M_1(x)) \neq 0$ and (4.7.24) shows that the solution at large time depends only on the sign [i.e., $(h(x), M_1(x))/|(h(x), M_1(x))|$] of the leading term of the Fourier series of $u(x,0)$ and not on its magnitude. The large time behavior of the solution is independent of t, so that the solution $u(x,t)$ with $\hat{\lambda} \geq \hat{\lambda}_c$ does not grow unboundedly as predicted by the linear theory. Instead, it approaches a *steady state* whose approximate description (correct to order ϵ) is given in (4.7.24). We note that (4.7.24) is an approximate solution of the stationary form of (4.7.1) [i.e., with $u_t(x,t) = 0$] when $\hat{\lambda} \approx \hat{\lambda}_c = 1$.

The foregoing approximation method may be extended to deal with nonlinear heat sources of the form $\hat{\lambda}f(u)$, where $f(0) = 0$ and $f(u)$ has a Taylor expansion around $u = 0$. Also, hyperbolic and elliptic equations with similar types of nonlinearities can be treated by this method. Some examples are presented in the exercises.

Exercises 4.7

4.7.1. Show that (4.7.23) is a solution of the equation (4.7.20).

4.7.2. Obtain the solution (4.7.22) of the equation (4.7.20).

4.7.3. Consider the nonlinear heat equation $u_t(x,t) - u_{xx}(x,t) = \hat{\lambda}u(x,t)[1 - u(x,t)]$, $-\infty < x < \infty$, $t > 0$. Carry out a linear stability analysis around the two solutions $u_0(x,t) = 0$ and $u_1(x,t) = 1$.

4.7.4. Solve the Cauchy problem for $u_t(x,t) - u_{xx}(x,t) = \hat{\lambda}u(x,t)[1-u(x,t)]$, $-\infty < x < \infty$, $t > 0$, with the initial condition $u(x,0) = \epsilon$, $0 < \epsilon \ll 1$, by looking for a solution independent of x. Expand the solution for small ϵ and compare the result with that obtained by solving a linearized version of the problem. Obtain the limit of the solution of the nonlinear problem as $t \to \infty$.

4.7.5. Apply the method of eigenfunction expansions to solve (approximately) the nonlinear heat equation, $u_t(x,t) - u_{xx}(x,t) = \hat{\lambda}u(x,t)(1 - u(x,t))$, $0 < x < \pi$, $t > 0$ if $u(x,0) = \epsilon h(x)$, $0 < \epsilon \ll 1$, $u(0,t) = u(\pi,t) = 0$. Determine the critical value $\hat{\lambda}_c$ of $\hat{\lambda}$ and solve the problem if $\hat{\lambda}$ is slightly larger than $\hat{\lambda}_c$.

4.7.6. Consider the boundary value problem for the nonlinear elliptic equation $u_{xx}(x,y) + u_{yy}(x,y) = -\hat{\lambda}u(x,y)(1 - u^2(x,y))$, $0 < x < \pi$, $y > 0$ with $u(0,y) = u(\pi,y) = 0$, $u(x,0) = \epsilon h(x)$. Requiring that $u(x,y) \to 0$ as $y \to \infty$, apply a linear stability analysis to obtain a critical value $\hat{\lambda}_c$ of $\hat{\lambda}$. Expand the solution in the form $u(x,y) = \epsilon \sum_{k=1}^{\infty} N_k(y) \left(\sqrt{2/\pi} \sin kx\right)$ and obtain an (approximate) nonlinear equation for $N_1(y)$ and linear equations for the $N_k(y)$ with $k \geq 2$. Discuss the behavior of the nonlinear equation for $N_1(y)$ as far as possible without necessarily solving it, and consider the behavior of the solution $u(x,y)$ as $y \to \infty$.

4.7.7. Consider the nonlinear equation $u_{tt}(x,t) - u_{xx}(x,t) = -\hat{\lambda}u(x,t)(1-u^2(x,t))$, $-\infty < x < \infty$, $t > 0$, of hyperbolic type. (a) Show that if the equation is linearized for small $u(x,t)$ and $\hat{\lambda} > 0$, there exist exponentially growing solutions. (b) Set $u(x,t) = 1 + \epsilon v(x,t)$ and obtain a linearized equation for $v(x,t)$ in the form of the Klein-Gordon equation if $\hat{\lambda} > 0$, and show that it is neutrally stable.

4.7.8. Consider the initial and boundary value problem for the nonlinear hyperbolic equation $u_{tt}(x,t) - u_{xx}(x,t) = -\hat{\lambda}u(x,t)(1 - u^2(x,t))$, $0 < x < \pi$, $t > 0$, $u(x,0) = \epsilon h(x)$, $u_t(x,0) = 0$, $u(0,t) = u(\pi,t) = 0$. Use a linear stability analysis around the solution $u_0(x,t) = 0$ to determine a critical value $\hat{\lambda}_c$ such that exponentially growing normal mode solutions exist if $\hat{\lambda} > \hat{\lambda}_c$. Using eigenfunction expansions in the manner presented in the text, discuss this nonlinear initial and boundary value problem if $\hat{\lambda}$ is slightly greater than $\hat{\lambda}_c$.

4.8 MAPLE METHODS

The Maple procedure *pdsolve* can find general solutions for PDEs in various forms. We have seen in Chapter 2 that it has some success in dealing with initial value problems for first order PDEs in that it finds the characteristic equations, and these can then be solved using the ODE procedure *dsolve*. General solutions of the PDEs can also be used for this purpose. When dealing with second or higher order PDEs, the results are less useful. Although some general information is dispensed by Maple, exact solutions (even in series form) of initial and boundary value problems for these PDEs are hard to obtain using *pdsolve* without a good deal of further analysis.

For example, if we invoke $pdsolve(WaveEquation)$ where the argument is the one-dimensional wave equation (4.4.1), Maple's output is the general solution

$u(x,t) = F_1(ct+x)+F_2(ct-x)$. The output of $pdsolve(WaveEquation, HINT = M(x)N(t))$, where we ask for a solution of the wave equation using separation of variables in the form $u(x,t) = M(x)N(t)$, is $u(x,t) = M(x)N(t), \{d^2N(t)/dt^2 = c^2\lambda N(t), d^2M(x)/dx^2 = \lambda M(x)\}$. For the heat equation (4.4.25), $pdsolve(HeatEquation)$ yields $u(x,t) = M(x)N(t)$, and $\{dN(t)/dt = c^2\lambda N(t), d^2M(x)/dx^2 = \lambda M(x)\}$. $pdsolve(LaplaceEquation)$ gives a closed-form solution for Laplace's equation (4.4.31) in complex form. If we add the argument HINT = $M(x)N(y)$, Maple's output is $u(x,y) = M(x)N(y), \{d^2N(y)/dy^2 = -\lambda N(y), d^2M(x)/dx^2 = \lambda M(x)\}$.

For each of the PDEs considered above, Maple indicates that they are separable and exhibits the separated ODEs. Similar results are given for higher dimensional versions of these PDEs. To use these results to determine solutions of initial and boundary value problems for these PDEs, eigenvalues and eigenfunctions and their expansions must be determined. Although Maple can solve many boundary value problems for ODEs, it does not have a procedure for finding eigenvalues and eigenfunctions for ODEs.

Eigenvalue Problems for ODEs

We have constructed two procedures that determine *eigenvalues* and *eigenfunctions* for *second order ODEs*, which need not necessarily be self-adjoint. These procedures can be applied to solve *regular* and *singular Sturm-Liouville problems*. The main procedure *EVProbODE* determines the characteristic equation for the determination of the eigenvalues and the corresponding eigenfunctions. Maple's *dsolve* procedure is used to solve the given ODE, and the boundary conditions are used to determine an equation for the eigenvalues. Generally, the eigenvalue equation cannot be solved explicitly to give a full set of eigenvalues. The related procedure $Evalues$ deals with this problem. It determines the eigenvalues in a specified interval numerically if the problem has no parameters. The equation for the determination of the eigenvalues and the general expression for the eigenfunctions must be entered. These can be given as the global variables $CharEqn$ and $EFunc$ as determined from the procedure $EVProbODE$.

The procedure $EVProbODE(EVEquation, x = a..b, BCL, BCR, v(x), \lambda)$ has the following arguments. The first four arguments give the ODE for the eigenvalue problem, with the boundary conditions BCL and BCR at the left and right endpoints $x = a$ and $x = b$, respectively, of the interval $[a, b]$. The dependent variable is v and the eigenvalue parameter is λ.

We apply *EVProbODE* to the eigenvalue problem of Section 4.3 and obtain the following output, which agrees essentially with (4.3.32)–(4.3.33);

$$\text{Eigenvalues} = (\alpha_1\alpha_2 - \lambda\beta_1\beta_2)\sin(\sqrt{\lambda}\,l) + \sqrt{\lambda}(\alpha_1\beta_2 + \beta_1\alpha_2)\cos(\sqrt{\lambda}\,l) = 0,$$

$$\text{Eigenfunctions} = \beta_1\sqrt{\lambda}\cos(\sqrt{\lambda}\,x) + \alpha_1\sin(\sqrt{\lambda}\,x). \qquad (4.8.1)$$

As a further result, not given in the example, we obtain the squared norm for the eigenfunctions found above (valid for $\lambda \neq 0$). They are determined by using Maple's *int* procedure, and found to be

$$\text{Squared Norm} = \frac{\left(\beta_1\alpha_1^2\alpha_2 - \alpha_1^3\beta_2 + \beta_1^3\lambda\alpha_2 - \lambda\beta_1^2\beta_2\alpha_1\right)\cos(\sqrt{\lambda}\,l)^2}{2(\beta_1\lambda\beta_2 - \alpha_1\alpha_2)}$$

$$+\frac{2\beta_1^2\alpha_1\lambda\beta_2 - 2\beta_1\alpha_1^2\alpha_2 + \lambda^2\beta_1^3 l\beta_2 - \lambda\beta_1^2 l\alpha_1\alpha_2 + \alpha_1^2 l\beta_1\lambda\beta_2 - \alpha_1^3 l\alpha_2}{2(\beta_1\lambda\beta_2 - \alpha_1\alpha_2)}.$$

For simple problems such as those of the subsections of Section 4.3 that deal with the basic trigonometric eigenfunctions, the eigenvalue equation can be solved exactly to determine closed-form expressions for the eigenvalues. However, if we consider a mixed problem for (4.3.29) with $\alpha_1 = 1$, $\alpha_2 = 4$, $\beta_1 = 0$, $\beta_2 = 1$, $l = 1$, we obtain the eigenvalue equation

$$\cos\left(\sqrt{\lambda}\right)\sqrt{\lambda} + 4\sin\left(\sqrt{\lambda}\right) = 0 \tag{4.8.2}$$

with the unnormalized eigenfunctions $\sin\left(\sqrt{\lambda}\,x\right)$. The eigenvalues must be determined numerically, and we use the procedure $Evalues$ to do so.

The procedure $Evalues(CharEqn, EFunc, \lambda, int, n)$ finds eigenvalues in a specified interval numerically. The eigenvalue equation and eigenfunction must be given. They can be given as the global variables $CharEqn$ and $EFunc$ as determined on applying the procedure $EVProbODE$ to the same eigenvalue problem. The eigenvalue parameter equals λ. In keeping with our use of the procedures for Sturm-Liouville problems, the left endpoint of the λ interval is $\lambda = 0$. The interval chosen is divided into n subintervals, with n as the fifth argument in the procedure. The length of the interval equals $n \times int$, where int is the fourth argument in the procedure.

For the foregoing example with the eigenvalue or characteristic equation (4.8.2), we invoke $Evalues(CharEqn, EFunc, \lambda, 10, 90)$ after $EVProbODE$ has been entered as above. The λ-interval $[0, 900]$ is broken up into 90 subintervals of length 10. Numerically determined eigenvalues $\{\lambda_k\}$, $k = 1, 2, \ldots, 10$ are found to be [6.607, 28.666, 68.939, 128.48, 207.59, 306.37, 424.86, 563.06, 721.00, 898.67]. The corresponding unnormalized eigenfunctions are $\{\sin(\sqrt{\lambda_k}\,x)\}$, $k = 1, 2, \ldots, 10$. Their norms can be found by specializing to this case the general result for the squared norm given above.

In the following example we consider a singular Sturm-Liouville problem that arises when the *Schrödinger equation* of quantum physics,

$$i\hbar u_t(x,t) = -\frac{\hbar^2}{2m}\nabla^2 u(x,t) + V(x,t)u(x,t), \tag{4.8.3}$$

is solved by separation of variables.

Example 4.10. An Eigenvalue Problem that Involves Hermite Polynomials. Separation of variables in the quantum mechanical *harmonic oscillator* equation in one dimension leads to the following *singular Sturm-Liouville problem* over an infinite interval

$$-v''(x) + x^2\,v(x) = \lambda\,v(x), \quad -\infty < x < \infty, \quad v(x) \text{ bounded as } |x| \to \infty. \tag{4.8.4}$$

We apply the procedure $EVProbODE(-v''(x)+x^2v(x)=\lambda v(x), x=-\infty..\infty, v(x)=bounded, v(x)=bounded, v(x), \lambda)$ (the third and fourth arguments ask for bounded solutions at $\pm\infty$) and obtain for the eigenvalues

$$\frac{1}{\sqrt{x}}WhittakerW(\lambda/4, 1/4, x^2) = \text{bounded}, \tag{4.8.5}$$

and for the eigenfunctions

$$\frac{1}{\sqrt{x}}WhittakerW(\lambda/4, 1/4, x^2). \tag{4.8.6}$$

Maple solves the ODE in terms of the Whittaker function $WhittakerW$.

The eigenvalues λ must be chosen to satisfy (4.8.5). On trying a set of positive integral values of λ in the ODE, *dsolve* shows that if $\lambda = 2k+1$ where $k = 0, 1, 2, \ldots$, there is one set of solutions of the form $e^{-x^2/2}P_k(x)$ with $P_k(x)$ as a polynomial of degree k. (They are multiples of *Hermite polynomials*.) They are bounded at infinity and it turns out that only these values of λ yield bounded solutions. To see this, the behavior of the solutions at infinity must be determined. $convert(1/\sqrt{x}\,WhittakerW(\lambda/4, 1/4, x^2), Hermite)$ yields the expression $HermiteH(-1/2+\lambda/2, x)2^{(1/2-\lambda/2)}\exp(-x^2/2)$, where $Hermite(\alpha, x)$ is a Hermite function. With the foregoing choice of λ, the Hermite functions become Hermite polynomials.

Noting this relation, it is easily shown that the transformation $v(x) = e^{-x^2/2}w(x)$ in (4.8.4) yields *Hermite's differential equation*, $w''(x)-2xw'(x)+(\lambda-1)w(x) = 0$. The polynomial solutions $P_k(x)$ of this equation with $\lambda = 2k+1$, $k = 0, 1, 2, \ldots$ are commonly chosen to be the *Hermite polynomials* $H_k(x)$. The first five Hermite polynomials are $H_0(x) = 1$, $H_1(x) = 2x$, $H_2(x) = 4x^2 - 2$, $H_3(x) = 8x^3 - 12x$, $H_4(x) = 16x^4 - 48x^2 + 12$. (In Maple notation the Hermite polynomials are given as $HermiteH(k, x)$.) The eigenvalues and orthonormalized set of eigenfunctions for the Sturm-Liouville problem (4.8.4) are

$$\lambda_k = 2k+1, \quad v_k(x) = \frac{1}{\sqrt{2^k\,\pi^{1/2}\,k!}}e^{-x^2/2}H_k(x), \quad k = 0, 1, 2, \ldots. \tag{4.8.7}$$

Continuing our discussion of singular Sturm-Liouville problems, we consider an eigenvalue problem for an ODE related to *Laguerre's differential equation*.

Example 4.11. An Eigenvalue Problem that Involves Laguerre Polynomials. We consider the following *singular Sturm-Liouville problem* over a semi-infinite interval:

$$-xv''(x) - v'(x) + \left(\frac{x}{4} - \frac{1}{2}\right)v(x) = \lambda v(x), \qquad 0 < x < \infty,$$
$$v(x) \text{ bounded as } x \to 0, \quad x \to \infty. \tag{4.8.8}$$

Note that for this problem, $q(x) = x/4 - 1/2$ is nonnegative only for $x \geq 2$ and that $x = 0$ is a regular singular point for the ODE. As a result, we ask for a solution that is bounded at $x = 0$ and at infinity.

We apply the procedure $EVProbODE(-x\,v''(x) - v(x) + (x/4 - 1/2)\,v(x) = \lambda v(x), x = 0..\infty, v(x) = bounded, v(x) = bounded, v(x), \lambda)$ and obtain for the eigenvalue equation $WhittakerM(1/2 + \lambda, 0, x)/\sqrt{x}$ = bounded, and for the eigenfunctions $WhittakerM(1/2 + \lambda, 0, x)/\sqrt{x}$. The solution is given by Maple in terms of the $WhittakerM$ function. (Replacing the $WhittakerM$ function by the $WhittakerW$ function also yields a solution of the ODE.)

The Maple command *series* shows that $WhittakerW(1/2 + \lambda, 0, x)/\sqrt{x}$ has an infinite limit at $x = 0$, whereas $WhittakerM(1/2 + \lambda, 0, x)/\sqrt{x}$ is finite there. Therefore, the $WhittakerM$ function is appropriate for this problem. The eigenvalues are determined from the condition that $WhittakerM(1/2 + \lambda, 0, x)/\sqrt{x}$ must be bounded at infinity. We try a set of positive integral values of λ in the ODE, and *dsolve* shows that if $\lambda = k$ where $k = 0, 1, 2, \ldots$, there is one set of solutions of the form $e^{-x/2}P_k(x)$ with $P_k(x)$ as a polynomial of degree k. (They are multiples of *Laguerre polynomials*.) These solutions are bounded at infinity and it turns out that only these values of λ yield bounded solutions. The Maple procedure $convert(1/\sqrt{x}WhittakerM(\lambda + 1/2, 0, x), Laguerre)$ yields the expression $Laguerre(\lambda, x)\exp(-x/2)$, where $Laguerre(\lambda, x)$ is a Laguerre function. With the foregoing choice of λ, the Laguerre functions become Laguerre polynomials.

It is easily shown that the transformation $v(x) = e^{-x/2}w(x)$ in (4.8.8) yields *Laguerre's differential equation*, $xw''(x) + (1 - x)w'(x) + \lambda w(x) = 0$. The polynomial solutions $P_k(x)$ of the equation with $\lambda = k$, $k = 0, 1, 2, \ldots$, are commonly chosen to be the *Laguerre polynomials* $L_k(x)$. The first five Laguerre polynomials are $L_0(x) = 1$, $L_1(x) = 1 - x$, $L_2(x) = 1 - 2x + x^2/2$, $L_3(x) = -x^3/6 + 3x^2/2 - 3x + 1$, $L_4(x) = x^4/24 - 2x^3/3 + 3x^2 - 4x + 1$. [In Maple notation the Laguerre polynomials are given as $LaguerreL(k, x)$.] The eigenvalues and orthonormal set of eigenfunctions for the Sturm-Liouville problem (4.8.8) are

$$\lambda_k = k, \qquad v_k(x) = e^{-x/2}L_k(x), \qquad k = 0, 1, 2, \ldots. \tag{4.8.9}$$

Trigonometric Fourier Series

We restrict our discussion to *Fourier series* given over the interval $-l \leq x \leq l$ and to *Fourier sine series* and *Fourier cosine series* given over the interval $0 \leq x \leq l$.

The procedure *FourierCoeff*$(f(x), x, n, l)$ determines the nth Fourier coefficient(s) of the Fourier series of $f(x)$ over $-l \leq x \leq l$. The output for $n \neq 0$ is

$$a_n = \frac{1}{l}\int_{-l}^{l} f(x)\cos\left(\frac{n\pi x}{l}\right)dx, \qquad b_n = \frac{1}{l}\int_{-l}^{l} f(x)\sin\left(\frac{n\pi x}{l}\right)dx. \tag{4.8.10}$$

If a fifth argument *numeric* is added in the procedure, the Fourier coefficient integrals are evaluated numerically. For example, if $f(x) = (x - 6)^3\cos(x^3)$ is

expanded in a Fourier series over $-4 \leq x \leq 4$, the integrals that determine the Fourier coefficients cannot be evaluated in closed form. The use of *FourierCoeff*$((x-6)^3\cos(x^3), x, 33, 4, numeric)$ determines that the thirty-third Fourier coefficients are $\frac{1}{4}\int_{-4}^{4}(x-6)^3\cos(x^3)\cos\left(\frac{33}{4}\pi x\right)dx = -46.78574768$, and $\frac{1}{4}\int_{-4}^{4}(x-6)^3\cos(x^3)\sin\left(\frac{33}{4}\pi x\right)dx = -14.99891264$.

The procedure $FourierSeries(f(x), x, n, l)$ determines $n+1$ terms in the Fourier series of $f(x)$ over $-l \leq x \leq l$. (We can set $n = \infty$.) For example, the Fourier series of the Dirac delta function $\delta(x)$ over the interval $-l \leq x \leq l$ is given as

$$\delta(x) = \frac{1}{2l} + \frac{1}{l}\sum_{k=1}^{\infty}\cos\left(\frac{k\pi x}{l}\right). \tag{4.8.11}$$

The expansion of $f(x) = (x-6)^3\cos(x^3)$ in a Fourier series over $-4 \leq x \leq 4$ up to six terms is carried out numerically by invoking $FourierSeries((x-6)^3\cos(x^3), x, 5, 4, numeric)$. The output (with only four significant digits retained) is $(x-6)^3\cos(x^3) \approx -49.01 - 78.94\cos(.7855\,x) + .9671\sin(.7855\,x) - 84.1\cos(1.571\,x) + 7.594\sin(1.571\,x) - 53.95\cos(2.356\,x) + 19.71\sin(2.356\,x) - 21.88\cos(3.142\,x) + 29.86\sin(3.142\,x) + 36.76\cos(3.928\,x) + 24.18\sin(3.928\,x)$. The numerical result can be evaluated at any point in the given interval and can be plotted. Clearly, the approximation can be improved by finding more terms in the series. We note that $f(x) = (x-6)^3\cos(x^3)$ does not satisfy $f(-4) = f(4)$, one of the periodic boundary conditions for the Fourier series given in (4.3.43). As a result, the periodic extension of $f(x)$ is discontinuous and the convergence rate of the Fourier series is slow, so that many terms in the series are needed to yield a good approximation. Indeed, we find that $f(0) = -216$, while the partial sum found above has the value -251.12 at $x = 0$.

The procedure *FourierSineCoeff*$(f(x), x, n, l)$ determines the nth Fourier sine coefficient of the Fourier sine series of $f(x)$ over $0 \leq x \leq l$. The output is $b_n = (2/l)\int_0^l f(x)\sin(n\pi x/l)\,dx$. If a fifth argument $numeric$ is added in the procedure, the Fourier sine coefficient integrals are evaluated numerically. As an example, we find the nth Fourier sine coefficient of $f(x) = x$ to be $b_n = -2l(-1)^n/n\pi$.

The procedure $FourierSineSeries(f(x), x, n, l)$ determines n terms in the Fourier sine series of $f(x)$ over $0 \leq x \leq l$. (We can set $n = \infty$.) For example, the Fourier sine series of $f(x) = x$ is

$$x = \sum_{k=1}^{\infty}\frac{2(-1)^{1+k}l}{k\pi}\sin\left(\frac{k\pi x}{l}\right). \tag{4.8.12}$$

We note that the rate of convergence of the series is only $O(1/k)$, because the odd periodic extension of x is discontinuous. However, the sine series for $f(x) = x^2(x-l)^2$ is given as

$$x^2(x-l)^2 = \sum_{k=1}^{\infty}\frac{4l^4}{k^5\pi^5}\left(-12 + k^2\pi^2\right)\left(-1 + (-1)^k\right)\sin\left(\frac{k\pi x}{l}\right). \tag{4.8.13}$$

The series converges on the order of $O(1/k^5)$ because the function and its first derivatives vanish at $x = 0$ and $x = l$, so that the odd periodic extension is a smooth function.

Finally, the procedure *FourierCosineCoeff*$(f(x), x, n, l)$ determines the nth Fourier cosine coefficient of the Fourier cosine series of $f(x)$ over $0 \leq x \leq l$. The output is (for $n \neq 0$) $a_n = (2/l)\int_0^l f(x)\cos(n\pi x/l)\,dx$. If a fifth argument *numeric* is added in the procedure, the Fourier cosine coefficient integrals are evaluated numerically. As an example, we find the nth Fourier cosine coefficient of $f(x) = x$ to be $a_n = [2l(-1 + (-1)^n)]/n^2\pi^2$.

The procedure $FourierCosineSeries(f(x), x, n, l)$ determines n terms in the Fourier cosine series of $f(x)$ over $0 \leq x \leq l$. (We can set $n = \infty$.) For example, the Fourier cosine series of $f(x) = x$ is

$$x = \frac{1}{2\,l} + \sum_{k=1}^{\infty} \frac{2\,l}{k^2\pi^2}\left(-1 + (-1)^k\right)\cos\left(\frac{k\pi x}{l}\right). \tag{4.8.14}$$

The cosine series converges more rapidly that the sine series because the even periodic extension of $f(x) = x$ is continuous.

As indicated, partial sums of the three types of Fourier series considered above can be plotted. This is of interest in demonstrating the *Gibbs phenomenon* graphically. It asserts that partial sums of Fourier series of a function $f(x)$ overshoot and undershoot the values of the extended function at points of discontinuity.

Fourier-Bessel and Fourier-Legendre Series

The procedure *FourierBesselCoeff*$(f(x), x, n, l, m)$ finds the Fourier coefficients $b_{n,m}$ of *Fourier-Bessel series* of the function $f(x)$ associated with Bessel functions of order n in the interval $0 < x \leq l$. The eigenvalues are related to the zeros of the Bessel functions as in (4.3.59). We consider only the case where the function vanishes at $x = l$ and is bounded at $x = 0$. The output of the procedure is

$$b_{n,m} = 2\,\frac{\int_0^l x f(x)\,BesselJ\,(n, BesselJZeros\,(n,m)\,x/l)\,dx}{l^2\,BesselJ\,(1+n,\,BesselJZeros\,(n,m))^2}, \tag{4.8.15}$$

where $BesselJ(n, x)$ is the Bessel function $J_n(x)$ of order n in Maple notation. (If a sixth argument *numeric* is added, the integral is evaluated numerically.) The expression $BesselJZeros(n, m)$ represents the mth positive zero of the Bessel function $J_n(x)$. It is given as α_{mn} in (4.3.59). [The coefficient $b_{n,m}$ agrees with (4.3.62).] $FourierBesselSeries(f(x), r, n, N, l)$ yields a finite (N-term) or infinite Fourier-Bessel series that corresponds to the case above.

As an example, we obtain a 10-term Fourier-Bessel series of $f(x) = 1 - x^2$ in terms of the zero order Bessel function $J_0(x)$ over the interval $0 \leq x \leq 1$, evaluated at $x = 0.5$. We observe that $f(x)$ vanishes at $x = 1$, so that we expect good convergence for the series. The procedure is $FourierBesselSeries(1 - x^2, x, 0, 10, 1, BesselJZeros(0, 1..10))$, where the last argument gives the first 10 positive zeros of $J_0(x)$. (This results in the evaluation of the Bessel zeros that occur in the Fourier-Bessel coefficients.) The exact and approximate results are $(1 - x^2)|_{x=0.5} = .75 \approx .7502247424$.

Next we consider *Fourier-Legendre series*. The procedure *FourierLegendreCoeff* $(f(x), x, n)$ finds the nth coefficient L_n in the series. The output is $L_n = (2n+1)/2$ $\int_{-1}^{1} f(x)\, P(n,x)\, dx$, where $P(n,x)$ is the Maple notation for the nth-degree Legendre polynomial $P_n(x)$. The full Fourier-Legendre series is given by $FourierLegendreSeries(f(x), x, infinity)$ as

$$f(x) = \sum_{k=1}^{\infty} \frac{2k+1}{2} \int_{-1}^{1} f(x)P(k,x)\, dx\; P(k,x). \tag{4.8.16}$$

We obtain two terms in Fourier-Legendre series of $\cos(x)$ by invoking $FourierLegendreSeries(\cos(x), x, 2)$. The output is $\cos(x) \approx \sin(1)\, P_0(x) + (-10\sin(1) + 15\cos(1))\, P_2(x) = \sin(1) + (-10\sin(1) + 15\cos(1))\,(-1/2 + 3x^2/2)$. The first terms on the right side of the equality exhibit the Fourier-Legendre coefficients and the Legendre polynomials $P_0(x)$ and $P_2(x)$. The second terms express the result as a second-degree polynomial. $FourierLegendreSeries(\cos(x), x, 2, numeric)$ yields $\cos(x) \approx 0.8415P_0(x) - 0.3102P_2(x) = 0.9966 - 0.4653x^2$ and evaluates the foregoing result numerically.

Finite Fourier Transforms: Eigenfunction Expansions

In Section 4.6 we constructed solutions of initial and boundary value problems as well as boundary value problems for a class of PDEs. The solutions were obtained in the form of eigenfunction expansions $u = \sum_{k=1}^{\infty} N_k M_k$, with M_k as the eigenfunctions. Two procedures have been constructed to generate these expansions. The procedure *FiniteFourTransTerm* finds the coefficients N_k in these expansions and *FiniteFourTransSeries* yields the partial or complete eigenfunction expansion. The PDEs must have two independent variables, and the appropriate eigenfunctions for each problem must be known and prescribed within the procedure. The *finite Fourier transform method* is used to find solutions in all cases.

For a time-dependent PDE such as the *wave* or *heat equation*, the first procedure is $FiniteFourTransTerm(PDE, BCL, BCR, x = 0..l, ICList, t = t_0..\infty, u(x,t), M_k(x))$, which generates specific terms $N_k(t)M_k(x)$ in the eigenfunction expansions of the form $u(x,t) = \sum_{k=1}^{\infty} N_k(t)\, M_k(x)$. The PDE must be given as the first argument in the procedure. The initial condition is given at $t = t_0$ and the x-interval is $0 < x < l$. The terms BCL and BCR represent the boundary conditions at $x = 0$ and $x = l$, respectively. They may be given as $u(x,t) = g(t)$ or $u_x(x,t) = h(t)$, for example. $ICList$ represents a list of initial conditions at $t = t_0$. For the heat equation we have $[u(x,t) = f(x)]$, while for the wave equation we must enter $[u(x,t) = f(x), u_t(x,t) = g(x)]$. The eigenfunction $M_k(x)$ must be given. (It need not be normalized.) Then the procedure determines $N_k(t)$, if possible, and displays $N_k(t)M_k(x)$. If an arbitrary ninth argument is added in the procedure, the full initial and boundary value problem is displayed as well as the equation for $N_k(t)$ and its solution.

For example, $FiniteFourTransTerm(u_t(x,t) - u_{xx}(x,t) = x^2, u(x,t) = 0, u(x,t) = 0, x = 0..\pi, [u(x,t) = x], t = 0..\infty, u(x,t), \sin(2x), full)$ yields the

output $u_t(x,t)-u_{xx}(x,t)=x^2$, $x=0..\pi$, $t=0..\infty, u(x,t)=x$, $t=0$, $u(x,t)=0$, $x=0$, $u(x,t)=0$, $x=\pi$. $Eigenfunction=\sin(2x)$, $x=0..\pi$. $(\pi/2)N'(t)+2\pi N(t)=-\pi^3/4$, $N(0)=-\pi/2$, and the term $(-\pi/4+\pi e^{-4t}/4-e^{-4t})\sin(2x)$. The eigenvalue problem is $v''(x)+\lambda v(x)=0$, $v(0)=0$, $v(\pi)=0$. The (unnormalized) eigenfunctions are $\sin(kx)$, $k=1,2,\ldots$. The eigenfunction $M_2(x)=\sin(2x)$ is selected for this problem. The last part of the output represents $N_2(t)M_2(x)$. [On removing the last argument $full$ in the procedure only $N_2(t)M_2(x)$ is displayed.]

For a time-independent PDE such as Poisson's equation, the first procedure is $FiniteFourTransTerm(PDE, BCL, BCR, x=0..l, BVList, y=a..b, u(x,y), M_k(x))$, which generates specific terms $N_k(y)M_k(x)$ in the eigenfunction expansions of the form $u(x,y)=\sum_{k=1}^{\infty}N_k(y)M_k(x)$. The only change from the time-dependent case is that $BVList$ is a list of boundary conditions at $y=a$ and $y=b$. If Dirichlet conditions are assigned, we enter $[u(x,y)=g(x), u(x,y)=h(x)]$ with $g(x)$ and $h(x)$ as the values at $y=a$ and $y=b$, respectively.

As an example, we consider the procedure $FiniteFourTransTerm(u_{xx}+u_{yy}=-2\pi^2\sin(\pi x)\sin(\pi y), u(x,y)=0, u(x,y)=0, x=0..1, [u(x,y)=0, u(x,y)=0], y=0..1, u(x,y), \sin(k\pi x))$. It finds the term $N_k(y)M_k(x)$ for the given Poisson equation in the unit square with $u(x,y)=0$ on the boundary. The eigenvalue problem for this case is $v''(x)+\lambda v(x)=0$, $v(0)=0$, $v(1)=0$. The (unnormalized) eigenfunctions are $M_k(x)=\sin(\pi kx)$, $k=1,2,\ldots$.The output gives

$$N_k(y)M_k(x)=\frac{-4\sin(\pi y)\sin(k\pi)\,k\sin(\pi kx)}{(k^4-1)(-\cos(k\pi)\sin(k\pi)+k\pi)}. \tag{4.8.17}$$

Since k is a positive integer, we conclude that $N_k(y)M_k(x)=0$ for all $k>1$. Thus, $N_1(y)M_1(x)$ must give the exact solution of the problem. However, when we put $k=1$ in (4.4.11), we have an indeterminate form. Taking the limit as k tends to zero yields $u(x,y)=\sin(\pi x)\sin(\pi y)$. This is the exact solution of the given problem, as is immediately seen. It can be obtained directly by reapplying the procedure and entering $\sin(\pi x)$ as the last argument.

The second procedure $FiniteFourTransSeries$ finds partial or complete *eigenfunction expansions*. The first seven arguments in the procedure are identical to those in $FiniteFourTransTerm$. If the eighth argument is $M_k(x)$, as before, the full eigenfunction expansion is found, but if it is $[seq(M_k(x), k=1..N)]$, only N terms in the expansion are displayed. [On invoking the procedure, if k is assumed to be an integer (using Maple's *assume* facility), an incorrect result may be obtained, say, when $k=0$ yields an eigenvalue. Then, it is recommended that no assumption be made about k initially, and only after a series is obtained should the assumption that k is an integer be introduced. In some problems, however, unless k is assumed to be an integer, a messy result is obtained. Then $FiniteFourTransTerm$ should be used to get the correct term for $k=0$, and the full series should be adjusted accordingly. In general, the output of the procedure indicates which values of k can lead to problems.] To check the full formulation of the problem, $FiniteFourTransTerm$ can be invoked.

We reconsider the boundary value problem for Poisson's equation given above, but now ask for the complete eigenfunction expansion. On invoking $FiniteFourTrans\,Series(u_{xx} + u_{yy} = -2\,\pi^2 \sin(\pi x) \sin \pi y), u(x, y) = 0, u(x, y) = 0, x = 0..1, [u(x, y) = 0, u(x, y) = 0], y = 0..1, u(x, y), \sin(k\pi x))$, we obtain

$$u(x,y) = \sum_{k=1}^{\infty} -4\,\frac{\sin(\pi y)\sin(k\pi)k\sin(\pi k x)}{(k^4-1)\,(-\cos(k\pi)\sin(k\pi)+k\pi)}. \tag{4.8.18}$$

Clearly, all the terms agree with those found using $FiniteFourTransTerm$. If we assume that k is an integer, the output is $u(x, y) = \sum_{k=1}^{\infty} 0$, which is incorrect. Now, part of the output of $FiniteFourTransSeries$ for this problem identifies $k = 1$ as a singular point for the corresponding term in the series, putting the result for that case in question. All the other terms in the series vanish. Then, to get the correct value for the $k = 1$ term, we must use $FiniteFourTransTerm$, as was done above.

Finally, we consider an initial and Neumann boundary value problem for the wave equation. We use $FiniteFourTransSeries(u_{tt}(x, t) - u_{xx}(x, t) = 0, u_x(x, t) = 0, u_x(x, t) = 1, x = 0..l, [u(x, t) = 0, u_t(x, t) = 1], t = 0..\infty, u(x, t), \cos(k\pi\, x/l))$. The eigenvalue problem for this case is $v''(x) + \lambda v(x) = 0,\ v'(0) = 0,\ v'(l) = 0$. The (unnormalized) eigenfunctions are $M_k(x) = \cos(\pi k x/l),\ k = 0, 1, 2, \ldots$. The eigenfunction expansion is given as

$$u\,(x,t) = t + \frac{t^2}{l} - \frac{2l}{\pi^2 k^2}\sum_{k=1}^{\infty}\left(\cos\left(\frac{k\pi t}{l}\right) - 1\right)\left(1 + (-1)^k\right)\cos\left(\frac{k\pi x}{l}\right). \tag{4.8.19}$$

Stationary Inhomogeneities and Modified Eigenfunction Expansions

We have seen in Section 4.6 that for time-dependent problems, if the coefficients and the nonhomogeneous terms in the PDE and the boundary conditions are stationary (i.e., time independent), it is possible to separate out the effect of these terms and construct a new problem in which the PDE and the boundary conditions are homogeneous. The procedure $SteadyStateSol$ deals with this case. A steady-state solution is obtained from the solution $N(x)$ of the corresponding time-independent boundary value problem as given in (4.6.35). [Here we put $v(x) = N(x)$.] (The procedure works for one space dimension.) If we express the solution of the given problem as $u(x, t) = N(x) + w(x, t)$, in the manner of (4.6.36), then $w(x, t)$ satisfies a homogeneous PDE with homogeneous boundary conditions. The eigenfunction expansion for $w(x, t)$ has better convergence properties than that for the given problem. [We note that the steady-state problem may not have a solution if $u(x, t)$ does not approach a steady state.]

If the boundary conditions for the problem are not time independent, then the procedure *SteadyStateSol* applied to the homogeneous version of the PDE with the given (time-dependent) boundary conditions determines a function $N(x, t)$ that satisfies the boundary conditions. Proceeding as above [and replacing $N(x)$ by $N(x, t)$] now yields a problem for $w(x, t)$ with an inhomogeneous PDE (in general) but with

homogeneous boundary conditions. This again yields an improved eigenfunction expansion of the solution. [The choice of a function with the properties of $N(x,t)$ for a given problem is not unique, and simpler expressions may be found than those given by the procedure.]

The procedure is $SteadyStateSol(PDE, BCL, BCR, x = 0..l, u(x,t))$, with the arguments defined as before. We consider an example. The procedure $SteadyStateSol(u_t(x,t) - c^2\, u_{xx}(x,t) = 1, u(x,t) = 2, u(x,t) = 1, x = 0..l, u(x,t))$ exhibits the steady state problem as part of its output:

$$\text{Steady State: } -c^2\, N''(x) = 1,\ x = 0..l,\ N(0) = 2,\ N(l) = 1, \tag{4.8.20}$$

and obtains the steady-state solution

$$N(x) = -\frac{x^2}{2c^2} - \frac{(-l^2 + 2c^2)x}{2lc^2} + 2. \tag{4.8.21}$$

If the constant boundary condition at $x = 0$ is replaced by the boundary condition $u(0,t) = t$, there is no steady state solution. Then the output of the procedure $SteadyStateSol(u_t(x,t) - c^2\, u_{xx}(x,t) = 1, u(x,t) = t, u(x,t) = 1, x = 0..l, u(x,t))$ is

$$N(x,t) = -\frac{x^2}{2c^2} - \frac{(-l^2 + 2\,tc^2 - 2\,c^2)x}{2lc^2} + t. \tag{4.8.22}$$

We have $N(0,t) = t$ and $N(l,t) = 1$. The problem for $w(x,t)$ determined from $u(x,t) = N(x,t) + w(x,t)$ satisfies homogeneous boundary conditions, but the PDE satisfied by $w(x,t)$ is nonhomogeneous. Nevertheless, the finite Fourier transform eigenfunction expansion converges more rapidly than that for $u(x,t)$ because the boundary conditions are homogeneous.

Having determined $N(x)$ or $N(x,t)$ for a given initial and boundary value problem for a PDE, the procedure $FiniteFourTransSeriesMod$ makes use of the $SteadyStateSol$ result to convert the problem to one with homogeneous boundary conditions. (If the given problem has a steady state, the PDE for the new problem will be homogeneous.) This should accelerate the convergence rate of the series. The procedure retains all eight arguments of $FiniteFourTransSeries$, but the result $N(x)$ or $N(x,t)$ obtained from $SteadyStateSol$ must be inserted between the seventh and eighth arguments of $FiniteFourTransSeries$. (Then there are nine arguments in the procedure.) The procedure carries out the necessary modifications of the given problem and displays the full solution in the form $u = N + w$.

To illustrate the use of the foregoing procedures, we consider an initial and boundary value problem for the nonhomogeneous heat equation

$$u_t(x,t) - 4\, u_{xx}(x,t) = -1,\ u(x,0) = 0,\ u(0,t) = 0,\ u(\pi,t) = 1. \tag{4.8.23}$$

Solving with $FiniteFourTransSeries$, we obtain for $u = u(x,t)$

$$u = \sum_{k=1}^{\infty} \frac{\left(\left(1 + 4\,(k^2 - 1)\,(-1)^k\right) e^{-4\,k^2 t} + 4\,(-1)^{1+k}\,k^2 + (-1)^k - 1\right)\sin(kx)}{2\,k^3\pi}. \tag{4.8.24}$$

The output of $SteadyStateSol$ is $N(x) = x^2/8 - ((\pi^2 - 8)x)/8\pi$. Making use of $N(x)$, we apply $FiniteFourTransSeriesMod$ and obtain

$$u(x,t) = \frac{x^2}{8} - \frac{(\pi^2 - 8)\,x}{8\,\pi} + \sum_{k=1}^{\infty} \frac{\left(1 + (4\,k^2 - 1)\,(-1)^k\right) e^{-4\,k^2 t} \sin(kx)}{2\,k^3 \pi}. \tag{4.8.25}$$

We observe that the terms in the series given by $FiniteFourTransSeries$ decay like $1/k$, but those in the series given by $FiniteFourTransSeriesMod$ decay at an exponential rate.

Exercises 4.8

4.8.1. Apply the Maple procedure *pdsolve* to the wave, heat, and Laplace's equations, with and without the $Hint$ that requests separated solutions, and obtain the results presented in the text.

4.8.2. Use the procedure $EVPProbODE$ to obtain the results given in (4.8.1).

4.8.3. Find the squared norm of the eigenfunctions determined in Exercise 4.8.2 using the Maple procedure *int*.

4.8.4. Apply $EVPProbODE$ to obtain the eigenvalue equation (4.8.2) and then invoke the procedure $Evalues$ to determine the first 10 eigenvalues.

4.8.5. Reproduce the results of Example 4.10.

4.8.6. Reproduce the results of Example 4.11.

4.8.7. Apply the procedures that determine Fourier, Fourier sine, and Fourier cosine coefficients and series to obtain the results given in the subsection on **Trigonometric Fourier Series** in the book.

4.8.8. Obtain 10 terms of the Fourier-Bessel series of $f(x) = 1 - x^2$ over the interval $0 \le x \le 1$ using the procedure $FourierBesselSeries$.

4.8.9. Invoke the procedure $FourierLegendreSeries$ to obtain five terms in the Fourier-Legendre expansion of $\cos(x)$.

4.8.10. Apply $FiniteFourTransTerm(u_t(x,t) - u_{xx}(x,t) = x^2, u(x,t) = 0, u(x,t) = 0, x = 0..\pi, [u(x,t) = x], t = 0..\infty, u(x,t), \sin(2x), full)$ to generate the output given in the text.

4.8.11. Obtain (4.8.17) by using the procedure $FiniteFourTransTerm$.

4.8.12. Obtain (4.8.18) by using the procedure $FiniteFourTransSeries$.

4.8.13. Obtain the results (4.8.24)–(4.8.25) by invoking $SteadyStateSol$, $FiniteFourTransSeries$, and $FiniteFourTransSeriesMod$.

CHAPTER 5

INTEGRAL TRANSFORMS

5.1 INTRODUCTION

In Chapter 4 we dealt with initial and boundary value problems for PDEs given over bounded spatial regions. The method of separation of variables and the closely related finite Fourier transform method were used to obtain solutions of these problems. This chapter deals for the most part with PDEs defined over unbounded spatial regions. The tools we use for solving Cauchy and initial and boundary value problems are *integral transforms*. Specifically, we consider the *Fourier transform*, the *Fourier sine and cosine transforms*, the *Hankel transform*, and the *Laplace transform*.

Instead of proceeding directly to a discussion of each of these transforms, as will be done in the following sections, we begin by showing how the separation of variables method motivates some of the results and suggests which transforms are appropriate for given problems.

As was done Chapter 4, we restrict our discussion to the second order equations (4.2.1), (4.2.5) and (4.2.7) of *hyperbolic, parabolic*, and *elliptic types*, respectively. On applying the method of separation of variables to each of these equations, we set

Partial Differential Equations of Appplied Mathematics, Third Edition. By Erich Zauderer

$u(\mathbf{x}, t) = M(\mathbf{x})N(t)$ (with t replaced by y in the elliptic case) and obtain, as before,

$$LM(\mathbf{x}) = \lambda^2 \rho(\mathbf{x}) M(\mathbf{x}), \tag{5.1.1}$$

with the separation constant λ replaced by λ^2 for convenience.

The equations for N are given as

$$\begin{cases} N''(t) + \lambda^2 N(t) = 0, & \text{hyperbolic case,} \\ N'(t) + \lambda^2 N(t) = 0, & \text{parabolic case,} \\ N''(y) - \lambda^2 N(y) = 0, & \text{elliptic case.} \end{cases} \tag{5.1.2}$$

Since we are dealing with unbounded spatial regions, we shall generally require that the solution $M(\mathbf{x})$ of (5.1.1) be bounded or vanish at infinity. If the region is a semi-infinite interval or a half-plane, one of the boundary conditions given in Chapter 4 is applied on the finite portion of the boundary. All boundary conditions are assumed to be of homogeneous type, so that $M(\mathbf{x}) = 0$ satisfies (5.1.1) and the boundary conditions. To obtain nonzero $M(\mathbf{x})$, restrictions must be placed on the parameter λ^2. The values of λ (rather than λ^2) for which nonzero $M(\mathbf{x})$ can be found are the eigenvalues, and the corresponding $M(\mathbf{x})$ are the eigenfunctions for the eigenvalue problem.

The fundamental distinction between the eigenvalue problem for (5.1.1) over an unbounded region and that for a bounded region considered in Chapter 4 is that, in general (and this is the case for the problems considered here), the *spectrum* (i.e., the set of eigenvalues) is *continuous* in the unbounded case, whereas in the bounded case the spectrum is *discrete*, as we have seen in Chapter 4. Consequently, if λ_k $(k = 1, 2, \ldots)$ is the set of eigenvalues in the discrete case and $u_k = M_k N_k$ is the corresponding set of solutions of the given PDE, the general solution is obtained by *superposition* of the u_k as

$$u = \sum_{k=1}^{\infty} u_k = \sum_{k=1}^{\infty} M_k N_k. \tag{5.1.3}$$

In the case of a continuous spectrum, the eigenvalues λ range over the set D, which may be the interval $-\infty < \lambda < \infty$, $0 \leq \lambda < \infty$, or some other uncountable set. Let $u_\lambda = M_\lambda N_\lambda$ be the separated solution, corresponding to λ, of the given partial differential equation. We then obtain a general solution by the formal *superposition*

$$u = \int_D u_\lambda \, d\lambda = \int_D M_\lambda N_\lambda \, d\lambda. \tag{5.1.4}$$

Since the spectrum is continuous and an orthogonality property for the eigenfunctions M_λ does not occur in a simple and natural way (as for the discrete spectrum), the specification of the N_λ in (5.1.4) in terms of initial and boundary data cannot be carried out in a general way as was done in Chapter 4.

Therefore, we concentrate on a number of specific eigenvalue problems for (5.1.1) in this chapter that lead to the consideration of *Fourier*, *Hankel*, and *Laplace transforms*. (In this respect our discussion parallels that given in Section 4.3 for the

Sturm-Liouville problem.) Only the most basic properties of each of the transforms are presented and their use is demonstrated in a number of examples. Once it is determined that a particular transform is relevant for the solution of a given problem, the solution is obtained by transforming the given equation and solving for the transform function. (This approach is identical to that used in the finite Fourier transform method of Chapter 4.) Simple and useful expressions for the solution do not often result once its transform is known, so approximate methods for evaluating Laplace and Fourier transforms and integrals are presented at the end of this chapter.

5.2 ONE-DIMENSIONAL FOURIER TRANSFORMS

We consider the one-dimensional form of (5.1.1) with $\rho(x) = p(x) = 1$ and $q(x) = 0$, given over the infinite interval $-\infty < x < \infty$; that is,

$$M''(x) + \lambda^2 M(x) = 0, \qquad -\infty < x < \infty, \tag{5.2.1}$$

with the auxiliary condition that $M(x)$ remain bounded as $|x| \to \infty$. The general solution of (5.2.1) in complex form is

$$M_\lambda(x) = \alpha(\lambda)e^{i\lambda x} + \beta(\lambda)e^{-i\lambda x}, \tag{5.2.2}$$

where $\alpha(\lambda)$ and $\beta(\lambda)$ are constants. The boundedness condition restricts the eigenvalue parameter λ to be real valued. Thus the eigenvalues λ lie in the interval $-\infty < \lambda < \infty$, so that the spectrum is continuous.

Corresponding to the eigenfunction expansions given in Chapter 4, we would like to be able to represent an arbitrary function $f(x)$ (under suitable conditions) in terms of the eigenfunctions $M_\lambda(x)$ in the form

$$f(x) = \int_{-\infty}^{\infty} \left(\alpha(\lambda)e^{i\lambda x} + \beta(\lambda)e^{-i\lambda x}\right)\, d\lambda = \int_{-\infty}^{\infty} \gamma(\lambda)e^{-i\lambda x}\, d\lambda, \tag{5.2.3}$$

where $\gamma(\lambda) = \alpha(-\lambda) + \beta(\lambda)$ [the integral in $\gamma(\lambda)$ is obtained from the first integral via a simple change of variables]. Since the eigenfunctions for this eigenvalue problem are not orthogonal (in an elementary sense), it is not immediately apparent how to determine $\gamma(\lambda)$ in terms of $f(x)$ in (5.2.3). However, a generalized *orthogonality property* for the eigenfunctions $M_\lambda(x)$ is determined in Section 7.2 when Dirac delta functions are discussed. It is shown that $(1/2\pi)\int_{-\infty}^{\infty} e^{-i\lambda x}e^{i\hat{\lambda}x}\, dx = \delta(\lambda - \hat{\lambda})$. Thus, if we multiply across in (5.2.3) by $(1/2\pi)e^{i\hat{\lambda}x}$, integrate with respect to x from $-\infty$ to ∞, and assume that it is valid to interchange the order of integration, we conclude that

$$\gamma(\lambda) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\lambda x} f(x)\, dx, \tag{5.2.4}$$

on using the substitution property of the Dirac delta function $\delta(z)$ and replacing $\hat{\lambda}$ by λ.

An alternative (formal) method for determining $\gamma(\lambda)$ in (5.2.3) can be based on the expansion (4.3.52) for the Fourier series of the function $v(x)$ given in Section 4.3. On expressing the trigonometric functions in complex form and letting $l \to \infty$, it can be shown (see the exercises) that a plausible result is the *Fourier integral formula*

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-i\lambda(x-t)} f(t)\, dt\, d\lambda, \tag{5.2.5}$$

where we have used $f(x)$ in place of $v(x)$. If we define the function $F(\lambda)$ in terms of the inner integral in (5.2.5) to be

$$F(\lambda) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{i\lambda x} f(x)\, dx, \tag{5.2.6}$$

(where we have changed the variable of integration from t to x), we obtain

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\lambda x} F(\lambda)\, d\lambda. \tag{5.2.7}$$

The coefficients in (5.2.6)–(5.2.7) have been chosen to be $1/\sqrt{2\pi}$ for the purpose of achieving symmetry in the formulas for $F(\lambda)$ and $f(x)$.

The function $F(\lambda)$ in (5.2.6) is called the *Fourier transform* of $f(x)$ and (5.2.7) is the *Fourier inversion formula* that gives the function $f(x)$ in terms of its transform $F(\lambda)$. In this context, the function $f(x)$ is referred to as the *inverse Fourier transform* of $F(\lambda)$. By comparing (5.2.7) with (5.2.4), we conclude that $\gamma(\lambda) = (1/\sqrt{2\pi})F(\lambda)$. (We remark that various alternative definitions of the Fourier transform and its inversion formula are given in the literature. In particular, Maple uses a different definition, as we will see. As a result, transforms of specific functions and certain formulas may assume different forms, but in their use for solving problems for differential equations, the solutions of the problems must be the same.)

General Properties

Once the Fourier integral formula has been obtained, sufficient conditions for its validity can be verified directly. Thus if $f(x)$ is piecewise continuously differentiable on each finite interval and $\int_{-\infty}^{\infty} |f(x)|\, dx < \infty$, it can be shown that the integral in (5.2.5) converges pointwise to the function $f(x)$ at points of continuity of $f(x)$. At a point x_0 where $f(x)$ has a jump discontinuity, the integral converges to $\frac{1}{2}f(x_0-) + \frac{1}{2}f(x_0+)$, the average of the limit values of $f(x)$. If we weaken the conditions on $f(x)$ and merely require that it be square integrable over $-\infty < x < \infty$, so that $\int_{-\infty}^{\infty} |f(x)|^2\, dx < \infty$, it can be shown that the integral converges to $f(x)$ in the mean square sense.

In applications of Fourier transforms it is often necessary to find the inverse transform of $H(\lambda) = F(\lambda)G(\lambda)$, where $F(\lambda)$ and $G(\lambda)$ are transforms of the known functions $f(x)$ and $g(x)$, respectively. The inverse transform of $H(\lambda)$ is given as

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-i\lambda x}F(\lambda)G(\lambda)\,d\lambda = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-i\lambda(x-t)}f(t)G(\lambda)\,dt\,d\lambda,$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} f(t)\int_{-\infty}^{\infty} e^{-i\lambda(x-t)}G(\lambda)\,d\lambda\,dt = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(t)g(x-t)\,dt, \quad (5.2.8)$$

where we have assumed that the interchange of orders of integration is valid. The last integral in (5.2.8) represents the *convolution* of $f(x)$ and $g(x)$. We have shown that the Fourier transform of the convolution of $f(x)$ and $g(x)$ equals the product of the Fourier transforms of $f(x)$ and $g(x)$. This result is known as the *convolution integral theorem* for Fourier transforms.

If we put $G(\lambda) = \overline{F(\lambda)}$, the complex conjugate of $F(\lambda)$, in (5.2.8), it is easily shown that we obtain the *Parseval equation* for Fourier transforms:

$$\int_{-\infty}^{\infty} |F(\lambda)|^2\,d\lambda = \int_{-\infty}^{\infty} |f(x)|^2\,dx. \quad (5.2.9)$$

This should be compared with the Parseval equality (4.3.14) for Fourier series.

A fundamental result for the application of Fourier transforms to differential equations relates the Fourier transform of derivatives of functions to the transform of the functions themselves. For the transform of $f'(x)$ we have

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{i\lambda x}f'(x)\,dx = \frac{-i\lambda}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{i\lambda x}f(x)\,dx = -i\lambda F(\lambda), \quad (5.2.10)$$

where $F(\lambda)$ is the Fourier transform of $f(x)$, assuming that $f(x)$ is a smooth function that vanishes at infinity. More generally, if $f^{(n)}(x)$ is the nth derivative of $f(x)$, and $f(x)$ and its first $n-1$ derivatives are smooth functions that vanish at infinity, we have

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{i\lambda x}f^{(n)}(x)\,dx = -(i\lambda)^n F(\lambda) \qquad n = 1, 2, 3, \ldots. \quad (5.2.11)$$

The formulas (5.2.10)–(5.2.11) are obtained on integrating by parts.

If $f(x)$ is absolutely integrable, [i.e., $\int_{-\infty}^{\infty} |f(x)|\,dx < \infty$], the *Riemann-Lebesgue lemma* asserts that its Fourier transform $F(\lambda)$ tends to zero as $|\lambda|$ tends to infinity. (See the exercises for a proof.) The rate at which $F(\lambda) \to 0$ depends on the smoothness of $f(x)$ and on how rapidly it decays as $|x| \to \infty$.

Further properties of Fourier transforms and transforms of specific functions are given in the exercises, in Section 5.8, and in tables of transforms.

Applications to ODEs and PDEs

We now consider several problems for ODEs and PDEs whose solutions are found by using Fourier transforms. In each case one of the independent variables has the infinite interval $(-\infty, \infty)$ as its domain of definition.

Example 5.1. An Ordinary Differential Equation. An example that exhibits the basic features of the Fourier transform method for solving differential equations is given by the following problem for the ODE

$$y''(x) - k^2 y(x) = -f(x), \qquad -\infty < x < \infty, \tag{5.2.12}$$

where k is a constant and $f(x)$ is prescribed. We require that

$$y(x),\ y'(x) \to 0, \quad \text{as} \quad |x| \to \infty, \tag{5.2.13}$$

and that $f(x)$ has a Fourier transform $F(\lambda)$. Let

$$Y(\lambda) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{i\lambda x} y(x)\, dx, \tag{5.2.14}$$

so that $Y(\lambda)$ is the Fourier transform of the solution $y(x)$.

To solve this boundary value problem, we (Fourier) transform the equation (5.2.12) by multiplying across by $(1/\sqrt{2\pi})\ e^{i\lambda x}$ and integrating both sides with respect to x from $-\infty$ to $+\infty$. Our assumptions (5.2.13) imply that on using (5.2.11), $(1/\sqrt{2\pi}) \int_{-\infty}^{\infty} e^{i\lambda x} y''(x)\, dx = (-i\lambda)^2 Y(\lambda) = -\lambda^2 Y(\lambda)$. Thus, the equation for $Y(\lambda)$ is algebraic and we obtain, with $G(\lambda) = 1/(\lambda^2 + k^2)$,

$$-(\lambda^2 + k^2) Y(\lambda) = -F(\lambda), \quad Y(\lambda) = \frac{F(\lambda)}{\lambda^2 + k^2} \equiv F(\lambda) G(\lambda). \tag{5.2.15}$$

To find $y(x)$, we must invert $Y(\lambda)$. This can be done using the convolution theorem (5.2.8) if the inverse transform of $G(\lambda)$ is known. Using a table of transforms or complex integration theory, it can be shown that

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\lambda x} G(\lambda)\, d\lambda = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{e^{-i\lambda x}}{\lambda^2 + k^2}\, d\lambda = \frac{\sqrt{2\pi}}{2k} e^{-k|x|}, \tag{5.2.16}$$

with $k > 0$. [Note that once the inverse transform of $G(\lambda)$ is given, as in (5.2.16), the Fourier transform of that function is easy to evaluate and thereby shown to equal $G(\lambda)$.] Using (5.2.16) in the convolution theorem, with $Y(\lambda) = F(\lambda)G(\lambda)$, we obtain the solution of (5.2.12)–(5.2.13) in the form

$$y(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\lambda x} F(\lambda) G(\lambda)\, d\lambda = \frac{1}{2k} \int_{-\infty}^{\infty} e^{-k|x-t|} f(t)\, dt. \tag{5.2.17}$$

It is of interest to apply the solution formula (5.2.17) for two special choices of $f(x)$. First we set $f(x) = 1$ and then we set $f(x) = \delta(x - \xi)$, where $\delta(x - \xi)$ is the Dirac delta function. With $f(x) = 1$ we obtain

$$y(x) = \frac{1}{2k} \int_{-\infty}^{x} e^{-k(x-t)}\, dt + \frac{1}{2k} \int_{x}^{\infty} e^{k(x-t)}\, dt = \frac{1}{k^2}. \tag{5.2.18}$$

Putting $f(x) = \delta(x - \xi)$ and recalling the properties of the delta function given in Example 1.2 gives

$$y(x) = \frac{1}{2k} \int_{-\infty}^{\infty} e^{-k|x-t|} \delta(t - \xi)\, dt = \frac{1}{2k}\, e^{-k|x-\xi|}. \tag{5.2.19}$$

Each of these *solutions* has a shortcoming. The function $y(x) = 1/k^2$ satisfies (5.2.12) but fails to vanish at infinity, as required by the boundary condition (5.2.13). The function $y(x) = (1/2k)\, e^{-k|x-\xi|}$ satisfies (5.2.12) at all $x \neq \xi$ and vanishes at infinity, but is not differentiable at $x = \xi$.

We do not expect the *solutions* above to satisfy all the conditions of the boundary value problem (5.2.12)–(5.2.13) since, in both cases, the Fourier transforms of the functions $f(x)$ are not defined in a conventional sense. We proceed formally. The function $f(x) = 1$ has a Fourier transform equal to $\sqrt{2\pi}\ \delta(\lambda)$. This can be seen by putting $F(\lambda) = \sqrt{2\pi}\ \delta(\lambda)$ in the inversion formula (5.2.7). Similarly, the Fourier transform of $f(x) = \delta(x - \xi)$ is $F(\lambda) = (1/\sqrt{2\pi})e^{i\lambda\xi}$. In either case we are required to deal with *generalized functions* (specifically, Dirac delta functions) and without extending the theory of Fourier transforms to include such functions (as is done in Section 7.2) and an appropriate formulation of the corresponding boundary value problems, it is not clear that the solution formula can be applied in such cases. Nevertheless, the solution $y(x) = 1/k^2$ in the case where $f(x) = 1$ is certainly plausible (even if it does not vanish at infinity), and the solution $y(x) = (1/2k)e^{-k|x-\xi|}$ when $f(x) = \delta(x - \xi)$ is known as the *Green's function* for (5.2.12)–(5.2.13). (The importance of Green's functions in solving problems for both ODEs and PDEs is demonstrated in Chapter 7.)

Although a direct approach can be developed to deal with each of the foregoing special cases which may or may not depend on Fourier transform theory, we now show how each of the solutions can be obtained as limits of sequences of solutions of the boundary value problem (5.2.12)–(5.2.13), each of which satisfies the conditions of the problem. As such, the two solutions obtained previously are called *generalized solutions*.

We define the sequence of functions

$$f_N(x) = \begin{cases} \alpha(N), & |x| \leq N, \\ 0, & |x| > N, \end{cases} \tag{5.2.20}$$

where $N > 0$, and consider two cases. First, we set $\alpha = 1$ for all N and find that $\lim_{N\to\infty} f_N(x) = 1$ for all x. This corresponds to the case $f(x) = 1$ given above. Second, we set $\alpha(N) = 1/2N$ and obtain, in accordance with our discussion in Example 1.2, $\lim_{N\to 0} f_N(x) = \delta(x)$. This corresponds to the case $f(x) = \delta(x)$, where we have put $\xi = 0$ for simplicity.

Clearly, each of the $f_N(x)$ has a Fourier transform, and the problem (5.2.12)–(5.2.13) may be solved in the manner given above with $f(x) = f_N(x)$. The solution

$y_N(x)$ corresponding to $f_N(x)$ is found to be

$$y_N(x) = \frac{\alpha(N)}{2k}\int_{-N}^{N} e^{-k|x-t|}\,dt = \begin{cases} [1 - e^{-kN}\cosh(kx)]\alpha(N)/k^2, & |x| \le N, \\ [e^{-k|x|}\sinh(kN)]\alpha(N)/k^2, & |x| > N. \end{cases} \tag{5.2.21}$$

We have $\lim_{|x|\to\infty} y_N(x) = 0$ and $\lim_{|x|\to\infty} y_N'(x) = 0$, so (5.2.13) is satisfied.

First we put $\alpha(N) = 1$ in (5.2.21). Then as $N \to \infty$, the pointwise limit of $y_N(x)$ is clearly $y_N(x) \to 1/k^2$ for all finite x. Next we set $\alpha(N) = 1/2N$ and the limit as $N \to 0$ is found to be, on using l'Hospital's rule, $y_N(x) \to (1/2k)e^{-k|x|}$. (This is the Green's function with $\xi = 0$.) The limits are nonuniform over the infinite interval, so the smoothness properties and the vanishing at infinity are not preserved in both cases, yet we have obtained these solutions as limits of strict solutions of the boundary value problem. It may be stated as a general principle that solutions of a given problem obtained by formal and plausible means can be characterized as limits obtained in an appropriate fashion of a sequence of strict solutions of the problem. We shall have further occasion to construct solutions of problems by formal means, but will not always show how they may be obtained as a limit of strict solutions. In Chapter 4, the eigenfunction expansions did not always constitute strict solutions of the given problem and had to be interpreted as generalized solutions of the problem even though this was not always indicated explicitly.

On replacing the differential equation (5.2.12) by

$$y''(x) + k^2 y(x) = -f(x), \qquad -\infty < x < \infty, \tag{5.2.22}$$

we can no longer require that $y(x)$ and $y'(x)$ tend to zero as $|x| \to \infty$, and use the Fourier transform as was done for (5.2.12). Indeed, if $f(x) = 0$ for $|x| > N$, the general solution of (5.2.22) for $|x| > N$ is a linear combination of $\sin(kx)$ and $\cos(kx)$. While it is bounded at infinity, it cannot be made to vanish there unless it is identically zero. Further if we formally apply the Fourier transform to (5.2.22) and neglect contributions from infinity, we obtain

$$Y(\lambda) = \frac{F(\lambda)}{\lambda^2 - k^2}, \tag{5.2.23}$$

and the inverse transform $y(x)$, [in terms of (5.2.7)] does not exist for arbitrary $f(x)$ in view of the singularity of $Y(\lambda)$ at $\lambda = \pm k$.

To proceed with the solution of (5.2.22), we adapt the solution formula (5.2.17) to fit the present problem. [This parallels the approach used in the two special cases considered for (5.2.12).] We observe that if we replace k by ik or by $-ik$ in (5.2.12), the equation takes the form (5.2.22). Thus, if we replace k by $\pm ik$ in (5.2.17) we must obtain a solution of (5.2.22) if $f(x)$ is a function for which the resulting integral converges. This gives

$$y_{\pm}(x) = \pm\frac{i}{2k}\int_{-\infty}^{\infty} e^{\pm ik|x-t|} f(t)\,dt, \tag{5.2.24}$$

where $y_{\pm}(x)$ correspond to the choice $k \to \mp ik$, respectively. Both $y_+(x)$ and $y_-(x)$ are solutions of (5.2.22) and are bounded at infinity. The solution cannot be required to vanish at infinity to obtain a unique result, as was the case for (5.2.12). Furthermore, asking that the solution be bounded does not yield uniqueness, as shown by the two solutions given in (5.2.24).

If (5.2.22) arises in the context of wave propagation as a reduced (inhomogeneous) wave equation (see Example 6.13), the *radiation condition* at infinity,

$$\lim_{|x|\to\infty} \left[\frac{\partial y(x)}{\partial |x|} - iky(x) \right] = 0, \tag{5.2.25}$$

picks out the solution $y_+(x)$ in (5.2.24) and guarantees a unique solution for the problem. If $f(x)$ vanishes outside a finite interval, the radiation condition (5.2.25) signifies that a wave traveling away from the source toward infinity is generated. This is represented by the solution $y_+(x)$. The solution $y_-(x)$ represents a wave that travels inward from infinity. (This is elaborated on in Example 6.13 and Section 10.1.) The choice $f(x) = \delta(x - \xi)$ yields the *Green's function* $y_+(x) = (i/2k)e^{ik|x-\xi|}$ for the boundary value problem.

Example 5.1 shows the effectiveness of the Fourier transform in solving a boundary value problem for ODEs over an infinite interval. It was seen, however, that modifications of the Fourier transform method, as originally presented, may be required to solve certain problems of great interest.

Example 5.2. The Cauchy Problem for the Heat Equation. We consider the heat (or diffusion) equation

$$u_t(x,t) - c^2 u_{xx}(x,t) = 0, \qquad -\infty < x < \infty,\ t > 0, \tag{5.2.26}$$

where c^2 is a constant, with the initial condition

$$u(x,0) = f(x), \qquad -\infty < x < \infty, \tag{5.2.27}$$

and use a Fourier transform in the x-variable. We assume that $f(x)$ has a Fourier transform and that u and u_x vanish at infinity, so that (5.2.11) is applicable. The solution obtained will be found to be valid under weaker conditions than are necessary for the Fourier transform method to be applicable.

Denoting the Fourier transform of $u(x,t)$ by

$$U(\lambda,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{i\lambda x} u(x,t)\, dx, \tag{5.2.28}$$

we multiply through by $(1/\sqrt{2\pi})e^{i\lambda x}$ in (5.2.26) and integrate with respect to x from $-\infty$ to $+\infty$. Using (5.2.11), we obtain an ODE for $U(\lambda,t)$,

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u_t e^{i\lambda x}\, dx + (c\lambda)^2 U(\lambda,t) = \frac{\partial U(\lambda,t)}{\partial t} + (c\lambda)^2 U(\lambda,t) = 0. \tag{5.2.29}$$

Fourier transforming the initial condition (5.2.27) yields

$$U(\lambda, 0) = F(\lambda) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{i\lambda x} f(x)\, dx. \tag{5.2.30}$$

The solution of the initial value problem (5.2.29)–(5.2.30) is

$$U(\lambda, t) = F(\lambda) e^{-(c\lambda)^2 t}. \tag{5.2.31}$$

To find $u(x, t)$, we must invert the Fourier transform $U(\lambda, t)$. We have

$$u(x, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-i\lambda(x-s) - \lambda^2 c^2 t} f(s)\, d\lambda\, ds, \tag{5.2.32}$$

where we have used (5.2.30) and interchanged the order of integration. To evaluate the inner integral we break it up into two integrals from $-\infty$ to 0 and 0 to ∞ to obtain

$$\int_{-\infty}^{\infty} e^{-i\lambda(x-s) - \lambda^2 c^2 t}\, d\lambda = 2 \int_{0}^{\infty} e^{-\lambda^2 c^2 t} \cos(\lambda(x - s))\, d\lambda. \tag{5.2.33}$$

To evaluate the second integral in (5.2.33), let the integral $I(\alpha)$ be defined as

$$I(\alpha) = 2 \int_{0}^{\infty} e^{-\lambda^2 c^2 t} \cos(\alpha\lambda)\, d\lambda. \tag{5.2.34}$$

In view of the exponential decay of the integrand (for $t > 0$) we can obtain $dI/d\alpha$ by differentiating under the integral sign. We have, on integrating by parts

$$\frac{dI(\alpha)}{d\alpha} = -2 \int_{0}^{\infty} \lambda e^{-\lambda^2 c^2 t} \sin(\alpha\lambda)\, d\lambda = -\frac{\alpha}{2c^2 t} I(\alpha). \tag{5.2.35}$$

Now $I(0)$ is given as

$$I(0) = 2 \int_{0}^{\infty} e^{-\lambda^2 c^2 t}\, d\lambda = \sqrt{\frac{\pi}{c^2 t}}. \tag{5.2.36}$$

The solution of the initial value problem (5.2.35)–(5.2.36) is found to be

$$I(\alpha)|_{\alpha = x-s} = 2 \int_{0}^{\infty} e^{-\lambda^2 c^2 t} \cos(\lambda(x - s))\, d\lambda = \sqrt{\frac{\pi}{c^2 t}} \exp\left[-\frac{(x-s)^2}{4c^2 t}\right]. \tag{5.2.37}$$

Introducing (5.2.37) into (5.2.32) yields the solution of the Cauchy problem (5.2.26)–(5.2.27) as

$$u(x, t) = \frac{1}{\sqrt{4\pi c^2 t}} \int_{-\infty}^{\infty} \exp\left[-\frac{(x-s)^2}{4c^2 t}\right] f(s)\, ds. \tag{5.2.38}$$

The term

$$G(x - \xi, t) = \frac{1}{\sqrt{4\pi c^2 t}} \exp\left[-\frac{(x-\xi)^2}{4c^2 t}\right] \tag{5.2.39}$$

is known as the *fundamental solution* of the heat or diffusion equation. With the formal substitution $f(s) = \delta(s-\xi)$ in (5.2.38), we obtain the fundamental solution (5.2.39). Thus (5.2.39) is a solution of the heat equation corresponding to a point source of heat at the initial time $t = 0$ located at the point $x = \xi$. The solution (5.2.38) can be thought of as a superposition of solutions due to point sources distributed along the x-axis, with density $f(x)$. The rapid decay at infinity of the exponential term (when $t > 0$) in the integral in (5.2.38) indicates that the restriction $\int_{-\infty}^{\infty} |f(x)|\, dx < \infty$ placed on $f(x)$ to guarantee the existence of its Fourier transform, can be relaxed considerably.

We have already encountered a solution of the form (5.2.39) in Section 1.1. In fact, given the diffusion equation (with drift) for $v(x,t)$,

$$v_t(x,t) + cv_x(x,t) = -\frac{1}{2}Dv_{xx}(x,t), \tag{5.2.40}$$

considered in (1.1.15), if we set $\hat{x} = x - ct$, $\hat{t} = t$, and $u(\hat{x},\hat{t}) = v(x,t)$, then $u(\hat{x},\hat{t})$ satisfies (5.2.26) with c^2 replaced by $D/2$ and (x,t) by $(\hat{x},\hat{t})$. Further, $u(\hat{x},0) = u(x,0) = v(x,0)$. Thus with the initial condition (1.1.18) for $v(x,t)$ [i.e., $v(x,0) = \delta(x)$], we obtain the solution (1.1.20) for $v(x,t)$ on comparing with the fundamental solution (5.2.39).

Either from the solution (5.2.38) with $f(x) = 0$ outside a finite interval or from the fundamental solution (5.2.39), we conclude that $u(x,t)$ is instantaneously greater than zero at each x as soon as the time t increases from zero. [If $u(x,t)$ represents temperature, it is generally assumed that $u(x,0) = f(x) \geq 0$.] Thus heat propagates at infinite speed according to the heat equation (5.2.26). A similar observation was made and discussed in connection with the diffusion equation of Section 1.1.

For the initial value problem for the nonhomogeneous heat equation

$$u_t(x,t) - c^2u_{xx}(x,t) = F(x,t),\ -\infty < x < \infty,\ t > 0, \quad u(x,0) = 0, \tag{5.2.41}$$

we easily obtain, on using Duhamel's principle,

$$u(x,t) = \int_0^t \int_{-\infty}^{\infty} G(x-s, t-\tau)F(s,\tau)\, ds\, d\tau, \tag{5.2.42}$$

where $G(x-s,t-\tau)$ is defined as in (5.2.39). The result (5.2.42) can also be obtained by the use of Fourier transforms. The fundamental solution (5.2.39) is thus seen to play an important role in the solution of the Cauchy problem for the heat equation. As we shall see, it also plays a role in the solution of problems over semi-infinite intervals and even problems over a finite interval.

The function $G(x-\xi,t)$ is also of interest for the reason that when $t \downarrow 0$, $G(x-\xi,t) \to \delta(x-\xi)$. Thus $G(x-\xi,t)$ represents a smooth function (for $t > 0$) which in the limit tends to the Dirac delta function. This is in contrast to the functions $\delta_\epsilon(x)$ defined in Example 1.2, which tend to $\delta(x)$ as $\epsilon \to 0$ but have jump discontinuities at $x = \pm\epsilon$ and are not smooth functions.

Example 5.3. The Cauchy Problem for the Wave Equation. The Cauchy problem for the wave equation asks for a function $u(x,t)$ that satisfies, for $-\infty < x < \infty$,

$$u_{tt}(x,t) - c^2 u_{xx}(x,t) = 0,\ t > 0, \quad u(x,0) = f(x),\ u_t(x,0) = g(x). \quad (5.2.43)$$

The Fourier-transformed initial value problem is

$$\frac{\partial^2 U(\lambda,t)}{\partial t^2} + (c\lambda)^2 U(\lambda,t) = 0,\ t > 0, \quad U(\lambda,0) = F(\lambda),\ \frac{\partial U(\lambda,0)}{\partial t} = G(\lambda), \quad (5.2.44)$$

where U, F, and G are Fourier transforms of u, f, and g, respectively. It is assumed that the operations necessary to obtain (5.2.44) are valid.

The solution of the initial value problem (5.2.44) is easily found to be

$$U(\lambda,t) = \left[\frac{1}{2}F(\lambda) + \frac{1}{2i\lambda c}G(\lambda)\right] e^{i\lambda ct} + \left[\frac{1}{2}F(\lambda) - \frac{1}{2i\lambda c}G(\lambda)\right] e^{-i\lambda ct}. \quad (5.2.45)$$

Inverting this transform gives the solution

$$u(x,t) = \frac{1}{2\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-i\lambda(x-ct)}F(\lambda)\,d\lambda + \frac{1}{2\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-i\lambda(x+ct)}F(\lambda)\,d\lambda$$

$$+\frac{1}{2c\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-i\lambda(x-ct)}\frac{G(\lambda)}{i\lambda}\,d\lambda + \frac{1}{2c\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-i\lambda(x+ct)}\frac{G(\lambda)}{i\lambda}\,d\lambda. \quad (5.2.46)$$

The integrals involving $G(\lambda)$ in (5.2.46) are to be interpreted in the *Cauchy principal value sense* in the neighborhood of $\lambda = 0$.

If we express the function $g(x)$ as

$$g(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-i\lambda x}G(\lambda)\,d\lambda, \quad (5.2.47)$$

the indefinite integral of $g(x)$ is given as

$$\int^{x} g(s)\,ds = -\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-i\lambda x}\frac{G(\lambda)}{i\lambda}\,d\lambda, \quad (5.2.48)$$

with the integral in $G(\lambda)$ interpreted in the Cauchy principal value sense. Then (5.2.46) can be given as

$$u(x,t) = \frac{1}{2}[f(x-ct) + f(x+ct)] + \frac{1}{2c}\int_{x-ct}^{x+ct} g(s)\,ds, \quad (5.2.49)$$

on using (5.2.7) and (5.2.48). The expression (5.2.49) is just *d'Alembert's solution*, which was obtained in Example 2.4 by another method.

Example 5.4. Laplace's Equation in a Half-Plane and Stokes' Rule. We begin by considering *Dirichlet's problem* for Laplace's equation in a half-plane. We require that $u(x, y)$ satisfy the PDE with the boundary condition

$$u_{xx}(x, y) + u_{yy}(x, y) = 0, \ -\infty < x < \infty, \ y > 0, \ u(x, 0) = f(x), \quad (5.2.50)$$

and the additional conditions

$$u(x, y) \to 0 \text{ as } |x| \to \infty, \quad u(x, y) \text{ bounded as } y \to \infty. \quad (5.2.51)$$

To solve (5.2.50)–(5.2.51) we introduce the Fourier transform in x of the function $u(x, y)$ defined as

$$U(\lambda, y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{i\lambda x} u(x, y)\, dx. \quad (5.2.52)$$

Then the Fourier-transformed problem (5.2.50)–(5.2.51) becomes

$$\frac{\partial^2 U(\lambda, y)}{\partial y^2} - \lambda^2 U(\lambda, y) = 0, \qquad 0 < y < \infty, \quad (5.2.53)$$

with the boundary conditions

$$U(\lambda, 0) = F(\lambda), \qquad U(\lambda, y) \quad \text{bounded as} \quad y \to \infty. \quad (5.2.54)$$

Here $F(\lambda)$ is the transform of $f(x)$ and the boundedness condition is a consequence of (5.2.51). The solution of (5.2.53)–(5.2.54) is

$$U(\lambda, y) = F(\lambda) e^{-|\lambda| y}, \quad (5.2.55)$$

where the $|\lambda|$ occurs, since we require that $U(\lambda, y)$ is bounded as $y \to \infty$, and λ ranges from $-\infty$ to $+\infty$. Inverting the transform $U(\lambda, y)$ gives

$$u(x, y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-i\lambda(x-t) - |\lambda| y} f(t)\, d\lambda\, dt \quad (5.2.56)$$

on using the inverse transform for $F(\lambda)$ and inverting the order of integration. By breaking it into two parts, the inner integral can be evaluated as

$$\int_{-\infty}^{\infty} e^{-i\lambda(x-t) - |\lambda| y}\, d\lambda = \frac{2y}{(x-t)^2 + y^2}. \quad (5.2.57)$$

Thus we have

$$u(x, y) = \frac{y}{\pi} \int_{-\infty}^{\infty} \frac{f(t)}{(x-t)^2 + y^2}\, dt. \quad (5.2.58)$$

It may be verified directly that for $y > 0$, (5.2.58) satisfies Laplace's equation. However, a careful examination of the integral is required to show that the boundary

condition in (5.2.50) is satisfied. Although the coefficient y of the integral vanishes at $y = 0$, the integral is singular at the point $t = x$ when $y = 0$. The combination of these two effects leads to the satisfaction of the boundary condition. We do not carry out this verification in the general case.

Instead, we consider a special example with $f(x) = H(x)$, where $H(x)$ is the *Heaviside unit step function* defined as $H(x) = 0$, $x < 0$ and $H(x) = 1$, $x \geq 0$. Inserting $f(t) = H(t)$ into (5.2.58) yields an integral that can be evaluated explicitly in terms of the inverse tangent function. The result is

$$u(x, y) = \frac{\pi + 2\tan^{-1}(x/y)}{2\pi}. \tag{5.2.59}$$

We note that $\lim_{y\to\infty} u(x, y) = 1/2$, so that the boundedness condition in (5.2.51) is met. However, the condition on $u(x, y)$ as $|x| \to \infty$ is not satisfied. (This condition was imposed to enable the simplification of the transform of the u_{xx} term.) Indeed, $H(x)$ does not even have a conventional Fourier transform since $H(x) = 1$ for $x \geq 0$. Nevertheless, (5.2.59) is a solution of Laplace's equation for $y > 0$ that satisfies the boundary condition at $y = 0$. Furthermore, $u(x, y)$ is a smooth, infinitely differentiable function for $y > 0$ even though its limit at $y = 0$ has a jump discontinuity at $x = 0$. The discontinuity in the boundary value is immediately smoothed out for $y > 0$, consistent with our observation in Section 3.2 that Laplace's equation characterizes equilibrium processes where everything is smoothed out.

In the exercises, we apply the solution formula (5.2.58) to several cases where $f(x)$ is either constant or piecewise constant, and may not be Fourier transformable, if $f(x)$ does not vanish at infinity. In all cases, as in the foregoing example, the solutions are bounded at infinity. It can be shown that the boundedness requirement implies that the solutions are unique.

If we consider the *Neumann problem* for Laplace's equation, the boundary condition in (5.2.50) is replaced by

$$\frac{\partial u(x, 0)}{\partial y} = g(x), \qquad -\infty < x < \infty, \tag{5.2.60}$$

and $u(x, y)$ must be bounded in the upper half-plane. Solving this problem directly, leads to certain difficulties with the convergence of the relevant Fourier transforms that we prefer to avoid. Instead, we solve it by using a general principle known as *Stokes' rule*. Here it takes the following form.

Let $u(x, y)$ satisfy

$$u_{xx}(x, y) + u_{yy}(x, y) = 0, \qquad -\infty < x < \infty,\ y > 0, \tag{5.2.61}$$

and the boundary condition

$$u_y(x, 0) = g(x), \qquad -\infty < x < \infty. \tag{5.2.62}$$

Put $v(x, y) = \partial u(x, y)/\partial y$. Then

$$\nabla^2 v(x, y) = \frac{\partial}{\partial y}\nabla^2 u(x, y) = 0, \quad v(x, 0) = u_y(x, 0) = g(x). \tag{5.2.63}$$

Thus if we can solve the Dirichlet problem (5.2.63) for $v(x, y)$, the solution of the Neumann problem (5.2.61)–(5.2.62) is $u(x, y) = \int^y v(x, s)\, ds$ and is determined up to an arbitrary constant. Using the solution (5.2.58) gives

$$u(x,y) = \frac{1}{\pi}\int^y \int_{-\infty}^{\infty} \frac{sg(t)}{(x-t)^2+s^2}\, dt\, ds \tag{5.2.64}$$

$$= \frac{1}{\pi}\int_{-\infty}^{\infty}\left(\int^y \frac{s\, ds}{(x-t)^2+s^2}\right) g(t)\, dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} \log[(x-t)^2+y^2]\, g(t)\, dt.$$

It can be verified that (5.2.64) satisfies (5.2.61)–(5.2.62).

We have considered the application of Fourier transforms to some basic but simple problems. In each case it was possible to obtain the solution in a fairly straightforward manner by inverting relevant Fourier transforms. Equations of parabolic, hyperbolic, and elliptic types were chosen to show the applicability of the Fourier transform technique for equations of each of the three basic types. The solutions obtained in the examples are discussed further, later in the text and in the exercises.

In general applications of transform methods it is not expected that the inverse transforms can be simplified as far as was done in the examples above. However, approximation methods are available that yield results in various regions of interest in the domain of the dependent variable. Certain methods of this type are presented in Section 5.7. Additionally, it is possible to evaluate inverse transforms numerically and approximately by replacing the infinite limits in the inversion integral by finite limits.

Exercises 5.2

5.2.1. Show that the (general) Fourier series (4.3.52) may be expressed in complex form as $f(x) = \sum_{k=-\infty}^{\infty} c_k \exp(-i\pi kx/l)$, with the Fourier coefficients c_k given as $c_k = (1/2l)\int_{-l}^{l} f(t)\exp(i\pi kt/l)$. This result may be obtained directly by using the orthogonality property of the functions $\exp[i(\pi k/l)x]$ over the interval $-l < x < l$, in the manner of Exercise 4.3.4.

5.2.2. Show that the results in Exercise 5.2.1 may be expressed as

$$f(x) = \frac{1}{2l}\sum_{k=-\infty}^{\infty}\int_{-l}^{l} f(t)\exp\left[-i\,\frac{\pi k}{l}\,(x-t)\right] dt.$$

Assume that the λ-axis ($-\infty < \lambda < \infty$) is subdivided into the intervals $[\lambda_{k-1}, \lambda_k]$, with $\lambda_k = \pi k/l$, whose length is $\Delta\lambda_k = \pi/l$. Rewrite the expression as

$$f(x) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty}\int_{-l}^{l} f(t)\exp[-i\lambda_k(x-t)]\, dt\, \Delta\lambda_k,$$

and show that as $l \to \infty$ this (formal) Riemann sum tends to the Fourier integral formula (5.2.5) if the limit exists.

5.2.3. Prove that if $f(x)$ is absolutely integrable, its Fourier transform (5.2.6) tends to zero as $|\lambda| \to \infty$. (This is the *Riemann-Lebesgue lemma.*) *Hint*: Let $x = y + \pi/\lambda$ in (5.2.6) and show that $F(\lambda)$ can be expressed in the form (5.2.6) with $f(x)$ replaced by $-f(x+\pi/\lambda)$. Take the average of the two expressions for $F(\lambda)$ and let $|\lambda| \to \infty$.

5.2.4. Consider the Fourier transforms found in Exercises 5.2.7, 5.2.9, and 5.2.10. Show how the rates at which these transforms tend to zero as $|\lambda| \to \infty$ are related to the smoothness properties of the inverse transforms $f(x)$.

5.2.5. Show that the Fourier transform of $xf(x)$ is $-iF'(\lambda)$.

5.2.6. Obtain the Fourier transform $Y(\lambda)$ of the Airy equation $y''(x) - xy(x) = 0$, with $Y(0) = 1/\sqrt{2\pi}$. Show that the inverse transform $Ai(x)$ has the integral representation given in Exercise 5.7.2.

5.2.7. Show that the Fourier transform of the function $F_N(x)$ defined in (5.2.20) is given as $F_N(\lambda) = \sqrt{2/\pi}\,\alpha(N)\,\sin(\lambda N)/\lambda$.

5.2.8. Let $\alpha(N) = 1/2N$ in Exercise 5.2.7 and show that the limit of $F_N(\lambda)$ as $N \to 0$ is $1/\sqrt{2\pi}$. Conclude that the Fourier transform of $\delta(x)$ should be $1/\sqrt{2\pi}$.

5.2.9. Obtain the Fourier transform of $f(x) = \exp[-k|x|]$, $k > 0$ and verify the result in (5.2.16).

5.2.10. Adapt the results of Example 5.2 to obtain the Fourier transform of $f(x) = \exp(-c^2x^2)$.

5.2.11. Use the formula (5.2.10) and the result of Exercise 5.2.10 to obtain the Fourier transform of the error integral $\operatorname{erf}(x) = (2/\sqrt{\pi}) \int_0^x e^{-s^2}\,ds$.

5.2.12. Obtain the (closed-form) solution of the problem (5.2.12)–(5.2.13) if $f(x)$ is defined as $f(x) = \begin{cases} 1, & |x| < a, \\ 0, & |x| > a. \end{cases}$

5.2.13 Obtain the result (5.2.42) by using (a) Duhamel's principle; (b) Fourier transforms.

5.2.14. Show that the fundamental solution $G(x-\xi, t)$ of the heat equation as defined in (5.2.39) has the property that as $t \to 0$, $\lim_{t\to 0} G(x-\xi, t) = 0$, $x \neq \xi$. Also demonstrate that for $t > 0$ we have $\int_{-\infty}^{\infty} G(x-\xi, t)\,dx = 1$. Conclude from these results that $G(x-\xi, t)$ tends to the Dirac delta function $\delta(x-\xi)$ as $t \to 0$.

5.2.15. Assuming that $u(x,t)$ and its derivatives vanish as $|x| \to \infty$, integrate the heat equation (5.2.26) over the x-axis and obtain the result $\int_{-\infty}^{\infty} u(x,t)\,dx = \int_{-\infty}^{\infty} f(x)\,dx$, where $u(x,0) = f(x)$. Using Exercise 5.2.14, verify the result for the solution given in (5.2.38).

5.2.16. Let the initial temperature $f(x)$ in (5.2.38) be $f(x) = \begin{cases} u_0, & |x| < a, \\ 0, & |x| > a, \end{cases}$ where u_0 = constant. Express (5.2.38) in terms of the error integral $\operatorname{erf}(x)$ defined in Exercise 5.2.11.

5.2.17. Use Fourier transforms to solve $u_{tt}(x,t) - c^2 u_{xx}(x,t) = F(x,t)$, $-\infty < x < \infty$, $t > 0$, with $u(x,0) = u_t(x,0) = 0$.

5.2.18. Solve the problem for the hyperbolic equation $u_{tt}(x,t) - c^2 u_{xx}(x,t) - a^2 u(x,t) = 0$, $-\infty < x < \infty$, $t > 0$ with the initial data $u(x,0) = f(x)$, $u_t(x,0) = 0$.

5.2.19. Use Fourier transforms to solve $u_{tt}(x,t) + u_{xxxx}(x,t) = 0$, $-\infty < x < \infty$, $t > 0$, with $u(x,0) = f(x)$, $u_t(x,0) = 0$. *Hint*: The inverse transform of $\cos(\lambda^2 t)$ is $(1/\sqrt{2t}) \cos\left(x^2/4t - \pi/4\right)$.

5.2.20. Let $f(t) = 1$ in (5.2.58) and evaluate the integral in terms of $\tan^{-1} z$ to show that the solution becomes $u(x,y) = 1$. Note that $u(x,y)$ does not tend to zero as $y \to \infty$ in this case and that $f(x) = 1$ does not have a Fourier transform.

5.2.21. Noting the result in Exercise 5.2.20 and the fact that $\lim_{y \to 0} K(x-t,y) = 0$, $x \neq t$, where $K(x-t,y) = (y/\pi)/[(x-t)^2 + y^2]$, show that $K(x-t,y)$ tends to $\delta(x-t)$ as $y \to 0$.

5.2.22. Let $f(t) = a$, $t < 0$ and $f(t) = b$, $t > 0$ in (5.2.58) and evaluate the integral in terms of the inverse tangent function.

5.2.23. Use Fourier transforms to solve the boundary value problem $u_{xx}(x,y) + u_{yy}(x,y) - c^2 u(x,y) = 0$, $-\infty < x < \infty$, $y > 0$ with the boundary conditions $u_y(x,0) = -f(x)$, $u(x,y)$ bounded as $y \to \infty$. *Hint*: The inverse transform of $\exp(-y\sqrt{\lambda^2 + c^2})/\sqrt{\lambda^2 + c^2}$ is $(\sqrt{2/\pi}) K_0[c\sqrt{x^2 + y^2}]$, where K_0 is the zero-order modified Bessel function of the second kind.

5.2.24. Solve Laplace's equation in a strip using Fourier transforms: $u_{xx}(x,y) + u_{yy}(x,y) = 0$, $-\infty < x < \infty$, $0 < y < L$, with $u(x,0) = e^{-|x|}$, $u(x,L) = 0$, $u(x,y) \to 0$ as $|x| \to \infty$.

5.3 FOURIER SINE AND COSINE TRANSFORMS

On considering the one-dimensional form of (5.1.1) over the semi-infinite interval $0 < x < \infty$ with $\rho(x) = p(x) = 1$ and $q(x) = 0$, we obtain the equation

$$M''(x) + \lambda^2 M(x) = 0, \qquad 0 < x < \infty. \tag{5.3.1}$$

Two sets of boundary conditions are considered for the function $u(x,t)$ [or $u(x,y)$], which give rise to the eigenvalue problem under consideration when the variables are separated as in Section 5.1. First, we assume that u vanishes at $x = 0$ and is bounded as $x \to \infty$. This implies the (homogeneous) boundary conditions

$$M(0) = 0, \qquad M(x) \text{ bounded as } \quad x \to \infty, \tag{5.3.2}$$

for the solution (5.3.1). Second, we assume that u_x vanishes at $x = 0$ and that u is bounded as $x \to \infty$. This yields the (homogeneous) boundary conditions

$$M'(0) = 0, \qquad M(x) \text{ bounded as } \quad x \to \infty, \tag{5.3.3}$$

for(5.3.1).

The eigenvalue problem (5.3.1)–(5.3.2) yields the eigenfunctions

$$M_\lambda(x) = \sin(\lambda x), \tag{5.3.4}$$

with the continuous spectrum $0 < \lambda < \infty$. Negative values of λ need not be considered since $\sin(-\lambda x) = -\sin(\lambda x)$.

The eigenvalue problem (5.3.1) and (5.3.3) gives rise to the eigenfunctions

$$M_\lambda(x) = \cos(\lambda x), \tag{5.3.5}$$

with the continuous spectrum $0 \leq \lambda < \infty$. Again, negative values of λ need not be considered since $\cos(-\lambda x) = \cos(\lambda x)$.

Let the function $f(x)$, defined for $x > 0$, satisfy conditions equivalent to those given in Section 5.2 for Fourier transformable functions. Then corresponding to the eigenvalue problem (5.3.1)–(5.3.2), we have the representation

$$F_s(\lambda) = \sqrt{\frac{2}{\pi}} \int_0^\infty \sin(\lambda x) f(x)\, dx, \tag{5.3.6}$$

with $F_s(\lambda)$ defined as the *Fourier sine transform* of $f(x)$. The *inversion formula* giving $f(x)$ in terms of its transform is

$$f(x) = \sqrt{\frac{2}{\pi}} \int_0^\infty \sin(\lambda x) F_s(\lambda)\, d\lambda. \tag{5.3.7}$$

The formula (5.3.7) yields an odd extension of $f(x)$ to the entire x-axis [i.e., $f(-x) = -f(x)$]. By considering the Fourier transform (5.2.6) and its inversion

formula (5.2.7) as applied to the odd extension of $f(x)$, we obtain (5.3.6)–(5.3.7), as shown in the exercises. Thus (5.3.7) converges to $f(x)$ at points of continuity and to its mean value at points of jump discontinuity.

For the eigenvalue problem (5.3.1) and (5.3.3) we have the representation

$$F_c(\lambda) = \sqrt{\frac{2}{\pi}} \int_0^\infty \cos(\lambda x) f(x)\, dx, \tag{5.3.8}$$

with $F_c(\lambda)$ defined as the *Fourier cosine transform* of $f(x)$. The *inversion formula* giving $f(x)$ in terms of its transform is

$$f(x) = \sqrt{\frac{2}{\pi}} \int_0^\infty \cos(\lambda x) F_c(\lambda)\, d\lambda. \tag{5.3.9}$$

These formulas result from the Fourier integral formulas (5.2.6)–(5.2.7) when applied to an even function $f(x)$ [i.e., $f(-x) = f(x)$], as is shown in the exercises. It is seen that (5.3.9) yields an even extension of $f(x)$ to the entire axis, and again, the convergence properties for (5.3.9) are equivalent to those for the Fourier transform.

General Properties

Convolution theorems for products of Fourier sine and cosine transforms are presented in the exercises. As was found for the Fourier transform, it is possible to represent Fourier sine and cosine transforms of derivatives of $f(x)$ in terms of Fourier sine and cosine transforms of $f(x)$. These results are obtained using integration by parts. We derive the results for the transforms of $f'(x)$ and $f''(x)$. Results for transforms of higher derivatives are easily obtained.

For the *sine transform* integral we have

$$\int_0^\infty f'(x) \sin(\lambda x)\, dx = f(x) \sin(\lambda x)\Big|_{x=0}^{x=\infty} - \lambda \int_0^\infty f(x) \cos(\lambda x)\, dx, \tag{5.3.10}$$

$$\int_0^\infty f''(x) \sin(\lambda x)\, dx = f'(x) \sin(\lambda x)\Big|_{x=0}^{x=\infty} - \lambda \int_0^\infty f'(x) \cos(\lambda x)\, dx$$

$$= f'(x) \sin(\lambda x)\Big|_{x=0}^{x=\infty} - \lambda f(x) \cos(\lambda x)\Big|_{x=0}^{x=\infty} - \lambda^2 \int_0^\infty f(x) \sin(\lambda x)\, dx. \tag{5.3.11}$$

Thus if we assume that $f(x)$ and $f'(x)$ vanish at infinity, we obtain on using the definitions (5.3.7) and (5.3.8),

$$\begin{cases} \sqrt{2/\pi} \int_0^\infty f'(x) \sin(\lambda x)\, dx = -\lambda F_c(\lambda), \\ \sqrt{2/\pi} \int_0^\infty f''(x) \sin(\lambda x)\, dx = \lambda \sqrt{2/\pi}\, f(0) - \lambda^2 F_s(\lambda). \end{cases} \tag{5.3.12}$$

We observe that the Fourier sine transform of the first derivative of a function $f'(x)$ is given in terms of the Fourier cosine transform of $f(x)$. The Fourier sine transform

of the second derivative $f''(x)$, is given in terms of the sine transform of $f(x)$. There is an additional boundary term $\lambda\sqrt{2/\pi}f(0)$ that vanishes if $f(0) = 0$, that is, if $f(x)$ satisfies the boundary condition of the relevant eigenvalue problem. The formulas (5.3.12) show that the use of the sine transform is not expected to be effective unless the differential equation contains only derivatives of even orders in the transformed independent variable and if only one boundary condition is assigned for each two derivatives of the transformed function.

A similar result is true for the *cosine transform*. We easily find on using (5.3.7) and (5.3.8) that

$$\begin{cases} \sqrt{2/\pi}\int_0^\infty f'(x)\cos(\lambda x)\,dx = -\sqrt{2/\pi}\,f(0) + \lambda F_s(\lambda), \\ \sqrt{2/\pi}\int_0^\infty f''(x)\cos(\lambda x)\,dx = -\lambda\sqrt{2/\pi}\,f'(0) - \lambda^2 F_c(\lambda), \end{cases} \tag{5.3.13}$$

if we assume that $f(x)$ and $f'(x)$ vanish at infinity. Again, if $f'(0) = 0$ so that $f(x)$ satisfies the boundary condition at $x = 0$ of the associated eigenvalue problem, the cosine transform of $f''(x)$ is given strictly in terms of the cosine transform of the function $f(x)$, as shown in (5.3.13).

The formulas (5.3.12) and (5.3.13) for the Fourier sine and cosine transforms of the second derivative of a function require that the function and its derivative, respectively, be specified at $x = 0$ if these transforms are to be useful in solving differential equations. Thus, the sine or cosine transforms are appropriate for problems over semi-infinite intervals in a spatial variable in which the function or its derivative are prescribed on the boundary. The Fourier sine and cosine transforms are generally not useful when applied with respect to a time variable in connection with initial conditions at $t = 0$. For example, the heat equation contains only one time derivative u_t, so that its sine transform is given in terms of its cosine transform. The wave equation contains a second time derivative u_{tt}, but since both u and u_t are specified at $t = 0$ neither the sine nor the cosine transform can account for the full initial data. As shown in Section 5.6, the *Laplace transform* is the proper choice when the time variable is transformed.

An additional property of the sine and cosine transforms that follows immediately from their definitions is that

$$\sqrt{\frac{2}{\pi}}\int_0^\infty \sin(\lambda x)xf(x)\,dx = -\frac{\partial F_c(\lambda)}{\partial\lambda}, \quad \sqrt{\frac{2}{\pi}}\int_0^\infty \cos(\lambda x)xf(x)\,dx = \frac{\partial F_s(\lambda)}{\partial\lambda}. \tag{5.3.14}$$

Further useful properties of Fourier sine and cosine transforms and transforms of specific functions are presented in the exercises, in Section 5.8, and are found in tables of Fourier sine and cosine transforms.

Applications to PDEs

We now consider several problems for PDEs that can be solved by the use of Fourier cosine and sine transforms. (A number of problems for ODEs are presented in the exercises.) It should be noted that some problems that can be solved by use of Fourier

transforms in one of the variables, such as in Example 5.4, can also be solved by the use of cosine or sine transforms in the other variable(s). Cosine or sine transforms can be used when the transformed spatial variable is restricted to a semi-infinite interval, and the choice of the cosine or sine transform is dictated by the boundary conditions, as indicated. As often happens in applications to PDEs, since more than one variable occurs in the problems, it may be that one variable is unrestricted whereas the other variable is bounded on one side. As a result, either the Fourier transform or the Fourier cosine or sine transform may be used to solve the given problem.

Example 5.5. The Heat Equation in a Semi-Infinite Interval. We consider the heat (or diffusion) equation

$$u_t(x,t) - c^2 u_{xx}(x,t) = 0, \qquad 0 < x < \infty,\ t > 0, \tag{5.3.15}$$

where c^2 is a constant, over the semi-infinite interval $0 < x < \infty$, with the initial condition

$$u(x,0) = f(x), \qquad 0 < x < \infty, \tag{5.3.16}$$

and either the boundary condition of the first kind,

$$u(0,t) = g(t), \qquad t > 0, \tag{5.3.17}$$

or the boundary condition of the second kind,

$$u_x(0,t) = h(t), \qquad t > 0. \tag{5.3.18}$$

For the first boundary value problem [i.e., where $u(0,t) = g(t)$] we apply the Fourier sine transform in x, since the term $u(0,t)$ corresponds to the $f(0)$ term that occurs in (5.3.12). For the second boundary value problem [i.e., where $u_x(0,t) = h(t)$] we use the Fourier cosine transform, in view of the relation between $u_x(0,t)$ and $f'(0)$ in (5.3.13).

Applying the *Fourier sine transform* to (5.3.15), we multiply through in (5.3.15) by the term $\sqrt{2/\pi}\sin(\lambda x)$ and integrate from 0 to ∞ to obtain

$$\frac{\partial U_s(\lambda,t)}{\partial t} + (\lambda c)^2 U_s(\lambda,t) = \lambda c^2 \sqrt{\frac{2}{\pi}}\, g(t) \tag{5.3.19}$$

on using (5.3.13) and (5.3.17). The function $U_s(\lambda,t)$ is the sine transform of $u(x,t)$; that is,

$$U_s(\lambda,t) = \sqrt{\frac{2}{\pi}} \int_0^\infty \sin(\lambda x) u(x,t)\, dx. \tag{5.3.20}$$

From the sine transform of the initial condition (5.3.16), we have $U_s(\lambda,0) = F_s(\lambda)$. The solution of the initial value problem for $U_s(\lambda,t)$ is

$$U_s(\lambda,t) = F_s(\lambda) e^{-\lambda^2 c^2 t} + \lambda c^2 \sqrt{\frac{2}{\pi}} \int_0^t e^{-\lambda^2 c^2 (t-\tau)} g(\tau)\, d\tau. \tag{5.3.21}$$

The inverse transform of $U_s(\lambda, t)$ yields the solution

$$u(x,t) = \frac{2}{\pi}\int_0^\infty \int_0^\infty e^{-\lambda^2 c^2 t}\sin(\lambda s)\sin(\lambda x) f(s)\, d\lambda\, ds$$

$$+\frac{2c^2}{\pi}\int_0^t \int_0^\infty \lambda e^{-\lambda^2 c^2 (t-\tau)} g(\tau)\sin(\lambda x)\, d\lambda\, d\tau$$

$$= \frac{1}{\pi}\int_0^\infty \int_0^\infty e^{-\lambda^2 c^2 t}\{\cos[\lambda(x-s)] - \cos[\lambda(x+s)]\} f(s)\, d\lambda\, ds$$

$$+\frac{2c^2}{\pi}\int_0^t \int_0^\infty \frac{\partial}{\partial x}\left[-e^{-\lambda^2 c^2 (t-\tau)}\cos(\lambda x)\right] g(\tau)\, d\lambda\, d\tau$$

$$= \int_0^\infty [G(x-s,t) - G(x+s,t)] f(s)\, ds - 2c^2 \int_0^t \frac{\partial G(x, t-\tau)}{\partial x} g(\tau)\, d\tau, \tag{5.3.22}$$

where we have used the notation $G(x,t)$ for the fundamental solution of the heat equation defined in (5.2.39). The foregoing result was obtained by interchanging the order of integration in both integrals and evaluating the inner integrals following the approach presented in (5.2.33)–(5.2.37). We have also used the identity $2\sin(\lambda x)\sin(\lambda s) = \cos[\lambda(x-s)] - \cos[\lambda(x+s)]$.

Using Duhamel's principle and (5.3.22), it is easy to see that the solution of the inhomogeneous heat equation

$$v_t(x,t) - c^2 v_{xx}(x,t) = F(x,t), \qquad x > 0,\ t > 0, \tag{5.3.23}$$

with data $v(x,0) = f(x)$ and $v(0,t) = g(t)$, is

$$v(x,t) = u(x,t) + \int_0^t \int_0^\infty [G(x-s,t-\tau) - G(x+s,t-\tau)] F(s,\tau)\, ds\, d\tau, \tag{5.3.24}$$

where $u(x,t)$ is given by (5.3.22).

In the case of a uniform initial temperature, $u(x,0) = f(x) = u_0 =$ constant, we readily obtain, on changing the variable of integration,

$$\int_0^\infty [G(x-s,t) - G(x+s,t)] f(s)\, ds = \int_{-x/2c\sqrt{t}}^\infty \frac{u_0 e^{-r^2}}{\sqrt{\pi}}\, dr - \int_{x/2c\sqrt{t}}^\infty \frac{u_0 e^{-r^2}}{\sqrt{\pi}}\, dr$$

$$= \frac{2u_0}{\sqrt{\pi}}\int_0^{x/2c\sqrt{t}} e^{-r^2}\, dr = u_0 \operatorname{erf}\left(\frac{x}{2c\sqrt{t}}\right), \tag{5.3.25}$$

where $\operatorname{erf}(z)$ is the *error function* integral, defined as

$$\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}}\int_0^z e^{-r^2}\, dr, \tag{5.3.26}$$

with the properties that erf(0) = 0 and erf(∞) = 1. Thus the solution of the initial and boundary value problem (5.3.15)–(5.3.17) with $u(x,0) = u_0$ and $u(0,t) = 0$ is given as

$$u(x,t) = u_0 \operatorname{erf}\left(\frac{x}{2c\sqrt{t}}\right). \tag{5.3.27}$$

We note that $u(x,0) = u_0 \operatorname{erf}(\infty) = u_0$, $u(0,t) = u_0 \operatorname{erf}(0) = 0$ and $u(x,t) \to 0$ as $t \to \infty$. Even though $f(x) = u_0$ does not have a sine transform, the solution (5.3.27) satisfies all conditions of the problem. Again it can be shown that (5.3.27) is the limit of a sequence of problems with Fourier transformable initial data, but we do not demonstrate this.

We observe that if the initial temperature $f(x)$ is not constant but $|f(x)| < M < \infty$ for all $x > 0$ (as we indeed assume), we have

$$\left|\int_0^\infty [G(x-s,t) - G(x+s,t)]f(s)\,ds\right| \le M \int_0^\infty [G(x-s,t) - G(x+s,t)]\,ds$$

$$= M \operatorname{erf}\left(\frac{x}{2c\sqrt{t}}\right) \to 0 \qquad \text{as} \qquad t \to \infty, \tag{5.3.28}$$

since $G(x-s,t) \ge G(x+s,t)$ for $x, s \ge 0$ and $\operatorname{erf}(z) \to 0$ as $z \to 0$. [We have used the extended mean value theorem for integrals applied to the finite interval $[0, N]$ and then let $N \to \infty$ to obtain the inequality in (5.3.28).] Thus as $t \to \infty$, the solution $u(x,t)$ [i.e., (5.3.22)] of the initial and boundary value problem (5.3.15)–(5.3.17) tends to the *steady state*

$$u(x,t) \approx -2c^2 \int_0^t \frac{\partial G(x,t-\tau)}{\partial x} g(\tau)\,d\tau, \qquad t \gg 1, \tag{5.3.29}$$

in the sense that the effect of the initial temperature distribution $u(x,0) = f(x)$ is dissipated. (The term *steady state* does not signify time-independence for this problem.)

In fact, if we consider the initial value problem for the heat equation (5.3.15) with $0 < x < \infty$ and $t > t_0$, and specify that $u(x,t_0) = \hat{f}(x)$ and $u(0,t) = \hat{g}(t)$, the solution is

$$u(x,t) = \int_0^\infty [G(x-s,t-t_0) - G(x+s,t-t_0)]\hat{f}(s)\,ds$$

$$-2c^2 \int_{t_0}^t \frac{\partial G(x,t-\tau)}{\partial x} \hat{g}(\tau)\,d\tau, \tag{5.3.30}$$

as follows from (5.3.22). We assume that $|\hat{f}(x)| < M$ and that $\hat{g}(t)$ is defined for all t. The estimate (5.3.28) shows that the absolute value of the first integral is bounded by $M \operatorname{erf}(x/2c\sqrt{t-t_0})$. Then as $t_0 \to -\infty$, if $u(x,t)$ is bounded for all time, we obtain in the limit

$$u(x,t) = -2c^2 \int_{-\infty}^t \frac{\partial G(x,t-\tau)}{\partial x} \hat{g}(\tau)\,d\tau. \tag{5.3.31}$$

This is the solution of the steady-state problem for (5.3.15) over the interval $0 < x < \infty$, where no initial condition is prescribed and the boundary condition is given for all $t > -\infty$ as $u(0,t) = \hat{g}(t)$. If

$$u(0,t) = \hat{g}(t) = \begin{cases} g(t), & t \geq 0, \\ 0, & t < 0, \end{cases} \tag{5.3.32}$$

(5.3.31) reduces to (5.3.29).

If the initial temperature $u(x,0) = f(x) = 0$ and the boundary temperature $u(0,t) = g(t) = u_1$ = constant, it is readily verified (as shown in the exercises) that the solution (5.3.22) can be expressed as

$$u(x,t) = u_1 \left[1 - \operatorname{erf}\left(\frac{x}{2c\sqrt{t}}\right)\right] = u_1 \operatorname{erfc}\left(\frac{x}{2c\sqrt{t}}\right), \tag{5.3.33}$$

where $\operatorname{erfc}(z)$ is the *complementary error function*, defined as $\operatorname{erfc}(z) = 1 - \operatorname{erf}(z)$. As we see from its definition, $\operatorname{erfc}(z) \to 1$ as $z \to 0$.

Thus from (5.3.28) we conclude that the temperature $u(x,t) \to 0$ as $t \to \infty$ if the boundary temperature $g(t) = 0$ and the initial temperature $u(x,0) = f(x)$ is uniformly bounded. If the boundary temperature $g(t) = u_1$ = constant and the initial temperature is uniformly bounded, (5.3.33) shows that the temperature $u(x,t)$ tends to the constant state u_1 as $t \to \infty$.

Finally, we remark that the function

$$w(x,t) = -2c^2 \frac{\partial G(x,t)}{\partial x}, \tag{5.3.34}$$

which occurs in the last integral in the solution (5.3.22), is a solution of the homogeneous heat equation (5.3.15) that satisfies the conditions

$$\lim_{t \to 0+} w(x,t) = 0, \ x > 0, \qquad \lim_{x \to 0+} w(x,t) = 0, \ t > 0. \tag{5.3.35}$$

However, as we approach the origin $(x,t) = (0,0)$ along the curve $x = 2c\sqrt{t}$, we have

$$\begin{aligned} \lim_{t \to 0+} \left[-2c^2 \frac{\partial G(x,t)}{\partial x}\right]_{x=2c\sqrt{t}} &= \lim_{t \to 0+} \left[\frac{x}{2\sqrt{\pi}\, ct^{3/2}} \exp\left(-\frac{x^2}{4c^2 t}\right)\right]_{x=2c\sqrt{t}} \\ &= \lim_{t \to 0+} \frac{e^{-1}}{\sqrt{\pi t}} \to \infty. \end{aligned} \tag{5.3.36}$$

Thus $w(x,t)$ is unbounded at the origin. Although $w(x,t)$ is a solution of the heat equation that satisfies homogeneous initial and boundary conditions, in view of (5.3.35), it is unbounded at the origin. To guarantee that solutions of the initial and boundary value problem (5.3.15)–(5.3.17) are determined uniquely, we must require that the solution $u(x,t)$ be bounded. Otherwise, we could add the function $w(x,t)$ multiplied by an arbitrary constant to any solution, without altering the initial and

boundary data for the problem. In fact, $w(x,t)$ is a solution of the heat equation for $-\infty < x < \infty$ and $t > 0$, and its limit as t tends to zero along any line x = constant is zero. However, it is not bounded at the origin in view of (5.3.36). Consequently, unless we require that the solution of the initial value problem for the heat equation is bounded for all $t \geq 0$, the solution is not unique.

Using the *Fourier cosine transform* and proceeding as above, we readily obtain the solution of the second boundary value problem (5.3.15)–(5.3.16) and (5.3.18) in the form

$$u(x,t) = \int_0^\infty [G(x-s,t) - G(x+s,t)]f(s)\,ds - 2c^2 \int_0^t G(x,t-\tau)h(\tau)\,d\tau$$
$$+ \int_0^t \int_0^\infty [G(x-s,t-\tau) + G(x+s,t-\tau)]F(s,\tau)\,ds\,d\tau, \qquad (5.3.37)$$

where the last term occurs if the inhomogeneous problem for the heat equation [i.e., (5.3.23)] is considered.

With constant initial temperature $u(x,0) = f(x) = u_0$ = constant and $h = F = 0$ in (5.3.37) we have, after some manipulation of the integrals,

$$u(x,t) = \int_0^\infty [G(x-s,t) - G(x+s,t)]u_0\,ds = u_0 \left[\frac{2}{\sqrt{\pi}} \int_0^\infty e^{-r^2}\,dr\right] = u_0. \qquad (5.3.38)$$

This is to be expected since $\partial u(0,t)/\partial x$ is a measure of the amount of heat passing through the boundary $x = 0$ and we have $\partial u(0,t)/\partial x = 0$. Also, $F(x,t) = 0$, so that there are no heat sources. As a result, no heat escapes through the boundary $x = 0$ and no heat is generated. Consequently, the temperature remains fixed at $u(x,t) = u_0$.

It should also be noted that the function

$$\hat{w}(x,t) = -2c^2 G(x,t) \qquad (5.3.39)$$

that occurs in the second integral in (5.3.37) has the property that

$$\lim_{t\to 0+} \hat{w}(x,t) = 0,\ x > 0, \qquad \lim_{x\to 0+} \frac{\partial \hat{w}(x,t)}{\partial x} = 0,\ t > 0, \qquad (5.3.40)$$

as is readily seen. Also, $\hat{w}(x,t)$ is a solution of the heat equation (5.3.15). However, $\hat{w}(x,t)$ is unbounded at $(x,t) = (0,0)$ if we approach the origin along the curve given in (5.3.36), so we must require the solution of the second boundary value problem to be bounded in order to obtain a unique solution.

The solution of the third boundary value problem for the heat equation in the semi-infinite interval with the boundary condition

$$u_x(0,t) - hu(0,t) = r(t), \qquad (5.3.41)$$

where $h > 0$ is a constant and $r(t)$ is given, is considered in the exercises.

Some of the results obtained in the foregoing example are rederived in Chapter 7 using Green's function methods. We have expressed them in terms of the fundamental solution $G(x,t)$ for the purposes of comparison with later results.

Example 5.6. Laplace's Equation in a Strip and a Quarter-Plane. We consider *Laplace's equation* in a semi-infinite strip

$$u_{xx}(x,y) + u_{yy}(x,y) = 0, \qquad 0 < x < \infty,\ 0 < y < \alpha, \tag{5.3.42}$$

with the boundary data

$$u(0,y) = 0, \qquad u(x,y) \to 0 \quad \text{as} \quad x \to \infty, \quad \text{uniformly in} \quad y, \tag{5.3.43}$$

$$u(x,0) = f(x), \qquad u(x,\alpha) = 0 \qquad 0 < x < \infty. \tag{5.3.44}$$

Since $u(x,y)$ is specified on the boundary $x = 0$, the *Fourier sine transform* is appropriate for this problem. Let the sine transform of $u(x,y)$ be defined as

$$U_s(\lambda,y) = \sqrt{\frac{2}{\pi}} \int_0^\infty \sin(\lambda x) u(x,y)\, dx \tag{5.3.45}$$

and multiply through in (5.3.42) by $\sqrt{2/\pi}\sin(\lambda x)$ and integrate from 0 to ∞. Using (5.3.13) and noting that $u(0,y) = 0$, we obtain

$$\frac{\partial^2 U_s(\lambda,y)}{\partial y^2} - \lambda^2 U_s(\lambda,y) = 0. \tag{5.3.46}$$

With $F_s(\lambda)$ as the sine transform of $f(x)$ and in view of (5.3.44), we have

$$U_s(\lambda,0) = F_s(\lambda), \quad U_s(\lambda,\alpha) = 0, \tag{5.3.47}$$

and

$$U_s(\lambda,y) = F_s(\lambda)\frac{\sinh[\lambda(\alpha - y)]}{\sinh(\lambda\alpha)}. \tag{5.3.48}$$

The inverse sine transform yields

$$u(x,y) = \frac{2}{\pi}\int_0^\infty \int_0^\infty f(s)\sin(\lambda s)\sin(\lambda x)\frac{\sinh[\lambda(\alpha - y)]}{\sinh(\lambda\alpha)}\, ds\, d\lambda. \tag{5.3.49}$$

Note that if $\alpha\lambda \gg 1$, we have

$$\frac{\sinh[\lambda(\alpha - y)]}{\sinh(\lambda\alpha)} \approx e^{-\lambda y}, \qquad \alpha\lambda \gg 1. \tag{5.3.50}$$

Thus the λ-integral in (5.3.49) converges at an exponential rate. The s-integral also converges well if $f(x) \to 0$ as $x \to \infty$, as is, in fact, required by the boundary condition (5.3.43). We remark that if there are inhomogeneous boundary conditions

at $y = \alpha$ and $x = 0$, the resulting problem can be solved in a similar fashion. Also, if we replace the Dirichlet boundary condition in (5.3.43) with the Neumann condition $u_x(0, y) = g(y)$ at $x = 0$, the problem can be solved with the use of the *Fourier cosine transform.*

In the limit as $\alpha \to \infty$ in the problem above we obtain the quarter-plane problem for Laplace's equation, that is,

$$u_{xx}(x, y) + u_{yy}(x, y) = 0, \qquad x > 0,\ y > 0, \tag{5.3.51}$$

with the boundary conditions

$$u(0, y) = 0,\ u(x, 0) = f(x),\ \lim_{x \to \infty} u(x, y) \to 0 \text{ uniformly in } \quad y. \tag{5.3.52}$$

This problem may be solved by a direct application of the sine transform or by going to the limit as $\alpha \to \infty$ in the solution (5.3.49). Using the latter approach and noting (5.3.50), we obtain as the solution of (5.3.51)–(5.3.52),

$$u(x, y) = \frac{2}{\pi} \int_0^\infty \int_0^\infty f(s) \sin(\lambda s) \sin(\lambda x) e^{-\lambda y}\, ds\, d\lambda. \tag{5.3.53}$$

Interchanging the order of integration, we have for the inner integral,

$$\begin{aligned} 2 \int_0^\infty \sin(\lambda s) \sin(\lambda x) e^{-\lambda y}\, d\lambda &= \int_0^\infty e^{-\lambda y} \{\cos[\lambda(x - s)] - \cos[\lambda(x + s)]\}\, d\lambda \\ &= \frac{y}{y^2 + (x - s)^2} - \frac{y}{y^2 + (x + s)^2}, \end{aligned} \tag{5.3.54}$$

on adapting the result in equation (5.2.58). Then the solution $u(x, y)$ takes the form

$$u(x, y) = \frac{y}{\pi} \int_0^\infty \left[\frac{1}{y^2 + (x - s)^2} - \frac{1}{y^2 + (x + s)^2} \right] f(s)\, ds. \tag{5.3.55}$$

The formula (5.3.55) may be compared with that given in (5.2.58) for the solution of the half-plane problem for Laplace's equation. If we define $f(t)$ in (5.2.58) to be an odd function of t [i.e., $f(-t) = -f(t)$], the integral (5.2.58) can easily be transformed into (5.3.55). This is equivalent to extending the above quarter-plane problem to a half-plane problem by extending the solution $u(x, y)$ to the full infinite x-interval as an odd function [i.e., $u(-x, y) = -u(x, y)$] with $u(0, y) = 0$. Then if $f(x)$ is also extended as an odd function, we would expect the solutions (5.2.58) and (5.3.55) to agree, as they indeed do. A further connection between half- and quarter-plane problems is considered in Chapter 7 when we discuss Green's functions.

Exercises 5.3

5.3.1. Show that if $f(x)$ is an odd function of x [i.e., $f(-x) = -f(x)$], the formulas (5.2.6)–(5.2.7) for the Fourier transform yield the formulas (5.3.6)–(5.3.7) for the Fourier sine transform.

5.3.2. Show that if $f(x)$ is an even function of x [i.e., $f(-x) = f(x)$], the formulas (5.2.6)–(5.2.7) yield the formulas (5.3.8)–(5.3.9) for the Fourier cosine transform.

5.3.3. Determine the Fourier sine and cosine transforms of the following functions: (a) $f(x) = e^{-kx}$, $k > 0$; (b) $f(x) = \begin{cases} 1, & 0 \le x \le a, \\ 0, & x > a; \end{cases}$ (c) $f(x) = xe^{-x}$; (d) $f(x) = e^{-x}\cos(x)$.

5.3.4. By specializing the convolution theorem for the Fourier transform to even and odd functions $f(x)$ and $g(x)$ as needed, obtain the following convolution theorems for sine and cosine transforms:
(a) $\int_0^\infty \sin(\lambda x)F_s(\lambda)G_c(\lambda)\,d\lambda = \frac{1}{2}\int_0^\infty f(t)[g(|x-t|) - g(x+t)]\,dt$.
(b) $\int_0^\infty \cos(\lambda x)F_c(\lambda)G_c(\lambda)\,d\lambda = \frac{1}{2}\int_0^\infty f(t)[g(|x-t|) + g(x+t)]\,dt$.

5.3.5. Use the Fourier sine transform, Exercise 5.3.3(a), and the convolution theorem to solve $y''(x) - k^2y(x) = -f(x)$, $x > 0$, $y(0) = 0$, $y(x) \to 0$ as $x \to \infty$.

5.3.6. Use the Fourier cosine transform, Exercise 5.3.3(a), and the convolution theorem to solve $y''(x) - k^2y(x) = -f(x)$, $x > 0$, $y'(0) = 0$, $y(x) \to 0$ as $x \to \infty$.

5.3.7. Use the Fourier sine transform to solve the boundary value problem for the ordinary differential equation $y''(x) - k^2y(x) = e^{-x}$, $x > 0$, $k \neq 1$, $y(0) = 1$, $y(x) \to 0$ as $x \to \infty$.

5.3.8. Use the Fourier cosine transform to solve the problem in Exercise 5.3.7 if we replace the condition $y(0) = 1$ by the boundary condition $y'(0) = 1$.

5.3.9. Use Duhamel's principle to obtain the result (5.3.24).

5.3.10. Show that the solution (5.3.22) takes the form (5.3.33) if $f(x) = 0$ and $g(t) = u_1$ = constant.

5.3.11. Suppose that $u(x,t)$ satisfies the heat equation (5.3.15), the initial condition (5.3.16), and the boundary condition (5.3.41). Let $v(x,t) = u_x(x,t) - hu(x,t)$. Show that $v(x,t)$ is also a solution of the heat equation; that is, $v_t(x,t) - c^2v_{xx}(x,t) = 0$, $0 < x < \infty$, $t > 0$, with the initial condition $v(x,0) = f'(x) - hf(x)$, $x > 0$, and the boundary condition $v(0,t) = r(t)$, $t > 0$. Assuming that $u(x,t)$ is bounded as $x \to \infty$, show that the solution of the given problem for $u(x,t)$ is given in terms of $v(x,t)$ as $u(x,t) = e^{hx}\int_\infty^x e^{-hs}v(s,t)\,ds$.

5.3.12. Use the Fourier sine transform to solve the following initial and boundary value problem for the wave equation $u_{tt}(x,t) - c^2u_{xx}(x,t) = 0$, $0 < x < \infty$, $t > 0$, $u(x,0) = u_t(x,0) = 0$, $u(0,t) = g(t)$.

5.3.13. Solve the following problem for the wave equation using the Fourier cosine transform $u_{tt}(x,t) - c^2u_{xx}(x,t) = 0$, $0 < x < \infty$, $t > 0$, $u(x,0) = u_t(x,0) = 0$, $u_x(0,t) = g(t)$.

5.3.14. Solve the Dirichlet and Neumann problems for Laplace's equation in a half-plane given in Example 5.4 by the use of Fourier sine and cosine transforms, respectively.

5.3.15. Solve the boundary value problem (5.3.51)–(5.3.52) for Laplace's equation using the Fourier sine transform.

5.3.16. Use the Fourier sine transform to solve the problem $u_{xx}(x,y)+u_{yy}(x,y)=0,\ 0<x<\infty,\ 0<y<\alpha,\ u(x,0)=e^{-x},\ u(x,\alpha)=0,\ u(0,y)=1,\ u(x,y)\to 0$ as $x\to\infty$.

5.3.17. Apply the Fourier cosine transform to solve the problem $u_{xx}(x,y)+u_{yy}(x,y)=0,\ 0<x<\infty,\ 0<y<\infty$, with the boundary conditions $u_x(0,y)=0,\ y>0,\ u(x,0)=\begin{cases}1, & 0<x<1,\\ 0, & x>1.\end{cases}$

5.3.18. Use Stokes' rule in the manner indicated in Example 5.4 to construct a solution of the problem $u_{xx}(x,y)+u_{yy}(x,y)=0,\ 0<x<\infty,\ 0<y<\infty,\ u(0,y)=0,\ u_y(x,0)=g(x)$, from the solution of the problem (5.3.51)–(5.3.52) given in (5.3.55).

5.4 HIGHER-DIMENSIONAL FOURIER TRANSFORMS

Higher dimensional Fourier transforms may be characterized as before in terms of higher dimensional eigenvalue problems for (5.1.1) with $\rho=p=1$ and $q=0$ over the entire space. Alternatively, they may be obtained by a repeated application of one-dimensional Fourier transforms in each of the variables. The conditions of validity for the transforms are then readily carried over from the one-dimensional case. Instead of discussing the properties of the higher-dimensional transforms, we merely define the transforms and the inversion formulas and cite the necessary properties as they are needed.

Let $\mathbf{x}=[x_1,\ldots,x_n]$ and $\boldsymbol{\lambda}=[\lambda_1,\ldots,\lambda_n]$ be n-component vectors. Under hypotheses on $f(\mathbf{x})$ analogous to those given in the one-dimensional case, we have for the n-dimensional *Fourier transform* $F(\boldsymbol{\lambda})$ of $f(\mathbf{x})$ the representation

$$F(\boldsymbol{\lambda})=\frac{1}{(\sqrt{2\pi})^n}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}e^{i\boldsymbol{\lambda}\cdot\mathbf{x}}f(\mathbf{x})\,d\mathbf{x}, \tag{5.4.1}$$

and the *Fourier inversion formula*

$$f(\mathbf{x})=\frac{1}{(\sqrt{2\pi})^n}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}e^{-i\boldsymbol{\lambda}\cdot\mathbf{x}}F(\boldsymbol{\lambda})\,d\boldsymbol{\lambda}. \tag{5.4.2}$$

The integrals (5.4.1)–(5.4.2) are both n-dimensional and $d\mathbf{x}$ and $d\boldsymbol{\lambda}$ are n-dimensional volume elements and $\boldsymbol{\lambda}\cdot\mathbf{x}=\sum_{i=1}^{n}\lambda_i x_i$ is the scalar or dot product of the vectors $\boldsymbol{\lambda}$ and $\mathbf{x}$. The formulas (5.2.10)–(5.2.11) relating the transforms of derivatives to the transforms of given functions are valid in the higher-dimensional case as well. Here they must be applied to transforms of partial derivatives of $f(\mathbf{x})$ and we require that $f(\mathbf{x})$ and its partial derivatives vanish at infinity.

It is also possible to consider transforms equivalent to the sine and cosine transforms given above if the problem is given over a semi-infinite space. We remark that although Fourier transforms lead to a formally simple approach to the solution of initial and boundary value problems for partial differential equations, the evaluation

and simplification of the resulting integral representation of the solution is generally not a simple task. This was observed in the examples considered above. We restrict our discussion to two and three-dimensional Fourier transforms.

Cauchy Problem for the Three-Dimensional Wave Equation: Spherical Means and Stokes' Rule

We examine the Cauchy or initial value problem for the *wave equation in three dimensions*, with $u = u(x, y, z, t)$, in the form

$$u_{tt} - c^2 [u_{xx} + u_{yy} + u_{zz}] = 0, \qquad t > 0, \ -\infty < x, \ y, \ z < \infty, \tag{5.4.3}$$

where c^2 is a constant, with the initial conditions

$$u(x, y, z, 0) = 0, \quad u_t(x, y, z, 0) = f(x, y, z). \tag{5.4.4}$$

We consider only the case where $u(x, y, z, 0) = 0$, since according to *Stokes' rule* (see Example 5.4), the solution of the wave equation (5.4.3) with data $u(x, y, z, 0) = f(x, y, z)$, $u_t(x, y, z, 0) = 0$ is $\partial u(x, y, z, t)/\partial t$, where $u(x, y, z, t)$ satisfies (5.4.3)–(5.4.4). This is easily verified. Then the solution of the general initial value problem for (5.4.3) is just the sum of these two solutions.

To solve (5.4.3)–(5.4.4) by the Fourier transform method, we multiply (5.4.3) by $1/(\sqrt{2\pi})^3 \exp[i(\lambda_1 x + \lambda_2 y + \lambda_3 z)]$ and integrate with respect to x, y, and z from $-\infty$ to ∞. Let $U(\boldsymbol{\lambda}, t)$ denote the Fourier transform of $u(x, y, z, t)$. Using the analogs of the formulas (5.2.10)–(5.2.11), we obtain the transformed equation

$$\frac{\partial^2 U(\boldsymbol{\lambda}, t)}{\partial t^2} + c^2(\lambda_1^2 + \lambda_2^2 + \lambda_3^2) U(\boldsymbol{\lambda}, t) = 0, \tag{5.4.5}$$

on assuming that u and its first partial derivatives vanish at infinity. The initial conditions for $U(\boldsymbol{\lambda}, t)$ are

$$U(\boldsymbol{\lambda}, 0) = 0, \qquad U_t(\boldsymbol{\lambda}, 0) = F(\boldsymbol{\lambda}), \tag{5.4.6}$$

where $F(\boldsymbol{\lambda})$ is the Fourier transform of $f(x, y, z)$.

The solution of (5.4.5)–(5.4.6) is

$$U(\boldsymbol{\lambda}, 0) = F(\boldsymbol{\lambda}) \frac{\sin[|\boldsymbol{\lambda}| ct]}{|\boldsymbol{\lambda}| c}, \tag{5.4.7}$$

where $|\boldsymbol{\lambda}| = \sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}$. Inverting the Fourier transform yields

$$u(x, y, z, t) = \frac{1}{(\sqrt{2\pi})^3} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{F(\boldsymbol{\lambda})}{|\boldsymbol{\lambda}| c} \sin[|\boldsymbol{\lambda}| ct] e^{-i\boldsymbol{\lambda} \cdot \mathbf{x}} \, d\boldsymbol{\lambda}, \tag{5.4.8}$$

where $\boldsymbol{\lambda} = [\lambda_1, \ \lambda_2, \lambda_3]$, $\mathbf{x} = [x, y, z]$ and $d\boldsymbol{\lambda} = d\lambda_1 \, d\lambda_2 \, d\lambda_3$.

Writing $\sin[|\boldsymbol{\lambda}|ct] = (1/2i)(e^{i|\boldsymbol{\lambda}|ct} - e^{-i|\boldsymbol{\lambda}|ct})$, we express $u(x, y, z, t)$ as

$$u(x,y,z,t) = \frac{1}{2i(\sqrt{2\pi})^3} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{F(\boldsymbol{\lambda})}{|\boldsymbol{\lambda}|c}$$

$$\times \left\{ \exp\left[i|\boldsymbol{\lambda}| \left(ct - \frac{\boldsymbol{\lambda}\cdot\mathbf{x}}{|\boldsymbol{\lambda}|} \right) \right] - \exp\left[-i|\boldsymbol{\lambda}| \left(ct + \frac{\boldsymbol{\lambda}\cdot\mathbf{x}}{|\boldsymbol{\lambda}|} \right) \right] \right\} d\boldsymbol{\lambda}. \tag{5.4.9}$$

The terms $\exp[\pm i|\boldsymbol{\lambda}|\,(ct \mp \boldsymbol{\lambda}\cdot\mathbf{x}/|\boldsymbol{\lambda}|)]$ in (5.4.9) are *plane wave solutions* of the wave equation (5.4.3). That is, the solutions remain constant on planes $\boldsymbol{\lambda}\cdot\mathbf{x}$ = constant that move (parallel to themselves) at the speed c. Thus (5.4.9) is a superposition of plane wave solutions traveling in all possible directions. (These plane waves represent normal mode solutions of the wave equation.)

The integral (5.4.9) can be simplified by expressing $F(\boldsymbol{\lambda})$ in terms of $f(x, y, z)$. It is then possible to integrate out four of the resulting integrals, and one method for doing so that involves the use of the Dirac delta function is given in Example 7.9. Instead of carrying out this process, we merely quote the final result and indicate in the exercises an alternative, simpler method for deriving the final expression.

The solution $u(x, y, z, t)$ is given in terms of spherical coordinates with their origin at the point (x, y, z) and defined as

$$\xi = x + r\sin(\phi)\cos(\theta), \quad \eta = y + r\sin(\phi)\sin(\theta), \quad \zeta = z + r\cos(\phi), \tag{5.4.10}$$

$$u(x,y,z,t) = \tag{5.4.11}$$

$$\frac{t}{4\pi}\int_0^{\pi}\int_0^{2\pi} f(x+ct\sin(\phi)\cos(\theta), y+ct\sin(\phi)\sin(\theta), z+ct\cos(\phi))\sin(\phi)\,d\theta\,d\phi.$$

Given a sphere S of radius a with center at (x, y, z), we define the average value of the function $f(\xi, \eta, \zeta)$ over the sphere to be given as

$$M_a[f] = \frac{1}{4\pi a^2}\iint_S f\,ds, \tag{5.4.12}$$

where $4\pi a^2$ is the area of the sphere S and ds is the area element on the sphere. In the spherical coordinates (5.4.10), the area element is given as $ds = a^2\sin(\phi)\,d\theta\,d\phi$. Then it is easily seen that (5.4.11) can be written as

$$u(x, y, z, t) = tM_{ct}[f], \tag{5.4.13}$$

where $M_{ct}[f]$ is the average value of f over the sphere of radius ct with center at (x, y, z). $M_{ct}[f]$ is denoted as the *spherical mean* of f over the sphere, and as such our solution method is called the *method of spherical means*.

Recalling *Stokes' rule* and using superposition of the two solutions, we obtain as the solution $u(x, y, z, t)$ of the wave equation (5.4.3) with initial data

$$u(x,y,z,0) = g(x,y,z), \quad u_t(x,y,z,0) = f(x,y,z), \tag{5.4.14}$$

$$u(x,y,z,t) = tM_{ct}[f] + \frac{\partial}{\partial t}\{tM_{ct}[g]\}. \tag{5.4.15}$$

The expression (5.4.15) is often called *Kirchhoff's formula* for the solution of the wave equation.

Using *Duhamel's principle* (see Section 4.5), we readily obtain the solution of the inhomogeneous wave equation for $u(x,y,z,t)$;

$$u_{tt} - c^2\,[u_{xx} + u_{yy} + u_{zz}] = G(x,y,z,t), \quad t > 0,\ -\infty < x,\ y,\ z < \infty, \tag{5.4.16}$$

with initial data (5.4.14), as

$$u(x,y,z,t) = tM_{ct}[f] + \frac{\partial}{\partial t}\{tM_{ct}[g]\} + \int_0^t (t-\tau)M_{c(t-\tau)}[G(x,y,z,\tau)]\,d\tau. \tag{5.4.17}$$

The integral term on the right in (5.4.17) can be expressed as a triple integral over the interior of the sphere $r = ct$ with center at (x,y,z) and radius ct, where $r = \sqrt{(\xi-x)^2 + (\eta-y)^2 + (\zeta-z)^2}$. In rectangular coordinates, the integral takes a form referred to as a *retarded potential*:

$$\int_0^t (t-\tau)M_{c(t-\tau)}[G]\,d\tau = \frac{1}{4\pi c^2}\iiint_{r\le ct} \frac{1}{r} G\left(\xi,\eta,\zeta,t-\frac{r}{c}\right) d\xi\,d\eta\,d\zeta. \tag{5.4.18}$$

In the retarded potential the contributions to the solution at the point (x,y,z) and at time t come only from the points in the interior and on a sphere of radius ct with center at (x,y,z). Since the speed of wave propagation is c, the function G, evaluated at the retarded or earlier time $t - r/c$, is integrated over a sphere of radius r, as the effect of these points reaches the center of the sphere at the time t.

Cauchy Problem for the Two-Dimensional Wave Equation: Hadamard's Method of Descent

Before discussing the solution (5.4.15), we obtain from it using *Hadamard's method of descent* the solution of the initial value problem for the *wave equation in two dimensions*. In this method we look for a solution of the wave equation (5.4.3) independent of z, with initial data

$$u(x,y,t)\Big|_{t=0} = g(x,y), \qquad \frac{\partial u(x,y,t)}{\partial t}\Big|_{t=0} = f(x,y), \tag{5.4.19}$$

so that $u(x,y,t)$ satisfies the wave equation in two dimensions,

$$u_{tt}(x,y,t) - c^2[u_{xx}(x,y,t) + u_{yy}(x,y,t)] = 0, \quad t > 0,\ -\infty < x,\ y < \infty. \tag{5.4.20}$$

We introduce $f(x,y)$ and $g(x,y)$ into the solution (5.4.15) and because the data are independent of z we may take the center of the sphere $r = ct$ to be at $(x,y,0)$. Introducing spherical coordinates on the sphere $r = \sqrt{(\xi-x)^2+(\eta-y)^2+\zeta^2} = ct$ with the surface element ds, we can write

$$u(x,y,t) = \frac{1}{4\pi c^2 t}\iint_{r=ct} f(\xi,\eta)\,ds + \frac{1}{4\pi c^2}\frac{\partial}{\partial t}\left\{\frac{1}{t}\iint_{r=ct} g(\xi,\eta)\,ds\right\}. \quad (5.4.21)$$

The sphere $r = ct$ has its center at $(x,y,0)$ and the (ξ,η)-plane cuts it in the circle $r = \sqrt{(\xi-x)^2+(\eta-y)^2} = ct$. Since the integrands in (5.4.21) are independent of ξ, we can integrate out the ξ dependence by effectively projecting the upper and lower hemispheres onto the circle in the (ξ,η)-plane. Using the relation $ds = \sec(\gamma)\,d\xi\,d\eta$, between the surface element ds on the sphere and the area element $d\xi\,d\eta$ in the plane, where

$$\sec(\gamma) = \frac{ct}{\sqrt{ct^2-(\xi-x)^2-(\eta-y)^2}}, \quad (5.4.22)$$

and noting that we get the same contribution from each hemisphere, we obtain

$$u(x,y,t) = \frac{1}{2\pi c}\iint_{r\le ct}\frac{f(\xi,\eta)\,d\xi\,d\eta}{\sqrt{(ct)^2-r^2}} + \frac{1}{2\pi c}\frac{\partial}{\partial t}\left\{\iint_{r\le ct}\frac{g(\xi,\eta)\,d\xi\,d\eta}{\sqrt{(ct)^2-r^2}}\right\}, \quad (5.4.23)$$

where $r = \sqrt{(\xi-x)^2+(\eta-y)^2}$. The integration in (5.4.23) is carried out over the interior of the circle of radius ct centered at (x,y). For the nonhomogeneous problem

$$u_{tt}(x,y,t) - c^2[u_{xx}(x,y,t)+u_{yy}(x,y,t)] = G(x,y,t),\ t>0,\ -\infty<x,\ y<\infty \quad (5.4.24)$$

with homogeneous initial data, *Duhamel's principle* yields the solution

$$u(x,y,t) = \frac{1}{2\pi c}\int_0^t\iint_{r\le c(t-\tau)}\frac{G(\xi,\eta,\tau)}{\sqrt{c(t-\tau)^2-r^2}}\,d\xi\,d\eta\,d\tau, \quad (5.4.25)$$

where $r = \sqrt{(\xi-x)^2+(\eta-y)^2}$.

This technique is called the *method of descent* since we descend from three to two space dimensions to construct a solution for the lower-dimensional problem in terms of the solution of the higher dimensional one. It is also possible to descend from two dimensions to one and construct d'Alembert's solution from the solution (5.4.23).

Huygens' Principle

We now discuss the solution (5.4.15) and (5.4.23) of the three- and two-dimensional wave equations, respectively. The basic distinction between the forms of the two solutions is that (5.4.15) represents integration over the surface of a sphere whereas (5.4.23) represents integration over the interior of a circle rather than on the circle itself. This characterizes a sharp difference in the nature of wave propagation in three and two dimensions.

We suppose that the initial data f and g are concentrated in a neighborhood of a point P_0 in the two- or the three-dimensional case. That is, we assume that they vanish outside an arbitrarily small neighborhood of P_0. In the three-dimensional case, (5.4.15) states that the solution $u(x, y, z, t)$ of the wave equation at the time t and at the point (x, y, z) depends on the data f and g only at points intersected by the sphere of radius ct with center at (x, y, z). If this sphere does not intersect the given neighborhood of P_0, the solution $u(x, y, z, t) = 0$. As t increases, the sphere must eventually intersect a neighborhood of P_0 [no matter where (x, y, z) is located]. The first time a disturbance is felt at (x, y, z) represents the time t it takes the *wave front*, originating at the point of the neighborhood of P_0 nearest to (x, y, z), to reach the point (x, y, z). Since the distance from the nearest point in the neighborhood of P_0 to (x, y, z) is ct, the speed of the propagation of the wave front is c. Because the data f and g are concentrated near P_0, as t increases the sphere will ultimately no longer intersect the given neighborhood of P_0 and the solution $u(x, y, z, t)$ will again be equal to zero.

The foregoing discussion shows that sharp signals or disturbances can occur in three-dimensional wave motion. Not only does it take a finite time for disturbances to travel from one point to another, but after the disturbance has passed, the solution returns to zero and the effect of the disturbance is no longer felt. This phenomenon is known as *Huygens' principle*. Huygens used this principle to construct solutions of wave propagation problems. Given the surface of a disturbance, or a (sharp) wave front, at the time t_0, a sphere of radius ct is drawn surrounding each point on the surface. The envelope of these spheres then gives the location of the wave front at the time $t_0 + t$. Although this construction only gives the front of the disturbance in general, in the case of concentrated sources (i.e., initial data concentrated at a point P_0) the construction gives the full disturbance (see Example 7.9). If the wave has a trailing edge, it disappears once the trailing edge has passed. (Huygens' construction of wave fronts served as the basis of the *wave theory of light*.)

For the two-dimensional case, if the initial data f and g are concentrated in a neighborhood of the point P_0, the solution (5.4.23) shows that $u(x, y, t)$ will be zero until the circle with center at (x, y) and radius ct intersects the neighborhood of P_0. Once the circle begins to intersect the neighborhood of P_0, the solution at the point (x, y) will not vanish as t increases, because the integration in (5.4.23) is carried out over the interior of the circle, not just its boundary. Thus although sharp wave fronts can exist in two-dimensional wave propagation, since it takes a finite time for the effect of the data near P_0 to reach the point (x, y), the effects of the disturbance linger and do not disappear sharply as in the three-dimensional case. Instead, they diffuse slowly to zero as $t \to \infty$ (see Example 7.9). As a result, Huygens' principle, which predicts sharp wave fronts and trailing edges, is not valid in the two-dimensional case. Yet, Huygens' construction can be used to determine the location of wave fronts.

The solutions given for the two- and three-dimensional wave equations can be used to determine domains of dependence and influence as was done earlier in the one-dimensional case, but this is not carried out here. Also, as seen in the exercises, Hadamard's method can be applied to other PDEs.

Helmholtz and Modified Helmholtz Equations

We begin by considering the (inhomogeneous) elliptic equation

$$u_{xx}(x,y) + u_{yy}(x,y) - k^2 u(x,y) = f(x,y), \qquad -\infty < x,\ y < \infty, \quad (5.4.26)$$

where the constant $k > 0$, $f(x,y)$ is specified, and $u(x,y)$ satisfies the additional condition

$$u(x,y) \to 0 \quad \text{as} \quad |x|,\ |y| \to \infty. \quad (5.4.27)$$

We assume that $f(x,y)$ has a Fourier transform. The equation (5.4.26) is sometimes referred to as the *modified Helmholtz equation*, in contrast to the case where $-k^2$ is replaced by $+k^2$, which is referred to as the *Helmholtz equation*. A stationary version of two-dimensional diffusion equations can give rise to (5.4.26).

To solve (5.4.26)–(5.4.27), we apply the two-dimensional Fourier transform to (5.4.26) and easily obtain

$$U(\boldsymbol{\lambda}) = -\frac{F(\boldsymbol{\lambda})}{\lambda_1^2 + \lambda_2^2 + k^2}, \quad (5.4.28)$$

where $\boldsymbol{\lambda} = [\lambda_1,\ \lambda_2]$, and $U(\boldsymbol{\lambda})$ and $F(\boldsymbol{\lambda})$ are the Fourier transforms of $u(x,y)$ and $f(x,y)$, respectively. [We have assumed in obtaining (5.4.28) that not only u but also its first partial derivatives vanish at infinity.]

Inverting the transform gives the solution

$$u(x,y) = -\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-i(\lambda_1 x + \lambda_2 y)} \frac{F(\boldsymbol{\lambda})}{\lambda_1^2 + \lambda_2^2 + k^2}\, d\lambda_1\, d\lambda_2. \quad (5.4.29)$$

Since $F(\boldsymbol{\lambda})$ has the form

$$F(\boldsymbol{\lambda}) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{i(\lambda_1 \xi + \lambda_2 \eta)} f(\xi,\eta)\, d\xi\, d\eta, \quad (5.4.30)$$

we obtain, on inserting (5.4.30) into (5.4.29),

$$u(x,y) = -\frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-i[\lambda_1(x-\xi)+\lambda_2(y-\eta)]} \frac{f(\xi,\eta)\, d\lambda_1\, d\lambda_2\, d\xi\, d\eta}{\lambda_1^2 + \lambda_2^2 + k^2}, \quad (5.4.31)$$

on interchanging the order of integration. Introducing polar coordinates in the two inner integrals in (5.4.31), that is, $\lambda_1 = \rho\cos\phi$, $\lambda_2 = \rho\sin\phi$, and also expressing $x - \xi$ and $y - \eta$ in polar form $x - \xi = r\cos\theta$, $y - \eta = r\sin\theta$, we readily obtain $\lambda_1(x-\xi) + \lambda_2(y-\eta) = \rho r\cos(\phi - \theta)$. Transforming to polar coordinates gives

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\exp(-i[\lambda_1(x-\xi) + \lambda_2(y-\eta)])}{\lambda_1^2 + \lambda_2^2 + k^2}\, d\lambda_1\, d\lambda_2$$

$$= \int_0^{\infty}\int_0^{2\pi} \frac{e^{-i\rho r\cos(\phi-\theta)}}{\rho^2 + k^2}\, \rho\, d\phi\, d\rho. \quad (5.4.32)$$

Using the integral representation of the *Bessel function of order zero* $J_0(z)$,

$$J_0(z) = \frac{1}{2\pi}\int_0^{2\pi} e^{iz\cos(\phi-\theta)}\,d\phi, \qquad \theta = \text{constant}, \tag{5.4.33}$$

we obtain, since $J_0(z) = J_0(-z)$,

$$\int_0^{\infty}\int_0^{2\pi} \frac{e^{-i\rho r\cos(\phi-\theta)}}{\rho^2+k^2}\,\rho\,d\phi\,d\rho = 2\pi\int_0^{\infty}\frac{\rho J_0(\rho r)}{\rho^2+k^2}\,d\rho. \tag{5.4.34}$$

A further result from the theory of Bessel functions yields

$$2\pi\int_0^{\infty}\frac{\rho J_0(\rho r)}{\rho^2+k^2}\,d\rho = -2\pi K_0(kr), \tag{5.4.35}$$

where K_0 is the modified Bessel function of the second kind, some of whose properties are given in Chapter 6. When (5.4.35) is introduced into (5.4.31), we obtain, with $r = \sqrt{(\xi-x)^2+(\eta-y)^2}$,

$$u(x,y) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K_0(kr)f(\xi,\eta)\,d\xi\,d\eta. \tag{5.4.36}$$

If $f(x,y) = \delta(x-x_0)\delta(y-y_0)$, the two-dimensional Dirac delta function (5.4.36) reduces to

$$u(x,y) = \frac{1}{2\pi}K_0(kr), \qquad r = \sqrt{(x-x_0)^2+(y-y_0)^2}. \tag{5.4.37}$$

This is known as the *free-space Green's function* for the modified Helmholtz equation. As shown in the exercises, (5.4.26) becomes a modified Bessel equation when expressed in a polar coordinate form independent of the angular variable. We return to this problem in later chapters.

Replacing k^2 by $-k^2$ in (5.4.26), results in the (inhomogeneous) *Helmholtz equation*. On making use of the Fourier transform to solve, we conclude that the transform of the solution is given by (5.4.28) with k^2 replaced by $-k^2$. Thus, the transform is singular and the inversion formula for the determination of $u(x,y)$ cannot be applied. [The situation is analogous to that encountered for (5.2.22).] Rather than consider the general case, we obtain the *free-space Green's function* for the Helmholtz equation. To do so, we replace k by $-ik$ in (5.4.37). Using a known relationship between the function K_0 of imaginary argument and the Hankel function of the first kind $H_0^{(1)}$ yields

$$u(x,y) = \frac{i}{4}H_0^{(1)}(kr), \qquad r = \sqrt{(x-x_0)^2+(y-y_0)^2}. \tag{5.4.38}$$

This solution satisfies the *radiation condition* of Section 10.1 in view of the behavior of the Hankel function for large argument given in (6.7.36). Replacing k by ik in (5.4.37) leads to a Hankel function of the second kind $H_0^{(2)}$, and the solution does not satisfy the radiation condition.

Exercises 5.4

5.4.1. Express the three-dimensional wave equation in spherical (spatial) coordinates and show that if we look for a solution $u = u(r,t)$ where r is the radial variable, the function $v(r,t) = ru(r,t)$ is a solution of the one-dimensional wave equation $v_{tt}(r,t) - c^2 v_{rr}(r,t) = 0$. Obtain the general solution $u(r,t) = F(r-ct)/r + G(r+ct)/r$, where F and G are arbitrary functions. These solutions represent spherical (propagating) waves.

5.4.2. Consider the function $\delta_\epsilon(x)$ defined in Example 1.2 and the integral

$$I(x,y,z,t;\epsilon) = \frac{1}{4\pi c}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\xi,\eta,\zeta)\frac{1}{r}\delta_\epsilon(r-ct)\,d\xi\,d\eta\,d\zeta,$$

where $r^2 = (x-\xi)^2+(y-\eta)^2+(z-\zeta)^2$. Since $\delta_\epsilon(r-ct)$ vanishes when $|r-ct| > \epsilon$, the integral extends only over a finite region. Conclude, in view of the fact that the integral represents a superposition of spherical waves, that $I(x,y,z,t;\epsilon)$ is a solution of the three-dimensional wave equation. Introduce the spherical coordinates defined in (5.4.10) and show that the integral takes the form

$$I[x,y,z,t;\epsilon] = \int_0^{\pi}\int_0^{2\pi}\int_0^{\infty} f[x+r\sin(\phi)\cos(\theta), y+r\sin(\phi)\sin(\theta), z+r\cos(\phi)] \times \delta_\epsilon(r-ct)r\sin(\phi)/4\pi c\,dr\,d\theta\,d\phi.$$

Carry out the limit as $\epsilon \to 0$, use the property (1.1.19) of δ_ϵ and the fact that $\lim_{\epsilon\to 0}\delta_\epsilon(r-ct)$ vanishes when $r \neq ct$, and conclude that $I(x,y,z,t;\epsilon)$ tends to (5.4.11) as $\epsilon \to 0$.

5.4.3. Verify by direct differentiation that (5.4.11) is a solution of (5.4.3)–(5.4.4).

5.4.4. Solve the problem for the wave equation in three dimensions $u_{tt}(x,y,z,t) - c^2[u_{xx}(x,y,z,t) + u_{yy}(x,y,z,t) + u_{zz}(x,y,z,t)] = 0$, with $-\infty < x,y,z < \infty$, $t > 0$, if the initial data are given as $u(x,y,z,0) = 0$, $u_t(x,y,z,0) = \begin{cases} 1, & r < a, \\ 0, & r > a, \end{cases}$ where $r^2 = x^2 + y^2 + z^2$.

5.4.5. Solve the following problem for the wave equation using the spherical wave solutions obtained in Exercise 5.4.1: $u_{tt}(x,y,z,t) - c^2[u_{xx}(x,y,z,t) + u_{yy}(x,y,z,t) + u_{zz}(x,y,z,t)] = 0$, with $-\infty < x,y,z < \infty$, $t > 0$, if the initial data are $u(x,y,z,0) = 1$, $u_t(x,y,z,0) = r^2$, where $r^2 = x^2 + y^2 + z^2$.

5.4.6. Obtain the solution of the Cauchy problem for the two-dimensional wave equation $u_{tt}(x,y,t) - c^2[u_{xx}(x,y,t) + u_{yy}(x,y,t)] = 0$, $-\infty < x,y < \infty$, $t > 0$, with $u(x,y,0) = 0$, $u_t(x,y,0) = x^2 + y^2$. Determine the value of $u(0,0,t)$.

5.4.7. Solve $u_{tt}(x,y,t) - c^2[u_{xx}(x,y,t) + u_{yy}(x,y,t)] = 0$, $-\infty < x,y < \infty$, $t > 0$, with $u(x,y,0) = x$, $u_t(x,y,0) = 0$.

5.4.8. Apply the method of descent to solve the Cauchy problem for the Klein-Gordon equation $v_{tt}(x,t) - c^2 v_{xx}(x,t) + a^2 v(x,t) = 0$, $-\infty < x < \infty$, $t > 0$,

with $v(x,0) = 0$, $v_t(x,0) = f(x)$. Let $u(x,y,t) = \cos(ay/c)v(x,t)$ and show that $u(x,y,t)$ satisfies the wave equation in two dimensions. Then descend from two dimensions to one.

5.4.9. Use the method of descent to solve the Cauchy problem for the hyperbolic equation, $v_{tt}(x,t) - c^2 v_{xx}(x,t) - a^2 v(x,t) = 0$, $-\infty < x < \infty$, $t > 0$, with $v(x,0) = 0$, $v_t(x,0) = f(x)$. Let $u(x,y,t) = \exp(ay/c)v(x,t)$ and show that $u(x,y,t)$ satisfies the wave equation in two dimensions. Then descend from two dimensions to one.

5.4.10. Express the Laplacian in polar coordinates and show that (5.4.37) is a solution of the homogeneous form of (5.4.26) when $r > 0$.

5.4.11. Use the two-dimensional Fourier transform to solve the following Cauchy problem for the heat equation, $u_t(x,y,t) = c^2[u_{xx}(x,y,t) + u_{yy}(x,y,t)]$, $-\infty < x, y < \infty$, $t > 0$, with $u(x,y,0) = f(x,y)$.

5.4.12. Use the three-dimensional Fourier transform to solve the Cauchy problem for the heat equation (as formulated in Exercise 5.4.11) in the case of three dimensions.

5.4.13. Consider the following Cauchy problem: $u_{tt}(x,y,z,t) - c^2[u_{xx}(x,y,z,t) + u_{yy}(x,y,z,t) + u_{zz}(x,y,z,t)] = F(x,y,z)e^{-i\omega t}$, with $-\infty < x, y, z < \infty$, $t > 0$ and initial data $u(x,y,z,0) = 0$, $u_t(x,y,z,0) = 0$, where ω is a constant. Assume that $F(x,y,z)$ vanishes outside the bounded region V. Use the retarded potential (5.4.18) to show that for large t the solution of this problem has the form

$$u(x,y,z,t) = \frac{e^{-i\omega t}}{4\pi c^2} \iiint_V F(\xi,\eta,\zeta)\frac{e^{ikr}}{r}\, d\xi\, d\eta\, d\zeta, \quad t \to \infty,$$

where $k = \omega/c$. Noting that u has the form $u(x,y,z,t) = v(x,y,z)e^{-i\omega t}$, show that $v(x,y,z)$ satisfies the inhomogeneous reduced wave equation $\nabla^2 v(x,y,z) + k^2 v(x,y,z) = -(1/c^2)F(x,y,z)$, $-\infty < x, y, z < \infty$. $v = (1/4\pi r)e^{ikr}$ is the free-space Green's function for the reduced wave equation (see Section 6.7).

5.5 HANKEL TRANSFORMS

We have seen in Chapter 4 and its exercises that certain problems lead naturally to the representation of the solutions in terms of Bessel functions. In bounded regions we are led to the Bessel function expansions considered in Chapter 4. For unbounded regions, in carrying out separation of variables in two or three space dimensions in cases where the Laplacian operator is expressed in polar or cylindrical coordinates, we are often led to consider the following eigenvalue problem in the radial variable r:

$$L[v(r)] = -\frac{d}{dr}\left(r\frac{dv(r)}{dr}\right) + \frac{n^2}{r}v(r) = \lambda^2 r v(r), \qquad 0 < r < \infty, \qquad (5.5.1)$$

where λ^2 is the eigenvalue parameter and the equation is written in the self-adjoint form (4.3.1). [In fact, (5.5.1) is Bessel's equation of order n in the variable λr.] Here n

can be any positive integer or zero. We require that the eigenfunctions $v(r)$ be bounded for all r. This implies that the (unnormalized) eigenfunctions are $v_{n,\lambda}(r) = J_n(\lambda r)$, where $J_n(z)$ is the Bessel function of order n and λ can be any nonnegative real number so that the spectrum is continuous. Each value of n determines a different family of eigenfunctions. [Since $J_n(-z) = (-1)^n J_n(z)$, we need not consider negative values of λ.]

A function $f(r)$ defined for $r \geq 0$ and satisfying the conditions given in Section 5.3 has the representation in terms of the eigenfunctions $v_{n,\lambda}(r)$,

$$f(r) = \int_0^\infty \lambda J_n(\lambda r) F_n(\lambda)\, d\lambda, \qquad n \geq 0, \tag{5.5.2}$$

where the *Hankel transform* $F_n(\lambda)$ is given by

$$F_n(\lambda) = \int_0^\infty r J_n(\lambda r) f(r)\, dr, \qquad n \geq 0. \tag{5.5.3}$$

In this context, (5.5.2) is the *inversion formula* for the Hankel transform and $f(r)$ is the *inverse Hankel transform* of $F_n(\lambda)$. For each integer $n \geq 0$, the integral (5.5.3) determines a Hankel transform of order n. [We note that alternative forms of the Hankel transform formulas are given in the literature.] The formulas (5.5.2)–(5.5.3) can be obtained from the two-dimensional Fourier transform formulas, and this derivation is considered in the exercises.

General Properties

As has been demonstrated, the usefulness of the Fourier, sine and cosine transforms in solving boundary value problems for differential equations rests on the fact that transforms of derivatives of functions can be related to the transforms of the functions themselves. The corresponding property for the Hankel transform is contained in the following equation. (The bracketed term on the left arises directly or after some intermediate calculations when the Laplacian is expressed in polar or cylindrical coordinates.) In applying the Hankel transform of order n to a given equation we are led to consider an expression of the form

$$\int_0^\infty r J_n(\lambda r) \left[f_{rr} + \frac{1}{r} f_r - \frac{n^2}{r^2} f \right] dr$$

$$= -\lambda^2 F_n(\lambda) + \left[r f_r J_n(\lambda r) - \lambda r f J_n'(\lambda r) \right]_{r=0}^{r=\infty}. \tag{5.5.4}$$

It is obtained on integrating by parts and is closely related to (4.3.58). [The operator L of (5.5.1) is used.] In general, the bracketed term on the right is required to vanish. This occurs if f and f_r are bounded at the origin and $\sqrt{r} f$ and $\sqrt{r} f_r$ vanish at infinity. Note that for $n \geq 0$ the Bessel function $J_n(\lambda r)$ is bounded at zero and vanishes like $1/\sqrt{r}$ at infinity. In certain problems, we permit f to be singular at $r = 0$ but in such a fashion that the limits $\lim_{r \to 0} r f_r J_n$ and $\lim_{r \to 0} r f J_n'$ are finite.

Although comprehensive tables of Hankel transforms are available (Maple can also be used), we list several results for use in the examples. The equation

$$\int_0^\infty e^{-\lambda z} J_0(\lambda r)\, d\lambda = \frac{1}{\sqrt{z^2+r^2}}, \qquad z,\ r > 0, \tag{5.5.5}$$

shows that the zero-order Hankel transform of $f(r,z) = 1/\sqrt{z^2+r^2}$ is given as $F_0(\lambda, z) = (1/\lambda)e^{-\lambda z}$. Using the transform formula (5.5.3), we obtain

$$\int_0^\infty \frac{rJ_0(\lambda r)}{\sqrt{z^2+r^2}}\, dr = \frac{1}{\lambda} e^{-\lambda z}. \tag{5.5.6}$$

The result (5.5.5) can be verified by using the integral representation of the Bessel function $J_0(\lambda r)$ and interchanging orders of integration (see Exercise 5.6.10). A second result that can be obtained by using the series expansion of the Bessel function $J_n(\lambda r)$ and integrating term by term is

$$\int_0^\infty \lambda J_n(\lambda r) e^{-t\lambda^2} \lambda^n\, d\lambda = \frac{1}{2t}\left(\frac{r}{2t}\right)^n \exp\left(-\frac{r^2}{4t}\right). \tag{5.5.7}$$

We cite additional results as they are required and now consider several examples in which Hankel transforms are used to solve problems for equations of elliptic, hyperbolic, and parabolic types.

Applications to PDEs

Example 5.7. Laplace's Equation: An Axial Source. We consider Laplace's equation in three dimensions with a source concentrated on the z-axis. Introducing cylindrical coordinates (r, θ, z) and noting the axial symmetry, we ask for a θ-independent solution $u(r,z)$ of Laplace's equation,

$$\nabla^2 u(r,z) = u_{rr}(r,z) + \frac{1}{r} u_r(r,z) + u_{zz}(r,z) = 0,\ r > 0,\ -\infty < z < \infty, \tag{5.5.8}$$

with the conditions on the z-axis (i.e., $r = 0$) given as

$$\lim_{r\to 0} r^2 u(r,z) = 0, \quad \lim_{r\to 0} 2\pi r \frac{\partial u(r,z)}{\partial r} = -f(z), \quad -\infty < z < \infty, \tag{5.5.9}$$

so that $f(z)$ is a measure of the strength of the source. [We discuss concentrated source problems in Section 6.7.]

Applying the zero-order Hankel transform to (5.5.8) [in view of (5.5.4)], we multiply across by $rJ_0(\lambda r)$ and integrate from 0 to ∞. [The zero-order transform is used because the term $(n^2/r^2)u(r,z)$ does not occur in equation (5.5.8).] Defining the zero-order Hankel transform $U_0(\lambda, z)$ of $u(r,z)$ as

$$U_0(\lambda, z) = \int_0^\infty rJ_0(\lambda r) u(r,z)\, dr \tag{5.5.10}$$

and using (5.5.4), we obtain

$$0 = \int_0^\infty rJ_0(\lambda r)\left[u_{rr}(r,z) + \frac{1}{r}u_r(r,z) + u_{zz}(r,z)\right] dr$$

$$= -\lambda^2 U_0(\lambda,z) + \frac{\partial^2 U_0(\lambda,z)}{\partial z^2} + [ru_r(r,z)J_0(\lambda r) - \lambda r u(r,z)J_0'(\lambda r)]_{r=0}^{r=\infty}. \quad (5.5.11)$$

Now $J_0(0) = 1$, $J_0'(\lambda r) = -J_1(\lambda r)$ and $J_1(\lambda r) \approx \lambda r/2$ as $r \to 0$. Assuming that the contributions from the limit at infinity vanish in (5.5.11) and noting that $ruJ_0'(\lambda r) \approx -(\lambda r^2/2)u$ as $r \to 0$, we obtain on using (5.5.9),

$$\frac{\partial^2 U_0(\lambda,z)}{\partial z^2} - \lambda^2 U_0(\lambda,z) = -\frac{1}{2\pi}f(z), \qquad -\infty < z < \infty. \quad (5.5.12)$$

Requiring that $U_0(\lambda,z) \to 0$ as $|z| \to \infty$, we have precisely the problem considered in Example 5.1. The solution is

$$U_0(\lambda,z) = \frac{1}{4\pi\lambda}\int_{-\infty}^{\infty} e^{-\lambda|z-s|}f(s)\,ds, \quad (5.5.13)$$

on comparing with (5.2.17). Inverting the transform $U_0(\lambda,z)$ gives

$$u(r,z) = \frac{1}{4\pi}\int_{-\infty}^{\infty}\int_0^\infty e^{-\lambda|z-s|}J_0(\lambda r)f(s)\,d\lambda\,ds = \frac{1}{4\pi}\int_{-\infty}^{\infty}\frac{f(s)\,ds}{\sqrt{r^2+(z-s)^2}}. \quad (5.5.14)$$

on interchanging the order of integration and using (5.5.5).

The term $(1/4\pi)[r^2+(z-s)^2]^{-1/2}$ is the *source function* or *free-space Green's function* for Laplace's equation, with the source point at $(x,y,z) = (0,0,s)$, (see Example 6.13). Thus, the solution (5.5.14) represents a superposition of point source functions over the z-axis with density $f(z)$.

Example 5.8. The Wave Equation in Two Dimensions. We consider the initial value problem for the two-dimensional wave equation over the entire spatial region with radially symmetric initial data. We introduce polar coordinates (r,θ) and look for a solution $u(r,t)$ independent of θ. This gives

$$u_{tt}(r,t) - c^2\left(u_{rr}(r,t) + \frac{1}{r}u_r(r,t)\right) = 0, \qquad r > 0,\ t > 0, \quad (5.5.15)$$

with initial data

$$u(r,0) = f(r), \qquad u_t(r,0) = g(r). \quad (5.5.16)$$

We again use the zero-order Hankel transform $U_0(\lambda,t)$ as defined in (5.5.10) (with z replaced by t), and obtain the problem

$$\frac{\partial^2 U_0(\lambda,t)}{\partial t^2} + c^2\lambda^2 U_0(\lambda,t) = 0, \qquad t > 0, \quad (5.5.17)$$

with the initial conditions

$$U_0(\lambda, 0) = F_0(\lambda), \qquad \frac{\partial U_0(\lambda, 0)}{\partial t} = G_0(\lambda), \tag{5.5.18}$$

where U_0, F_0, and G_0 are the zero-order Hankel transforms of u, f, and g, respectively, and (5.5.4) with $n = 0$ has been used. [It is assumed that u is such that the boundary terms in (5.5.4) vanish.] Solving (5.5.17)–(5.5.18) gives

$$U_0(\lambda, t) = F_0(\lambda)\cos(\lambda ct) + \frac{G_0(\lambda)}{\lambda c}\sin(\lambda ct). \tag{5.5.19}$$

Inverting the transform yields

$$u(r,t) = \int_0^\infty \lambda F_0(\lambda)\cos(\lambda ct)J_0(\lambda r)\, d\lambda + \frac{1}{c}\int_0^\infty G_0(\lambda)\sin(\lambda ct)J_0(\lambda r)\, d\lambda. \tag{5.5.20}$$

Example 5.9. The Diffusion Equation with Axially Symmetric Data. We consider the three-dimensional diffusion equation in cylindrical coordinates where the concentration $u(x, y, z, t)$ has an initial distribution that depends only on $r = \sqrt{x^2 + y^2}$. We introduce cylindrical coordinates (r, θ, z) and look for a solution $u(r, t)$ of the initial value problem

$$u_t(r,t) = D\left(u_{rr}(r,t) + \frac{1}{r}u_r(r,t)\right), \; t > 0, \quad u(r,0) = f(r), \; r > 0, \tag{5.5.21}$$

where $D > 0$ is the diffusion constant.

Applying the zero-order Hankel transform to (5.5.21), we obtain, on using the notation of Example 5.8,

$$\frac{\partial U_0(\lambda,t)}{\partial t} + D\lambda^2 U_0(\lambda,t) = 0, \; t > 0, \quad U_0(\lambda,0) = F_0(\lambda). \tag{5.5.22}$$

Thus

$$U_0(\lambda,t) = F_0(\lambda)e^{-D\lambda^2 t} \tag{5.5.23}$$

and

$$u(r,t) = \int_0^\infty \int_0^\infty \lambda s J_0(\lambda r)J_0(\lambda s)e^{-D\lambda^2 t}f(s)\, d\lambda\, ds, \tag{5.5.24}$$

on interchanging the order of integration in the last integral. We cite a known result to evaluate the inner integral. That is,

$$\int_0^\infty \lambda e^{-D\lambda^2 t}J_0(\lambda r)J_0(\lambda s)\, d\lambda = \frac{1}{2Dt}\exp\left(-\frac{r^2+s^2}{4Dt}\right)I_0\left(\frac{rs}{2Dt}\right), \tag{5.5.25}$$

where $I_0(z)$ is the modified zero-order Bessel function discussed in Chapter 6. We note the property of $I_0(z)$ that $I_0(0) = 1$, so that for $s = 0$ in (5.5.25) [since $J_0(0) = 1$] we find that (5.5.25) reduces to (5.5.7) with $n = 0$ and $D = 1$. Inserting (5.5.25) into (5.5.24) gives

$$u(r,t) = \frac{1}{2Dt}\int_0^\infty \exp\left(-\frac{r^2+s^2}{4Dt}\right) I_0\left(\frac{rs}{2Dt}\right) f(s)s\,ds. \tag{5.5.26}$$

It is of interest to show how this solution reduces to the fundamental solution for the two-dimensional heat or diffusion equation if we let $f(r)$ represent a concentrated source at $r = 0$. We require that with $f(r) \geq 0$,

$$\lim_{\epsilon\to 0} 2\pi\int_0^\epsilon f(r)r\,dr = 1, \tag{5.5.27}$$

where we assume that the source is concentrated in a circle of radius ϵ and let the radius tend to zero but keep the source strength fixed at unity. It may be assumed that $f(r)$ vanishes for $r > \epsilon$. Using the generalized mean value theorem for integrals, we have [since $f(r) \geq 0$]

$$u(r,t) = \frac{1}{2Dt}\exp\left(-\frac{r^2+\hat{s}^2}{4Dt}\right) I_0\left(\frac{r\hat{s}}{2Dt}\right)\int_0^\epsilon f(s)s\,dr, \tag{5.5.28}$$

where $\hat{s} = \hat{s}(\epsilon) \to 0$ as $\epsilon \to 0$. Then as $\epsilon \to 0$, we obtain

$$u(r,t) = \frac{1}{4\pi Dt}\exp\left(-\frac{r^2}{4Dt}\right) = \frac{1}{4\pi Dt}\exp\left(-\frac{x^2+y^2}{4Dt}\right), \tag{5.5.29}$$

on using the result (5.5.27) and $I_0(0) = 1$. This is the *fundamental solution of the diffusion equation* in the two-dimensional case. It should be compared with the one-dimensional fundamental solution $G(x,t)$ defined in (5.2.39).

If the initial value $f(r)$ is concentrated on a circle $r = a$, we require that $2\pi\int_0^\infty f(r)$ $r\,dr = 1$, so that the concentrated source has unit strength. To determine the solution in this case, we assume that $f(r) \geq 0$ and vanishes outside the interval $(a-\epsilon, a+\epsilon)$, thereby reducing the integral (5.5.26) to one with finite limits. Proceeding as in the point source problem, we obtain the solution (on letting $\epsilon \to 0$)

$$u(r,t) = \frac{1}{4\pi Dt}\exp\left(-\frac{r^2+a^2}{4Dt}\right) I_0\left(\frac{ar}{2Dt}\right). \tag{5.5.30}$$

In the limit as a tends to zero, (5.5.30) reduces to (5.5.29).

Exercises 5.5

5.5.1. Consider the Fourier transform formulas (5.4.1)–(5.4.2) in the case $n = 2$. Let $x_1 = r\cos(\theta)$, $x_2 = r\sin(\theta)$, $\lambda_1 = \lambda\sin(\phi)$, $\lambda_2 = \lambda\cos(\phi)$, and put $f(x_1, x_2) = e^{-in\theta}f(r)$. Show that the transform $F(\lambda_1, \lambda_2)$ [i.e., (5.4.1)] takes the form

$$F(\lambda_1, \lambda_2) = \frac{1}{2\pi}\int_0^\infty \int_{-\pi}^{\pi} \exp[i\lambda r\sin(\theta+\phi) - in\theta]f(r)r\, d\theta\, dr.$$

From the integral representation $J_n(\lambda r) = (1/2\pi)\int_{-\pi}^{\pi}\exp[i\lambda r\sin(\theta) - in\theta]\, d\theta$ for the Bessel function $J_n(\lambda r)$ of integral order, conclude that $e^{-in\phi}F(\lambda_1, \lambda_2) = \int_0^\infty rJ_n(\lambda r)f(r)\, dr = F(\lambda)$. Insert $e^{-in\phi}F(\lambda)$ and $e^{-in\theta}f(r)$ into the inversion formula (5.4.2) to obtain

$$f(r) = \frac{1}{2\pi}\int_0^\infty \int_{-\pi}^{\pi} \exp[in(\theta+\phi) - i\lambda r\sin(\theta+\phi)]F(\lambda)\lambda\, d\phi\, d\lambda.$$

Use the integral representation of $J_n(\lambda r)$ given previously to derive (5.5.2).

5.5.2. Verify the result given in (5.5.4).

5.5.3. Integrate by parts in (5.5.5) to obtain the result $\int_0^\infty e^{-\lambda z}J_1(\lambda r)\, d\lambda = (1/r)\left(1 - z/\sqrt{z^2+r^2}\right)$. *Hint*: $J_0'(z) = -J_1(z)$.

5.5.4. Evaluate the integral (5.5.14) if $f(z) = 1$.

5.5.5. Use the zero-order Hankel transform to solve the Neumann problem for Laplace's equation, $u_{rr}(r,z) + \frac{1}{r}u_r(r,z) + u_{zz}(r,z) = 0$, $r > 0$, $z > 0$, $u(r,z) \to 0$ as $z \to \infty$, $u_z(r,0) = -1/\pi\epsilon^2$, $r < \epsilon$, $u_z(r,0) = 0$, $r > \epsilon$. Show that the transform is $U_0(\lambda, z) = (1/\pi\epsilon\lambda^2)J_1(\lambda\epsilon)e^{-\lambda z}$. Let $\epsilon \to 0$ and show that the solution $u(r,z) \to (1/2\pi)(r^2+z^2)^{-1/2}$ in the limit.

5.5.6. Use the Hankel transform to solve the problem $u_{rr}(r,z) + u_r(r,z)/r + u_{zz}(r,z) - k^2u(r,z) = 0$, $r > 0$, $-\infty < z < \infty$, with the boundary conditions (5.5.9) given on the z-axis.

5.5.7. Use the Hankel transform to construct a (formal) solution of the following problem: $u_t(r,t) = D\left(u_{rr}(r,t) + \frac{1}{r}u_r(r,t)\right) + F(r,t)$, $r > 0$, $t > 0$, where $D > 0$ and $u(r,0) = f(r)$.

5.5.8. Solve the Cauchy problem for the damped wave equation, $u_{tt}(r,t) + a^2u_t(r,t) - c^2(u_{rr}(r,t) + \frac{1}{r}u_r(r,t)) = 0$, $r > 0$, $t > 0$, with the initial conditions $u(r,0) = f(r)$, $u_t(r,0) = g(r)$, $r > 0$, by using the Hankel transform.

5.5.9. Consider the Cauchy problem for the wave equation: $u_{tt}(r,\theta,t) - c^2(u_{rr}(r,\theta,t) + \frac{1}{r}u_r(r,\theta,t) + \frac{1}{r^2}u_{\theta\theta}(r,\theta,t)) = 0$, with $r > 0$, $0 \le \theta < 2\pi$, $t > 0$, and the initial conditions $u(r,\theta,0) = f(r)\cos(n\theta)$, $u_t(r,\theta,0) = g(r)\cos(n\theta)$, with n as a positive integer. Let $u(r,\theta,t) = v(r,t)\cos(n\theta)$. Solve the resulting Cauchy problem for $v(r,t)$ by means of the nth-order Hankel transform.

5.5.10. Consider the diffusion equation in three dimensions. Introduce spherical coordinates and if the initial value is given as $u(r,0) = f(r)$, where r is the radial variable, show that $u(r,t)$ satisfies the equation $u_t(r,t) = D\left(u_{rr}(r,t) + \frac{2}{r}u_r(r,t)\right)$, $r > 0$, $t > 0$, with D as the diffusion constant. This problem cannot be solved by means of the Hankel transform. However, let $v(r,t) = ru(r,t)$ and show that $v(r,t)$ satisfies the one-dimensional diffusion equation with $v(0,t) = 0$ and $v(r,0) = rf(r)$. Use the results of Example 5.5 to solve this problem and obtain the solution of the given problem for $u(r,t)$. As in Example 5.9, let $f(r)$ represent a concentrated source at $r = 0$, and, in analogy with (5.5.27), suppose that $\lim_{\epsilon\to 0} 4\pi \int_0^\epsilon f(r)r^2\, dr = 1$. Determine the limit of the solution $u(r,t)$ as $\epsilon \to 0$. This limit is the fundamental solution of the diffusion equation in the three-dimensional case.

5.6 LAPLACE TRANSFORMS

The solution of initial and boundary value problems for *hyperbolic* and *parabolic equations* by the use of the transform methods presented previously in this chapter involves the transformation of the spatial variable(s) in the problems, as we have seen. The problems are thus reduced to the solution of initial value problems for the transformed dependent variables. The *Laplace transform*, however, solves these problems by acting on the time variable in the hyperbolic and parabolic equations and yields boundary value problems for the transformed dependent variables. (In its application to Cauchy problems, the boundary conditions are given at infinity.) Even though the nature of their applications is rather different, the Laplace transform and its inversion formula can be derived from the corresponding formulas (5.2.6) and (5.2.7) for the Fourier transform and its inverse. This is the approach that we use. As the Laplace transform generally acts on the time variable, we formulate its properties in terms of the variable t.

Given a function $f(t)$ defined for $t \geq 0$ and integrable over any finite interval (say, it is piecewise continuous), we extend its definition to $t < 0$ by putting $f(t) = 0$ for $t < 0$. [It is not required that $f(0) = 0$.] We assume that $f(t)$ is of exponential order as $t \to \infty$. That is, there exist real constants $c > 0$, k, and M such that $|f(t)| \leq Me^{kt}$ for $t \geq c$. Let α be an arbitrary constant with $\alpha > k$. Then the Fourier transform $\hat{F}(\hat{\lambda})$ of $e^{-\alpha t}f(t)$ is given as

$$\hat{F}(\hat{\lambda}) = \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{(i\hat{\lambda}-\alpha)t} f(t)\, dt, \tag{5.6.1}$$

since $f(t)$ vanishes for negative t. (The notation for the Fourier transform has been changed slightly.) Let $\lambda = \alpha - i\hat{\lambda}$ in (5.6.1) and we have

$$F(\lambda) = \int_0^\infty e^{-\lambda t} f(t)\, dt, \tag{5.6.2}$$

where $F(\lambda) = (2\pi)^{1/2}\hat{F}[i(\lambda - \alpha)]$. From the expression (5.2.7) for the inverse Fourier transform, we obtain

$$e^{-\lambda t}f(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-i\hat{\lambda}t}\hat{F}(\hat{\lambda})\,d\hat{\lambda}. \tag{5.6.3}$$

Transforming again to the variable $\lambda = \alpha - i\hat{\lambda}$, we find that $\hat{\lambda} = \mp\infty$ corresponds to $\lambda = \alpha \pm i\infty$ and that $d\hat{\lambda} = -(1/i)\,d\lambda$. Therefore, the integration path along the real $\hat{\lambda}$-axis in (5.6.3) is transformed into a path along the line L in the complex λ-plane parallel to the imaginary axis. With $\lambda = \lambda_1 + i\lambda_2$, L is given as $\lambda = \alpha + i\lambda_2$, $-\infty < \lambda_2 < \infty$ and has the direction shown in Figure 5.1. In terms of λ, the inversion formula (5.6.3) takes the form

$$f(t) = \frac{1}{2\pi i}\int_{L} e^{\lambda t}F(\lambda)\,d\lambda, \tag{5.6.4}$$

since the term $e^{-\alpha t}$ cancels out.

The formula (5.6.2) determines the *Laplace transform* $F(\lambda)$ of $f(t)$. The *complex inversion formula* for the Laplace transform that determines $f(t)$ in terms of $F(\lambda)$ is given by (5.6.4). For a function $f(t)$ that satisfies the conditions given above, the formulas (5.6.2) and (5.6.4) are valid for all λ such that $\mathrm{Re}[\lambda] = \lambda_1 = \alpha > k$.

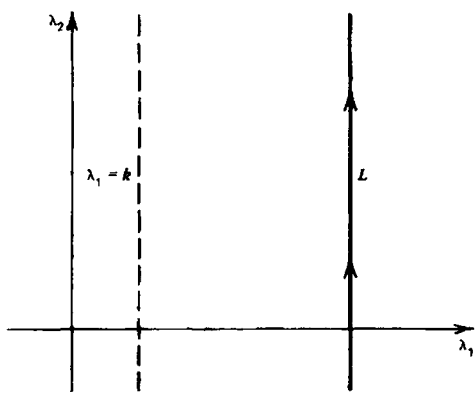

Figure 5.1 The line L.

General Properties

Given the transform $F(\lambda)$, the function $f(t)$ can often be determined from (5.6.4) by using the *residue theory of complex variables*. Otherwise, approximation methods are available to deal with (5.6.4). The inversion formula (5.6.4) will not play a (significant) role in our discussion of the Laplace transform. Laplace transforms are tabulated extensively (Maple contains a large number of built-in transforms in its

integral transform file). Additionally, it is often possible to invert Laplace transforms by reducing $F(\lambda)$ to a form for which the inverse transform is known. In this respect the linearity of the Laplace transform and the convolution theorem (see Exercise 5.6.3) play an important role. Thus if $F(\lambda)$ can be expressed as a finite (or even an infinite) sum, each of whose terms has a known inverse transform, we can invert $F(\lambda)$ given that the inverse of the (infinite) sum is the sum of the inverses. It is not always a straightforward matter to carry out this procedure, as will be seen.

The usefulness of the Laplace transform in solving problems for ODEs and PDEs that involve initial conditions is brought out by the following result, which relates the transform of derivatives to the transform of the given function. If $f^{(n)}(t) = d^n f(t)/dt^n$, we have

$$\int_0^\infty e^{-\lambda t} f^{(n)}(t)\, dt = \lambda^n \int_0^\infty e^{-\lambda t} f(t)\, dt - \lambda^{n-1} f(0) - \cdots - f^{(n-1)}(0), \quad (5.6.5)$$

on using integration by parts. We see that the transform of the nth derivative of $f(t)$ is given in terms of the transform of $f(t)$ and n initial values of $f(t)$ and its derivatives at $t = 0$. By contrast, the sine or cosine transform of an even-order nth derivative of $f(x)$ involves only $n/2$ values of $f(x)$ and its derivatives at $x = 0$. Additional general properties of Laplace transforms are given in the exercises, in Section 5.8, and in tables of transforms.

Applications to PDEs

Example 5.10. The Heat Equation in a Finite Interval. We again consider the *heat equation* in a finite interval,

$$u_t(x,t) - c^2 u_{xx}(x,t) = 0, \qquad 0 < x < l,\ t > 0, \quad (5.6.6)$$

where c^2 is a constant, with the boundary and initial conditions

$$u(0,t) = 0,\ u(l,t) = 0, \quad t > 0, \quad u(x,0) = f(x) \quad 0 < x < l. \quad (5.6.7)$$

This problem was solved by separation of variables in Example 4.3. However, we now obtain the solution in a different form and compare both results.

The problem (5.6.6)–(5.6.7) is solved by applying the Laplace transform in the t-variable. We multiply in (5.6.6) by $e^{-\lambda t}$ and integrate with respect to t from 0 to ∞. Defining the transform of $u(x,t)$ as

$$U(x,\lambda) = \int_0^\infty e^{-\lambda t} u(x,t)\, dt, \quad (5.6.8)$$

and using (5.6.5) and transforming the boundary conditions (5.6.7) yields

$$\lambda U(x,\lambda) - f(x) - c^2 \frac{\partial^2 U(x,\lambda)}{\partial x^2} = 0,\ 0 < x < l,\ U(0,\lambda) = U(l,\lambda) = 0. \quad (5.6.9)$$

The solution of the boundary value problem (5.6.9) is found to be

$$U(x,\lambda) = \frac{1}{c\sqrt{\lambda}\sinh(\sqrt{\lambda}\,l/c)} \int_0^x \sinh\left(\frac{\sqrt{\lambda}}{c}s\right) \sinh\left(\frac{\sqrt{\lambda}}{c}(l-x)\right) f(s)\,ds$$

$$+\frac{1}{c\sqrt{\lambda}\sinh(\sqrt{\lambda}\,l/c)} \int_x^l \sinh\left(\frac{\sqrt{\lambda}}{c}x\right) \sinh\left(\frac{\sqrt{\lambda}}{c}(l-s)\right) f(s)\,ds. \quad (5.6.10)$$

This result may be obtained by using the Green's function or variation of parameters method for ordinary differential equations.

The inversion of the transform $U(x,\lambda)$ is not straightforward, even if the inversion formula (5.6.4) is used, unless the transform is simplified first. In view of the discussion preceding (5.6.4), we may assume that λ (or $|\lambda|$) is not small. Consequently, we have the following result:

$$\frac{1}{\sinh(\sqrt{\lambda}\,l/c)} = \frac{2}{\exp(\sqrt{\lambda}\,l/c) - \exp(-\sqrt{\lambda}\,l/c)} = \frac{2\,\exp(-\sqrt{\lambda}\,l/c)}{1-\exp(-2\sqrt{\lambda}\,l/c)}$$

$$= 2\exp\left(-\frac{\sqrt{\lambda}\,l}{c}\right) \sum_{k=0}^{\infty} \exp\left(-\frac{2\sqrt{\lambda}\,kl}{c}\right), \quad (5.6.11)$$

on using the series $1/(1-x) = \sum_{k=0}^{\infty} x^k$, which is valid for $|x| < 1$, since it may be assumed that λ is large enough so that $\exp(-2\sqrt{\lambda}\,l/c) \ll 1$. The hyperbolic functions in the integrals (5.6.10) can also be expressed in terms of exponentials. We combine the exponentials and assume that in the inversion process, summation and integration, as well as orders of integration, can be reversed (this can be shown to be valid for this problem). Further, we find from a table of Laplace transforms that

$$\int_0^{\infty} e^{-\lambda t}\left[\frac{1}{\sqrt{\pi t}}\exp\left(\frac{-\alpha^2}{4t}\right)\right] dt = \frac{1}{\sqrt{\lambda}}e^{-\alpha\sqrt{\lambda}}, \qquad \alpha > 0. \quad (5.6.12)$$

Using (5.6.12), the solution $u(x,t)$ can then be written as (after a tedious but straightforward manipulation of the resulting series expansions)

$$u(x,t) = \int_0^l \sum_{k=-\infty}^{\infty} \{G(x-s-2kl,t) - G(x+s-2kl,t)\} f(s)\,ds, \quad (5.6.13)$$

where $G(x-s,t)$ is the fundamental solution of the heat equation as in (5.2.39).

The solution (5.6.13) has a form quite different from that given in terms of Fourier sine series in Example 4.3. By the uniqueness theorem for the initial and boundary value problem for the heat equation proven in Chapter 8, both solutions must be identical. The equivalence of both sums is demonstrated in Exercise 7.5.15 based on the *Poisson summation formula*.

We remark that both solution forms are useful in evaluating the solution $u(x,t)$ for different ranges of values of t. Because of the rapid exponential decay of the terms in

the Fourier series (4.4.29), the series converges rapidly for large values of t. However, the series of fundamental solutions converges most rapidly for small values of t since all the terms in the series except $G(x-s,t)$, which has delta function behavior in $0<x<l$, vanish at $t=0$. Thus for the purposes of computation each of the series is useful.

Example 5.11. The Wave Equation in a Finite Interval. We consider an initial and boundary value problem for the *wave equation* in a finite interval. We have

$$u_{tt}(x,t)-c^2u_{xx}(x,t)=0, \qquad 0<x<l,\ t>0, \tag{5.6.14}$$

with the homogeneous initial conditions and the boundary conditions

$$u(x,0)=0,\ u_t(x,0)=0, \quad u(0,t)=0,\ u(l,t)=g(t). \tag{5.6.15}$$

This problem can be solved by the method of finite sine transforms given in Section 4.6, but we now obtain the solution in a different form by using the Laplace transform to solve the problem.

Let $U(x,\lambda)$, defined as in (5.6.8), denote the Laplace transform of the solution $u(x,t)$. We transform (5.6.14) and use (5.6.5) and the initial conditions (5.6.15) to obtain

$$\lambda^2U(x,\lambda)-c^2\frac{\partial^2U(x,\lambda)}{\partial x^2}=0,\ 0<x<l, \quad U(0,\lambda)=0,\ U(l,\lambda)=G(\lambda), \tag{5.6.16}$$

where $G(\lambda)$ is the Laplace transform of $g(t)$. The solution of the boundary value problem for $U(x,\lambda)$ is

$$U(x,\lambda)=G(\lambda)\frac{\sinh[(\lambda/c)x]}{\sinh[(\lambda/c)l]}. \tag{5.6.17}$$

To invert this transform we proceed as in Example 5.10, expand $U(x,\lambda)$ for large λ, and express the hyperbolic functions in terms of exponentials. Then

$$U(x,\lambda)=\sum_{k=0}^{\infty}G(\lambda)\left\{\exp\left[\frac{\lambda}{c}(x-l-2kl)\right]-\exp\left[-\frac{\lambda}{c}(x+l+2kl)\right]\right\}. \tag{5.6.18}$$

The inverse transform of each term in this series can be found from the result of Exercise 5.6.4, and the solution to the given problem is

$$u(x,t) = \sum_{k=0}^{\infty} H\left(t - \frac{1}{c}[(1+2k)l - x]\right) g\left(t - \frac{1}{c}[(1+2k)l - x]\right)$$

$$-H\left(t - \frac{1}{c}[(1+2k)l + x]\right) g\left(t - \frac{1}{c}[(1+2k)l + x]\right), \quad (5.6.19)$$

where $H(z)$ is the Heaviside function.

The form of this solution differs from that obtained by means of the finite sine transform. That solution involves trigonometric functions and effectively represents $u(x,t)$ as a sum of *standing waves*. The solution form (5.6.19) is given in terms of *traveling waves*. The occurrence of the Heaviside functions signifies that at given times additional right- and left-traveling waves or wave fronts are generated as a result of reflection from the boundary. (A direct method for constructing this type of solution is given in Example 6.7.) Clearly, this form of the solution is only useful for small t. For large t, the standing wave representation is preferable.

Abelian and Tauberian Theories

As we have seen in the foregoing examples, it is not always a simple matter to invert a Laplace transform resulting from the solution of a partial differential equation. Consequently, it is of interest to examine methods that yield approximate expressions for inverse Laplace transforms. The *Abelian* and *Tauberian asymptotic theories* for Laplace transforms that relate the values of the transform $F(\lambda)$ as $\lambda \to \infty$ and $\lambda \to 0$ to the values of the function $f(t)$ as $t \to 0$ and $t \to \infty$, respectively, are useful in that regard. Given the presence of the exponentially decaying term in the formula (5.6.2) for $F(\lambda)$ and assuming that $f(t)$ is reasonably behaved, we expect that the main contributions to the integral for $F(\lambda)$ come from small values of λt. Thus if λ is large, t must be small and vice versa.

A useful *Abelian result* states that if

$$f(t) \approx \sum_{n=0}^{\infty} \alpha_n t^n, \qquad t \to 0, \quad (5.6.20)$$

then for the transform $F(\lambda)$ we have

$$F(\lambda) \approx \sum_{n=0}^{\infty} \alpha_n \frac{n!}{\lambda^{n+1}}, \qquad \lambda \to \infty. \quad (5.6.21)$$

The converse is also true, so that (5.6.21) implies (5.6.20). The result (5.6.21) can be obtained by formally substituting (5.6.20) in the integral (5.6.2) for $F(\lambda)$ and integrating term by term. The series (5.6.20) and (5.6.21) need not converge, but may have only asymptotic validity. This means that for small t and large λ, $f(t)$ and $F(\lambda)$ are well approximated by the leading terms in the given series even if the series diverge. (See Section 5.7 and Chapter 9 for a more complete discussion of the

meaning of asymptotic equalities and series.) In the theory of *asymptotic expansions of integrals*, the relation between $F(\lambda)$ and $f(t)$ given by (5.6.21) and (5.6.20) is a special case of a general result known as *Watson's lemma*. The converse, which is more useful for our purposes, is attributed to Heaviside.

A simple application of the foregoing Abelian result can be given in connection with the initial value problems formulated in Exercises 3.4.2 and 3.4.4 for the heat and telegrapher's equations. If these problems are solved by means of the Laplace transform and the transform function is expanded in inverse powers of λ as in (5.6.21), the power series expansions of the inverse transforms coincide with the results obtained in these exercises. The expansions obtained in the two examples in this section are also Abelian results that are valid and useful for small t, since they were obtained by expanding the transform for large λ. However, the solutions are not in the form of power series in t, since the large λ expansion does not have the form (5.6.21).

For small values of λ, a useful *Tauberian result* which requires a knowledge of complex variables theory for its application is as follows. Let $F(\lambda)$ be analytic for $\mathrm{Re}[\lambda] > k$. Assume further that $F(\lambda)$ has isolated real and simple poles in the finite λ-plane located at the set of points $\{\lambda_n\}$ with $\lambda_1 > \lambda_2 > \lambda_3 > \cdots$. With $f(t)$ assumed to be real, suppose that $F(\lambda)$ has (real) residues $\{\beta_n\}$ at the poles $\{\lambda_n\}$. Then we have

$$f(t) \approx \sum_{n=1}^{\infty} \beta_n e^{\lambda_n t}, \tag{5.6.22}$$

which is valid as $t \to \infty$. This result is a special case of a general technique for inverting Laplace transforms known as the *Heaviside expansion theorem*. To apply this theorem, we suppose that the transform function $F(\lambda)$ can be written as $F(\lambda) = G(\lambda)/H(\lambda)$, where $G(\lambda)$ is an entire function of λ and $H(\lambda)$ has simple zeros at the points $\lambda = \lambda_n$. Thus, $F(\lambda)$ has only simple poles at $\lambda = \lambda_n$, assuming that $G(\lambda_n) \neq 0$. (They need not be real valued.) Then, on using residue theory to evaluate the inversion integral (5.6.4), we obtain

$$F(\lambda) = \sum_{n=1}^{\infty} \frac{G(\lambda_n)}{H'(\lambda_n)} e^{\lambda_n t}. \tag{5.6.23}$$

The residue of $F(\lambda)$ at λ_n is given by $G(\lambda_n)/H'(\lambda_n)$. [The series (5.6.23) may have only asymptotic validity.] If the poles λ_n satisfy the additional conditions given above, we obtain a Tauberian result in that the terms in the series (5.6.23) grow or decay exponentially, so that $f(t)$ is well approximated by the leading terms in the series for large t. However, if the λ_n are all pure imaginary numbers, for example, each of the exponentials in (5.6.23) has unit magnitude, so that no simplification results for large values of t.

To obtain a Tauberian result for Example 5.10 we note that the transform function $U(x, \lambda)$ is apparently singular at the zeros of $\sinh(\sqrt{\lambda}\, l/c)$. These are located at the points $\lambda_n = -(\pi c/l)^2 n^2$ $(n = 0, 1, 2, \ldots)$. It is easily seen, however, that $\lambda_0 = 0$ is not a singular point of $U(x, \lambda)$. Since the zeros $\{\lambda_n\}$ $(n \geq 1)$ are all simple, $U(x, \lambda)$ has simple poles at these points. The λ_n satisfy the condition $\lambda_1 > \lambda_2 > \lambda_3 > \cdots$ and it can be verified that the residues $\{\beta_n\}$ of $U(x, \lambda)$ at $\{\lambda_n\}$ are all real, so that the

conditions for the aforementioned Tauberian theorem are met. The Tauberian result $u(x,t) = \sum_{n=1}^{\infty} \beta_n(x) e^{\lambda_n t}$ can be shown to be identical with the Fourier sine series (4.4.29). Again, that series is convergent and not merely asymptotic and is useful for large t.

In Example 5.11 the poles of the transform function (5.6.17) are determined from the zeros of $\sinh[(\lambda/c)l]$ and the poles of $G(\lambda)$. The hyperbolic function has simple zeros at the points $\lambda_n = i\pi cn/l$ $(n = \pm 1, \pm 2, \ldots)$ ($\lambda_0 = 0$ does not contribute a pole). If we assume that the poles of $G(\lambda)$ are all simple and do not coincide with the λ_n, the expansion form (5.6.23) (appropriately modified) is valid for this problem. However, if the poles of $G(\lambda)$ have a nonpositive real part, the expansion does not yield a useful Tauberian result. The pure imaginary zeros of the hyperbolic sine function contribute the most significant terms for large t, and none of these terms decays in magnitude as t tends to infinity.

Exercises 5.6

5.6.1. Determine the Laplace transforms of the following functions: (a) $f(t) = 1$; (b) $f(t) = e^{at}$; (c) $f(t) = \cos \omega t$; (d) $f(t) = \sin \omega t$; (e) $f(t) = t^n/n!$; (f) $f(t) = \delta(t - a)$, $a > 0$.

5.6.2. Show that if $F(\lambda)$ is the Laplace transform of $f(t)$, the Laplace transform of $e^{-at} f(t)$ is $F(\lambda + a)$.

5.6.3. Let $f(t)$ and $g(t)$ have Laplace transforms $F(\lambda)$ and $G(\lambda)$, respectively. Show that the Laplace transform of $\int_0^t f(s) g(t - s)\, ds$ is given as $F(\lambda)G(\lambda)$. (This is the *convolution theorem* for the Laplace transform.)

5.6.4. Show that the Laplace transform of $H(t - a) f(t - a)$, $a > 0$, where $H(t)$ is the Heaviside function, is $F(\lambda) e^{-a\lambda}$.

5.6.5. Show that the Laplace transform of $\int_0^t f(s)\, ds$ is given as $F(\lambda)/\lambda$.

5.6.6. Let $f(t; \alpha)$ have the Laplace transform $F(\lambda; \alpha)$, where α is a parameter. Show that (a) $\partial f(t; \alpha)/\partial \alpha$ has the transform $\partial F(\lambda; \alpha)/\partial \alpha$; (b) $\int^{\alpha} f(t; s)\, ds$ has the transform $\int^{\alpha} F(\lambda; s)\, ds$.

5.6.7. Use (5.6.12) and Exercise 5.6.6 to determine that the Laplace transform of $[\alpha/(2\sqrt{\pi}\, t^{3/2})] e^{-\alpha^2/4t}$ is $e^{-\alpha\sqrt{\lambda}}$

5.6.8. Use (5.6.12) and Exercise 5.6.6 to determine that the Laplace transform of $\operatorname{erf}(\alpha/2\sqrt{t})$ is $(1 - e^{-\alpha\sqrt{\lambda}})/\lambda$.

5.6.9. Use (5.6.12) and Exercise 5.6.6 to determine that the Laplace transform of $\operatorname{erfc}(\alpha/2\sqrt{t})$ is $(1/\lambda) e^{-\alpha\sqrt{\lambda}}$.

5.6.10. Determine the Laplace transform of the zero-order Bessel function $J_0(at)$. *Hint*: Use the integral representation given in Exercise 5.7.1.

5.6.11. Solve, using the Laplace transform: $u_t(x,t) - c^2 u_{xx}(x,t) = 0,\ 0 < x < \infty,\ t > 0,\quad u(x,0) = f(x),\ u(0,t) = 0.$

5.6.12. Solve, using the Laplace transform: $u_t(x,t) - c^2 u_{xx}(x,t) = 0,\ 0 < x < \infty,\ t > 0,\quad u(x,0) = 0,\ u(0,t) = g(t).$

5.6.13. Solve, using the Laplace transform: $u_t(x,t) - c^2 u_{xx}(x,t) = 0,\ 0 < x < \infty,\ t > 0,\quad u(x,0) = 0,\ u_x(0,t) = h(t).$

5.6.14. Use the Laplace transform to solve the Cauchy problem $u_t(x,t) - c^2 u_{xx}(x,t) = 0,\ -\infty < x < \infty,\ t > 0,\quad u(x,0) = \delta(x-a).$

5.6.15. Use the Laplace transform to solve: $u_t(x,t) - c^2 u_{xx}(x,t) = 0,\ 0 < x < l,\ t > 0, u(0,t) = e^{-t},\ u_x(l,t) = 0,\ t > 0,\quad u(x,0) = 0,\ 0 < x < l.$

5.6.16. Solve, using the Laplace transform, $u_{tt}(x,t) - c^2 u_{xx}(x,t) = 0,\ 0 < x < \infty,\ t > 0, u(0,t) = f(t),\ t > 0,\quad u(x,0) = u_t(x,0) = 0,\ 0 < x < \infty.$

5.6.17. Solve the problem of Exercise 5.6.16 if the Dirichlet boundary condition is replaced by the Neumann condition $u_x(0,t) = -g(t)$.

5.6.18. Use the Laplace transform to solve the Cauchy problem $u_{tt}(x,t) - c^2 u_{xx}(x,t) = 0,\ -\infty < x < \infty,\ t > 0, u(x,0) = 0,\ u_t(x,0) = \delta(x-a),\ -\infty < x < \infty.$

5.6.19. Solve the Cauchy problem for the telegrapher's equation using the Laplace transform: $u_{tt}(x,t) - \gamma^2 u_{xx}(x,t) + 2\hat{\lambda} u_t(x,t) = 0,\ -\infty < x < \infty,\ t > 0,$ $u(x,0) = 0,\ u_t(x,0) = g(x),\ -\infty < x < \infty$. *Hint*: Eliminate the u_t term by a transformation of the dependent variable. Then, make use the fact that the Laplace transform of $I_0(a\sqrt{t^2 - b^2})H(t-a)$ is given as $(\lambda^2 - a^2)^{-1/2}\exp(-b\sqrt{\lambda^2 - a^2})$, where $I_0(z)$ is the modified Bessel function of zero order and $H(z)$ is the Heaviside function.

5.6.20. Solve the initial and boundary value problem for the telegrapher's equation of Exercise 5.6.19 with $x > 0$ and $t > 0$ if $u(x,0) = u_t(x,0) = 0$ and $u_x(0,t) = -f(t)$.

5.6.21. Verify the relationship (5.6.20)–(5.6.21) by formally integrating term by term in the formula (5.6.2).

5.6.22. Use the Laplace transform to obtain the exact solution of the following problem: $u_t(x,t) - c^2 u_{xx}(x,t) = 0,\ 0 < x < l,\ t > 0, u(0,t) = u(l,t) = 0,\ t > 0,\quad u(x,0) = x + \sin(3\pi x/l),\ 0 < x < l$. Use the Abelian result (5.6.20)–(5.6.21) to discuss the solution of this problem for small t on the basis of the behavior of the Laplace transform. Compare your results with those obtained from the exact solution.

5.6.23. Suppose that the Laplace transform $F(\lambda)$ is a rational function; that is, $F(\lambda) = G(\lambda)/H(\lambda)$ where $G(\lambda)$ and $H(\lambda)$ are polynomials. If $H(\lambda)$ has only simple zeros, show that the partial fraction expansion of $F(\lambda)$ yields a (finite) Heaviside expansion (5.6.23) for the inverse transform if $G(\lambda)$ is a polynomial of lower degree than $H(\lambda)$.

5.6.24. Invert the transform (5.6.17) by use of the convolution theorem and the Heaviside expansion of the ratio of hyperbolic functions. Show that the resulting series solution agrees with that obtained by using finite sine transforms to solve the problem in Example 5.11.

5.6.25. Let $g(t) = e^{-t}$ in (5.6.15) and invert the resulting transform (5.6.17) by a direct application of the Heaviside expansion (5.6.23).

5.6.26. Demonstrate that the Laplace transform of the solution of the following problem satisfies the conditions required for the validity of the Tauberian result given in the text. $u_t(x,t) - c^2 u_{xx}(x,t) = 0,\ 0 < x < l,\ t > 0,\ u(0,t) = u(l,t) = 1,$ $t > 0, u(x,0) = 0,\ 0 < x < l$. Use the Tauberian result (5.6.22) to obtain an expansion of the solution $u(x,t)$ for large t and show that it is, in fact, an exact solution of the given problem.

5.6.27. Solve, using the Laplace transform: $u_{tt}(x,t) - c^2 u_{xx}(x,t) = e^{-t},\ 0 < x < \infty,\ t > 0, u(0,t) = \sin(t),\ t > 0, \quad u(x,0) = u_t(x,0) = 1,\ 0 < x < \infty.$

5.7 ASYMPTOTIC APPROXIMATION METHODS FOR FOURIER INTEGRALS

In this section we discuss two useful *approximation methods* for evaluating integrals that result from an application of *Fourier transform* techniques. We give only a brief discussion of these methods and apply them to two problems relating to dispersive and dissipative wave motion (these concepts were defined in Section 3.5). We refer to the literature for additional details regarding the validity and applicability of these and other asymptotic methods for evaluating integrals. We say that the function $f(x)$ is *asymptotic* to the function $g(x)$ as $x \to \alpha$ and denote the *asymptotic equality* by $f(x) \approx g(x)$ if the ratio $f(x)/g(x)$ tends to unity as $x \to \alpha$.

We first consider the *method of stationary phase*, which we apply to a solution of the Cauchy problem for the Klein-Gordon equation representative of *dispersive wave motion*. Then we consider a method developed by *Sirovich* (which is related to the Tauberian theory for Laplace transforms of Section 5.6) and apply it to a solution of the Cauchy problem for the telegrapher's or dissipative wave equation representative of *dissipative wave motion*. (Some properties of both these equations were given in Example 3.7.) Neither of these methods requires complex variables theory for its application.

It was seen in our discussion of transform methods in this chapter that one cannot, in general, expect to evaluate integrals resulting from transform methods in closed form. Consequently, the methods discussed are essential, in that they yield simple and useful representations of the solutions.

Method of Stationary Phase

The *method of stationary phase* is an asymptotic approximation method for evaluating integrals of the form

$$I(k) = \int_{-\infty}^{\infty} e^{ik\phi(t)} f(t)\, dt, \tag{5.7.1}$$

where k is a large real parameter and $f(t)$ is a real- or complex-valued function, referred to as the *amplitude term*. The real-valued function $\phi(t)$ is called the *phase term*. [The case of finite endpoints in the integral (5.7.1) is considered below.] Since k is large and $\phi(t)$ real valued, the integrand in (5.7.1) oscillates rapidly and cancellation occurs over most of the domain of integration. However, near the points t where $\phi'(t) = 0$, that is, the *stationary points* of the phase, the cancellation is reduced significantly because of the diminished oscillation of the integrand. Thus the major contribution to the integral comes from the neighborhood of these points.

We assume that $\phi(t)$ and $f(t)$ are smooth functions and proceed as follows. Let t_0 be an isolated stationary point and assume that $\phi''(t_0) \neq 0$ and $f(t_0) \neq 0$. Expand $\phi(t)$ and $f(t)$ in a Taylor series around $t = t_0$ and obtain [since $\phi'(t_0) = 0$]

$$\phi(t) = \phi(t_0) + \frac{1}{2}\phi''(t_0)(t - t_0)^2 + \cdots, \qquad f(t) = f(t_0) + \cdots. \tag{5.7.2}$$

Only the terms given in (5.7.2) are retained in the (leading order) stationary phase approximation. Let $\epsilon > 0$ be a small number. Then the method of stationary phase asserts that the main contribution to the integral $I(k)$ for large k, if there are no other stationary points, is given as

$$\begin{aligned} I(k) &\approx \exp[ik\phi(t_0)] f(t_0) \int_{t_0-\epsilon}^{t_0+\epsilon} \exp\left[\frac{ik}{2}\phi''(t_0)(t - t_0)^2\right] dt \\ &\approx \exp[ik\phi(t_0)] f(t_0) \int_{-\infty}^{\infty} \exp\left[\frac{ik}{2}\phi''(t_0)(t - t_0)^2\right] dt, \end{aligned} \tag{5.7.3}$$

where we have replaced the finite limits $t_0 \mp \epsilon$ by $\mp\infty$ since the main contribution to the infinite integral (see the discussion below) comes from the neighborhood of t_0. The last integral in (5.7.3) can be evaluated and we obtain the asymptotic result

$$I(k) \approx e^{ik\phi(t_0)} f(t_0) \exp\left[\frac{i\pi}{4}\frac{\phi''(t_0)}{|\phi''(t_0)|}\right] \left[\frac{2\pi}{k|\phi''(t_0)|}\right]^{1/2}, \qquad k \to \infty. \tag{5.7.4}$$

The formula (5.7.4) represents the dominant contribution to the integral $I(k)$ for $k \gg 1$ if $t = t_0$ is the only stationary point. If there is more than one stationary point, there is a contribution of the form (5.7.4) from each such point and we sum all the contributions. If $\phi''(t_0) = 0$, so that there is a higher order stationary point, the procedure above fails and a different discussion which requires the expansion of $\phi(t)$ up to cubic or higher powers, is necessary to yield the contribution from that stationary point.

If the phase $\phi(t)$ has no stationary points, so that $\phi'(t) \neq 0$ for all t, we determine the behavior of (5.7.1) by integrating by parts. To do so, we replace the infinite limits in the integral (5.7.1) by the finite limits a and b. After integrating by parts in the modified integral, we let a and b tend to minus and plus infinity, respectively. On integrating by parts once, we obtain

$$I(k) \approx \frac{f(t)}{ik\phi'(t)} e^{ik\phi(t)}\Big|_{t=a}^{t=b} - \frac{1}{ik}\int_a^b \frac{d}{dt}\cdot\left[\frac{f(t)}{\phi'(t)}\right] e^{ik\phi(t)}\, dt. \tag{5.7.5}$$

If the endpoint contributions vanish as a and b tend to infinity, we conclude that $I(k)$ decays at least like $1/k$ as $k \to \infty$. [Even if the endpoint contributions do not vanish in the limit, we can transform (5.7.1) into a Fourier integral and conclude from the *Riemann-Lebesgue lemma* that the integral vanishes as $k \to \infty$, but perhaps at a slower rate than $1/k$.] If n integrations by parts can be carried out and at each step the endpoint contributions are zero, the integral $I(k)$ decays at least like $1/k^n$ as $k \to \infty$. Clearly, if an unlimited number of integrations by parts can be carried out, $I(k)$ is in effect exponentially small for large k.

If one or both of the endpoints in (5.7.1) are finite (we denote them by a and b, as above), we must include the contributions from these endpoints in our asymptotic approximation of $I(k)$. [We assume that the endpoints are not stationary points of $\phi(t)$.] The contributions from the endpoints are obtained from the integration by parts formula (5.7.5). Thus, if $f(t)$ does not vanish at the endpoints, their contribution is of order $1/k$ as $k \to \infty$. If $\phi(t)$ has no stationary points, the endpoint contributions are dominant. Otherwise the stationary point contribution, given in (5.7.4), which is of the order of $1/\sqrt{k}$, is dominant. [If an endpoint is also a stationary point, (5.7.4) multiplied by $\frac{1}{2}$ gives the asymptotic value of $I(k)$ at that point, as is easily seen.] If the function $f(t)$ in (5.7.1) is piecewise smooth, we break up the integral into a sum of integrals in each of which $f(t)$ is smooth and proceed as above. We do not discuss the case of a higher-order stationary point.

Dispersive PDEs: Klein-Gordon Equation

We now apply the *method of stationary phase* to the solution of the Cauchy problem for the *Klein-Gordon equation*,

$$u_{tt}(x,t) - \gamma^2 u_{xx}(x,t) + c^2 u(x,t) = 0, \qquad t > 0,\ -\infty < x < \infty, \tag{5.7.6}$$

with initial data $u(x,0)$ and $u_t(x,0)$ specified at $t = 0$. Let $U(\lambda,t)$ be the Fourier transform of $u(x,t)$, that is,

$$U(\lambda,t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{i\lambda x} u(x,t)\, dx. \tag{5.7.7}$$

The Fourier-transformed equation (5.7.6) is given as

$$\frac{\partial^2 U(\lambda,t)}{\partial t^2} + (\gamma^2\lambda^2 + c^2)U(\lambda,t) = 0, \qquad t > 0. \tag{5.7.8}$$

Introducing the Fourier transforms of the initial data, we easily determine that $U(\lambda, t)$ can be written as

$$U(\lambda, t) = F_+(\lambda) \exp(i\sqrt{\gamma^2\lambda^2 + c^2}\, t) + F_-(\lambda) \exp(-i\sqrt{\gamma^2\lambda^2 + c^2}\, t). \quad (5.7.9)$$

The terms $F_\pm(\lambda)$ can be specified in terms of the transforms of the initial data for $u(x, t)$. Since we are only interested in a general characterization of dispersive wave motion, we are not concerned with the specific form of the functions $F_\pm(\lambda)$. We do require, however, that $F_\pm(\lambda)$ be smooth functions. This would be the case if the initial data vanish sufficiently rapidly at infinity.

Inverting the Fourier transform $U(\lambda, t)$ gives

$$u(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F_+(\lambda) \exp\{i[\omega_+(\lambda)t - \lambda x]\}\, d\lambda$$
$$+ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F_-(\lambda) \exp\{i[\omega_-(\lambda)t - \lambda x]\}\, d\lambda, \quad (5.7.10)$$

where $\omega_\pm(\lambda) = \pm\sqrt{\gamma^2\lambda^2 + c^2} \equiv \pm\,\omega(\lambda)$. The equation $\omega = \omega(\lambda)$ is the *dispersion relation* for the Klein-Gordon equation, and the terms $\exp\{(i[\pm\omega(\lambda)t - \lambda x]\}$ are the *normal mode solutions*. (In comparing with the results in Example 3.7, the parameter k used there is replaced here by $-\lambda$.) The solution (5.7.10) is thus a superposition of normal mode solutions with the phase terms

$$\phi_\pm(x, t, \lambda) = \pm\,\omega(\lambda)t - \lambda x. \quad (5.7.11)$$

The *phase velocity*, that is, the velocity dx/dt for which $\phi_\pm(x, t, \lambda)$ remains constant, is given as

$$\frac{dx}{dt} = \pm\,\frac{\omega(\lambda)}{\lambda} = \pm\,\frac{\sqrt{\gamma^2\lambda^2 + c^2}}{\lambda}. \quad (5.7.12)$$

We see that for each real value of λ we have $|dx/dt| > \gamma$, so that the phase speed of the normal modes exceeds the characteristic or wave front speed γ of the Klein-Gordon equation. The *characteristic speed* represents the maximum speed of propagation of disturbances for hyperbolic equations. This property has already been demonstrated for the wave equation and is shown to be valid for more general hyperbolic equations, such as (5.7.6), later in the text. Thus the physical significance of the phase velocity must be examined more closely.

Noting that the general solution (5.7.10) of the Klein-Gordon equation is a superposition of the normal modes with phase terms (5.7.11), each of which has a different phase speed, it appears that the appropriate object to consider is not a single normal mode but a group of normal modes or a *wave packet*. A wave packet is obtained by superposing a collection of normal modes with wave numbers λ ranging over some interval. Since each mode has a different phase velocity, the group of normal modes disperses as it propagates and fails to retain its shape. We now show by using the method of stationary phase that although individual modes in a group travel at different velocities, the wave packet as a whole retains its group character and travels

with a velocity appropriately known as the *group velocity*. (It can be shown that an *energy* associated with the wave packet travels with group velocity.) We apply the stationary phase method to each of the integrals in (5.7.10), choosing x and t, which are parameters in the integrals, to be large. A single stationary point λ is found for each choice of x and t. The method determines that the major contribution to each integral comes from values of λ near the stationary point, so that we obtain a wave packet whose properties we examine.

As shown in Exercise 5.7.3, the $F_-(\lambda)$ integral in (5.7.10) is the complex conjugate of the $F_+(\lambda)$ integral. Thus it suffices to consider only the $F_+(\lambda)$ integral. The solution $u(x,t)$ equals twice the real part of that integral. Let

$$I(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F_+(\lambda) \exp[i(\sqrt{\gamma^2\lambda^2 + c^2}\, t - \lambda x)]\, d\lambda, \tag{5.7.13}$$

and assume that both x and t are large. If we were to set $x = \alpha t$, where α is constant, then either x or t would play the role of the large parameter k in the integral (5.7.1) so that the stationary phase method could be applied to (5.7.13). We do so without bringing the integral into the exact form (5.7.1).

The stationary points of the phase $\phi_+(x,t,\lambda)$ are determined from

$$\frac{d\phi_+}{d\lambda} = \frac{d}{d\lambda}[\omega(\lambda)\, t - \lambda x] = \omega'(\lambda)\, t - x = \frac{\gamma^2\lambda t}{\sqrt{\gamma^2\lambda^2 + c^2}} - x = 0. \tag{5.7.14}$$

This implies that

$$\frac{x}{t} = \omega'(\lambda) = \frac{\gamma^2\lambda}{\sqrt{\gamma^2\lambda^2 + c^2}}, \tag{5.7.15}$$

and the stationary value λ is given as

$$\lambda = \frac{c}{\gamma}\frac{x}{t}\left[\gamma^2 - \left(\frac{x}{t}\right)^2\right]^{-1/2}. \tag{5.7.16}$$

The expression $\omega'(\lambda) = \gamma^2\lambda/\sqrt{\gamma^2\lambda^2 + c^2}$ is known as the *group velocity*. As is seen from (5.7.15), there is a common stationary value λ for all (x,t) that satisfy the equation $x/t = \omega'(\lambda)$ [i.e., different values of x and t do not lead to different stationary values λ if x and t are related by $x = \omega'(\lambda)t$]. It follows from the stationary phase method that the solution retains a fixed character or form at a point x moving with velocity $\omega'(\lambda)$ if the wave packet with wave numbers λ near the stationary value is observed. We note that the group speed $|\omega'(\lambda)|$ is less than γ (the characteristic speed) for all values of λ. In fact, there are (real) stationary values λ only for x and t such that $\gamma^2t^2 - x^2 > 0$, as is seen from (5.7.16). If x and t are such that $\gamma^2t^2 - x^2 < 0$, the stationary phase result is $I(x,t) \approx 0$ since there are no stationary points and we assume that an unlimited number of integrations by part are possible.

The lines $x = \pm\gamma t$ represent the boundaries of the domain of influence of the point $(0,0)$ for the Klein-Gordon equation (this has already been shown for the wave equation and is shown to be valid for the Klein-Gordon equation in Chapter 7). Thus

in the context of the stationary phase result we have $I(x,t) \approx 0$ and, equivalently, the solution $u(x,t) = 0$ outside the domain of influence corresponding to a concentrated or point initial source at $x = 0$. This suggests that for large x and t, the solution $u(x,t)$ (under the foregoing assumptions on the initial data) appears to correspond to the solution of a problem with initial data concentrated at the origin. The group velocity lines $x = \omega'(\lambda)t$ lie within the domain of influence, as shown in Figure 5.2. (This contrasts with the phase velocity lines, all of which lie outside the domain of influence.) As the values x and t approach the characteristic lines $x = \pm\gamma t$, (5.7.16) shows that the stationary value λ tends to infinity. Then the stationary phase method, as presented, fails and a different discussion, which may be found in the literature, is necessary to discuss the solution near the characteristic lines. We observe that as $|\lambda| \to \infty$, (5.7.15) implies that $|\omega'(\lambda)| \to \gamma$, so that the group speed tends to the characteristic speed.

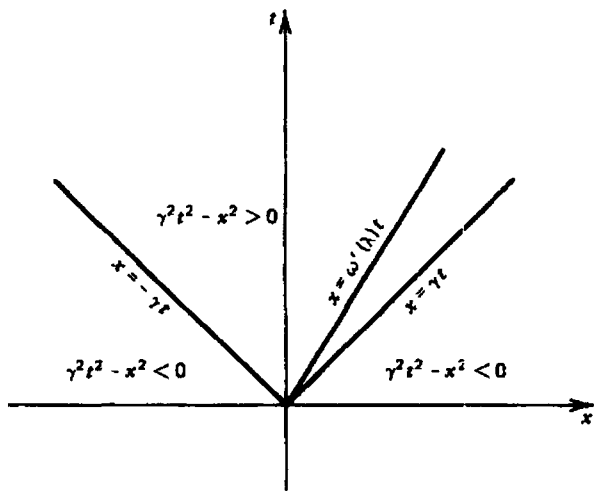

Figure 5.2 The group line.

To complete the determination of the asymptotic approximation of $I(x,t)$, we note that $\omega''(\lambda) = \gamma^2c^2(\gamma^2\lambda^2 + c^2)^{-3/2} > 0$ and $\phi''_+(\lambda) = \omega''(\lambda)t \geq 0$, so that $\phi''_+(\lambda)/|\phi''_+(\lambda)| = 1$. Consequently, we obtain from (5.7.4),

$$I(x,t) \approx \frac{1}{[t\omega''(\lambda)]^{1/2}} F_+(\lambda) \exp\left\{i\left[\omega(\lambda)t - \lambda x + \frac{\pi}{4}\right]\right\}$$

$$= \frac{1}{\sqrt{t}} H_+\left(\frac{x}{t}\right) \exp\left(i\frac{c}{\gamma}\sqrt{\gamma^2t^2 - x^2} + \frac{i\pi}{4}\right), \qquad (5.7.17)$$

where (5.7.16) was used to express λ as a function of x/t and we put $H_+ = F_+[\omega''(\lambda)]^{-1/2}$. This asymptotic formula determines the amplitude and the phase of a wave packet moving with the group velocity $dx/dt = \omega'(\lambda)$ $(-\infty < \lambda < \infty)$ and

is valid for $|x| < \gamma t$. The amplitude of the wave packet decays like $1/\sqrt{t}$ as $t \to \infty$. As $x \to \pm\gamma t$ or $|\lambda| \to \infty$, we see that $\omega''(\lambda) \to 0$, so that (5.7.17) becomes invalid near the characteristic lines, as was indicated. In the region $|x| > \gamma t$, that is, outside the characteristic sector indicated in Figure 5.2, we have $I(x,t) \approx 0$ if the conditions given above are met, since the integral has no stationary points.

We conclude by observing that the initial value problem for the Klein-Gordon can be solved exactly in terms of Bessel functions (as demonstrated in Chapter 7). However, the foregoing asymptotic characterization of dispersive wave propagation and the significance of the group velocity and wave packets are not readily apparent from the exact solution. Finally, we emphasize that (5.7.17) is valid only for large times t and that the full asymptotic solution of the given Cauchy problem equals twice the real part of (5.7.17).

Sirovich's Method

When dealing with Cauchy problems for *partial differential equations of dissipative type*, the Fourier integrals have a different form than those that occur when solving Cauchy problems for PDEs of dispersive type. The stationary phase method is inappropriate for these integrals and a different approach is needed to get asymptotic results. To this end, we introduce a method developed by Sirovich for approximately evaluating Fourier integrals of the type encountered on solving problems of dissipative type. This method differs from the method of stationary phase in that the integrands considered decay exponentially as the relevant parameter gets large. Thus the major contribution comes from the neighborhood of the point in the integration interval where the integrand has its maximum absolute value. In the stationary phase method, the integrand oscillates rapidly as the relevant parameter gets large and the main contributions come from the points where the oscillation is minimal. Sirovich's method is closely related to approaches used in approximating Laplace transform integrals, since for those integrals we also have exponential decay as the transform parameter gets large.

Sirovich's method deals with (Fourier-type) integrals of the form

$$I(x,t) = \int_{-\infty}^{\infty} F(\lambda) \exp[-g(\lambda)t - i\lambda x]\, d\lambda, \tag{5.7.18}$$

for which the following conditions are satisfied:

$$\begin{cases} \int_{-\infty}^{\infty} |F(\lambda)|\, d\lambda < M < \infty, \\ \max |F(\lambda)|_{\{\lambda\}} < M, \\ \text{Re}[g(\lambda)] \geq 0, \\ g(\lambda) = 0 \text{ if and only if } \lambda = 0, \\ g(\lambda) = i\alpha\lambda + \beta\lambda^2 + O(|\lambda|^3), \quad \alpha, \beta \text{ real}, \ \beta > 0, \ \lambda \to 0, \\ g(\lambda) \text{ continuous in } \lambda. \end{cases} \tag{5.7.19}$$

These conditions not only guarantee the convergence of the integral but indicate that for large t the main contribution to the integral comes from values of λ near the origin. Then as $t \to \infty$ we have

$$I(x,t) = \int_{-\infty}^{\infty} \exp[-\beta\lambda^2 t - i\alpha\lambda t - i\lambda x] F(\lambda)\, d\lambda + O\left(\frac{1}{t^{1-\delta}}\right), \qquad (5.7.20)$$

where δ is a small positive constant. The order term $O[\cdots]$ is understood to mean that if $F(x) = O[G(x)]$, then $|F(x)/G(x)| \to A$ for some constant A as $x \to \alpha$. Thus $O[1/t^{1-\delta}]$ means that as $t \to \infty$, the error term in (5.7.20) decays like $A/t^{1-\delta}$.

To simplify the integral in (5.7.20) we introduce a further approximation and replace $F(\lambda)$ by $F(0)$. We then obtain

$$I(x,t) \approx \int_{-\infty}^{\infty} \exp[-\beta\lambda^2 t - i\alpha\lambda t - i\lambda x] F(0)\, d\lambda. \qquad (5.7.21)$$

This integral can be evaluated exactly [see (5.2.33)–(5.2.37)] and we have

$$I(x,t) \approx \sqrt{\frac{\pi}{\beta t}}\, F(0) \exp\left[-\frac{(x+\alpha t)^2}{4\beta t}\right] \qquad t \to \infty. \qquad (5.7.22)$$

For proofs, further developments, and refinements of this method, we refer to Sirovich's work.

Dissipative PDEs: Dissipative Wave Equation

We consider the Cauchy problem for the *dissipative wave equation* (or the telegrapher's equation),

$$u_{tt}(x,t) - u_{xx}(x,t) + u_t(x,t) = 0, \qquad t > 0,\ -\infty < x < \infty, \qquad (5.7.23)$$

with smooth initial data $u(x,0)$ and $u_t(x,0)$ specified at $t = 0$. We solve the problem by means of the Fourier transform. An exact solution of the Cauchy problem can be given in terms of modified Bessel functions as shown in Chapter 7. However, we are interested in determining the behavior of the solution $u(x,t)$ as t tends to infinity. The large time behavior of the solution is not readily apparent from the Bessel function representation unless further analyses are carried out. Sirovich's method permits the analysis of the Fourier integral representation of the solution and yields important insights into how the solution behaves for large time.

We define the Fourier transform $U(\lambda,t)$ of $u(x,t)$ as in (5.7.7) and obtain the transformed equation of (5.7.23) as

$$\frac{\partial^2 U(\lambda,t)}{\partial t^2} + \frac{\partial U(\lambda,t)}{\partial t} + \lambda^2 U(\lambda,t) = 0, \qquad t > 0. \qquad (5.7.24)$$

The initial values $U(\lambda,0)$ and $\partial U(\lambda,0)/\partial t$ of $U(\lambda,t)$ are given in terms of the transforms of $u(x,0)$ and $u_t(x,0)$. We are not concerned with the specific form of

the data since we want to characterize the solution in general terms. Solving (5.7.24), we have

$$U(\lambda,t) = F_+(\lambda)e^{(-1+\sqrt{1-4\lambda^2})t/2} + F_-(\lambda)e^{(-1-\sqrt{1-4\lambda^2})t/2}, \tag{5.7.25}$$

and the inverse transform yields the solution as

$$u(x,t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} F_+(\lambda)e^{(-1+\sqrt{1-4\lambda^2})t/2-i\lambda x}\,d\lambda$$

$$+\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} F_-(\lambda)e^{(-1-\sqrt{1-4\lambda^2})t/2-i\lambda x}\,d\lambda, \tag{5.7.26}$$

where $F_+(\lambda)$ and $F_-(\lambda)$ are specified in terms of the initial data. We assume that

$$\int_{-\infty}^{\infty} |F_\pm(\lambda)|\,d\lambda < M < \infty. \tag{5.7.27}$$

To discuss the behavior of the solution for large values of t, we first notice that for $\lambda^2 > 1/4$ we have $\sqrt{1-4\lambda^2} = i\sqrt{4\lambda^2-1}$, since $4\lambda^2 - 1 > 0$. Thus

$$e^{(-1\pm\sqrt{1-4\lambda^2})t/2} = e^{-t/2}e^{\pm i\sqrt{4\lambda^2-1}\,t/2},\ \lambda^2 > 1/4. \tag{5.7.28}$$

Consequently, we obtain

$$\left|\frac{1}{\sqrt{2\pi}}\int_{\lambda^2>\frac{1}{4}} F_\pm(\lambda)e^{\pm i\sqrt{4\lambda^2-1}\,t/2-i\lambda x}e^{-t/2}\,d\lambda\right|$$

$$\le \frac{e^{-t/2}}{\sqrt{2\pi}}\int_{-\infty}^{\infty}|F_\pm(\lambda)|\,d\lambda \le \frac{1}{\sqrt{2\pi}}Me^{-t/2} \tag{5.7.29}$$

in view of (5.7.27). Therefore, the contributions to both integrals in (5.7.26) for values of λ such that $\lambda^2 > 1/4$, are exponentially small for large t. For $\lambda^2 \le 1/4$ we have

$$e^{(-1-\sqrt{1-4\lambda^2})t/2} \le e^{-t/2}, \qquad \lambda^2 \le 1/4, \tag{5.7.30}$$

$$e^{(-1+\sqrt{1-4\lambda^2})t/2} \le e^0 = 1, \qquad \lambda^2 \le 1/4, \tag{5.7.31}$$

as is easily seen. Thus the exponential in (5.7.30) is maximal at $\lambda = \pm\frac{1}{2}$, whereas that in (5.7.31) has its maximum at $\lambda = 0$. Combining the result (5.7.30) with (5.7.29), we obtain an estimate for the $F_-(\lambda)$ integral in (5.7.26) over the full interval of integration. We have

$$\left|\int_{-\infty}^{\infty} F_-(\lambda)e^{(-1-\sqrt{1-4\lambda^2})t/2-i\lambda x}\,d\lambda\right| \le e^{-t/2}\int_{-\infty}^{\infty}|F_-(\lambda)|\,d\lambda. \tag{5.7.32}$$

Therefore, the entire $F_-(\lambda)$ integral is bounded by $(M/\sqrt{2\pi})e^{-t/2}$ and we have shown it to decay exponentially, uniformly in x as $t \to \infty$. The $F_+(\lambda)$ integral does not exhibit uniform decay as $t \to \infty$, and we expect that the major contribution to the solution $u(x,t)$ as $t \to \infty$ must come from this integral. Furthermore, since the exponential (5.7.31) that occurs in this integral decays for all $\lambda \neq 0$, we expect that the major contribution to this integral for large t must come from the neighborhood of $\lambda = 0$.

We apply Sirovich's method to the $F_+(\lambda)$ integral in (5.7.26). To do so we identify $F_+(\lambda)/\sqrt{2\pi}$ with $F(\lambda)$ and the term $(1 - \sqrt{1 - 4\lambda^2})/2$ with $g(\lambda)$ in (5.7.18). We assume that $F_+(\lambda)$ satisfies the first two conditions in (5.7.19) so that, in addition to being absolutely integrable, $F_+(\lambda)$ must be uniformly bounded. Further, we have

$$\mathrm{Re}[g(\lambda)] = \mathrm{Re}\left[(1 - \sqrt{1 - 4\lambda^2})/2\right] \geq 0, \qquad -\infty < \lambda < \infty, \tag{5.7.33}$$

$$g(\lambda) = (1 - \sqrt{1 - 4\lambda^2})/2 = 0, \qquad \text{only at} \qquad \lambda = 0, \tag{5.7.34}$$

$$g(\lambda) = (1 - \sqrt{1 - 4\lambda^2})/2 = \lambda^2 + O(|\lambda|^3), \qquad \lambda \to 0, \tag{5.7.35}$$

on using the binomial expansion. Also, $g(\lambda)$ is continuous for all λ. Thus all the conditions on $g(\lambda)$ are met, and we have $\alpha = 0$ and $\beta = 1$ in the fifth condition in (5.7.19). Then (5.7.22) yields

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F_+(\lambda) e^{(-1+\sqrt{1-4\lambda^2})t/2 - i\lambda x}\, d\lambda \approx \frac{F_+(0)}{\sqrt{2t}} \exp\left(-\frac{x^2}{4t}\right), \tag{5.7.36}$$

which is valid as $t \to \infty$.

It may be remarked that to apply Sirovich's method to the integral in (5.7.18), we simply expand $g(\lambda)$ in a Taylor series around $\lambda = 0$, retaining terms up to quadratic order only. This yields the exponent in (5.7.20). The conditions given on g and F ensure that over the full range of integration the maximum contribution for large t comes from the neighborhood of $\lambda = 0$.

Combining the results (5.7.32) and (5.7.36), we obtain

$$u(x,t) \approx \frac{F_+(0)}{\sqrt{2t}} \exp\left(-\frac{x^2}{4t}\right), \qquad t \to \infty. \tag{5.7.37}$$

Recalling the form of the fundamental solution of the heat equation [see (5.2.39) with $\xi = 0$ and $c = 1$], that is, $G(x,t) = \exp(-x^2/4t)/\sqrt{4\pi t}$, we find that

$$u(x,t) \approx \sqrt{2\pi}\, F_+(0) G(x,t), \qquad t \to \infty. \tag{5.7.38}$$

Since the exponential in (5.7.37) essentially differs from zero only in the parabolic region $x^2/4t = O(1)$, the large time behavior of the solution $u(x,t)$ is as shown in Figure 5.3.

The solution (5.7.38) has the form of a constant multiple of the fundamental solution of the heat or diffusion equation with a source point at the origin. (Note that

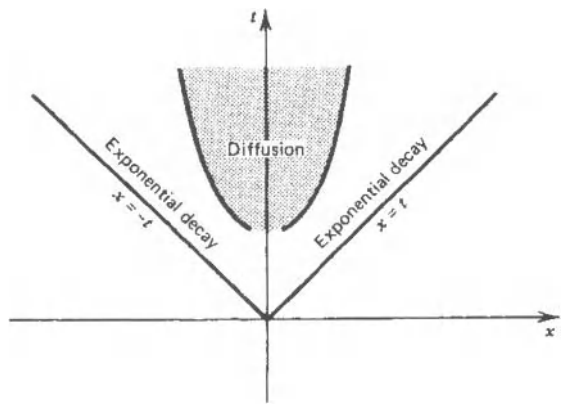

Figure 5.3 The large time behavior.

at large times the solution appears to represent the effect of a source at the origin just as it was the case for the large time solution of the Klein-Gordon equation.) The *diffusion effect* (5.7.38) occurs in the wake of the wave fronts or characteristics $x = \pm t$ after a long time, and yields the major contribution to the solution $u(x,t)$ since the *dissipative effect* has damped out everything else. Although (5.7.38) also dies out as $t \to \infty$, it decays only algebraically like $t^{-1/2}$ for $x^2 \leq 4t$ rather than exponentially as is the case for all other values of x.

For the more general *dissipative wave equation* or telegrapher's equation,

$$u_{tt}(x,t) - c^2 u_{xx}(x,t) + u_t(x,t) - \alpha u_x(x,t) = 0,\ t > 0,\ -\infty < x < \infty, \quad (5.7.39)$$

where α and c are positive constants and appropriate initial data $u(x,0)$ and $u_t(x,0)$ are assigned, the Fourier transform method readily yields

$$u(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F_+(\lambda) \exp\left[\left(-\frac{1}{2} + \frac{1}{2}\sqrt{1 - 4c^2\lambda^2 - 4i\lambda\alpha}\right) t - i\lambda x\right] d\lambda$$

$$+ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F_-(\lambda) \exp\left[\left(-\frac{1}{2} - \frac{1}{2}\sqrt{1 - 4c^2\lambda^2 - 4i\lambda\alpha}\right) t - i\lambda x\right] d\lambda. \quad (5.7.40)$$

with $F_\pm(\lambda)$ specified in terms of the initial data. The coefficient of t in the exponential of the first integral has the following (binomial) expansion around $\lambda = 0$:

$$\left(-1 + \sqrt{1 - 4c^2\lambda^2 - 4i\lambda\alpha}\right)/2 = -i\alpha\lambda + (\alpha^2 - c^2)\lambda^2 + O(\lambda^3). \quad (5.7.41)$$

Unless $\alpha^2 < c^2$, the coefficient of λ^2 is positive, and the exponential grows as t increases, if λ is small. Thus for $c^2 < \alpha^2$ the Cauchy problem for (5.7.39) is unstable since there are normal modes for small λ that grow unboundedly as t increases.

We assume, therefore, that $\alpha^2 < c^2$ in (5.7.39). It can then be shown that the second integral in (5.7.40) decays exponentially as $t \to \infty$ and that Sirovich's method can

be applied to the first integral. We obtain

$$u(x,t) \approx \frac{F_+(0)}{\sqrt{2(c^2-\alpha^2)t}} \exp\left[-\frac{(x+\alpha t)^2}{4(c^2-\alpha^2)t}\right], \qquad t \to \infty. \tag{5.7.42}$$

As was shown in Example 5.2, the expression (5.7.42) is—apart from a multiplicative constant—the fundamental solution of the diffusion equation (5.2.40) if we set $c = -\alpha$ and $D/2 = c^2 - \alpha^2$ in that equation. In contrast to the result obtained for (5.7.23), the diffusion effect is now concentrated along the line $x = -\alpha t$ rather than along the t-axis. To observe the diffusion as $t \to \infty$, we must let $x \to -\infty$ such that $x + \alpha t$ remains small. Since $\alpha < c$, the line of diffusion $x + \alpha t = 0$ lies within the characteristic lines $x = \pm ct$. As the *diffusion effect* may be thought to be traveling along the line $x = -\alpha t$ with velocity $-\alpha$, we may consider (5.7.39) to possess two speeds of wave propagation. For small times t, disturbances travel at the characteristic speed c to the right and/or to the left. At large times t, the *dissipative effects* have damped out most of the initial disturbances, and what remains in the wake of the wave fronts is a *diffusion effect* characterized by (5.7.42) that propagates with speed α to the left.

The relationship between dissipative wave propagation and diffusion has been discussed in the models presented in Chapter 1. The foregoing asymptotic results yield a further demonstration of the close relationship between both effects. We have further occasion to touch upon this question later in the book.

To conclude our discussion of asymptotic methods, we remark that even when exact solutions of problems are available, asymptotic approximations yield useful and often significant insights into the behavior of solutions in different regions. In addition, by suggesting general forms for solutions in different regions, it is often possible to construct approximate solutions of equations (say, when they have variable coefficients) for which no exact solution is available, by assuming that these solutions have the same general asymptotic form. This idea is exploited in Chapter 10 when we discuss direct asymptotic methods for partial differential equations.

Exercises 5.7

5.7.1. The Bessel function of integral order $J_n(x)$ has the integral representation $J_n(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \exp(ix\sin(\theta) - in\theta)\, d\theta$. Apply the method of stationary phase to obtain the asymptotic formula $J_n(x) \approx \sqrt{2/\pi x}\cos(x - \pi n/2 - \pi/4)$, valid as $x \to \infty$. *Hint*: There are two stationary points.

5.7.2. The Airy function $Ai(x)$ has the integral representation $Ai(x) = (1/\pi)\int_0^{\infty} \cos(\tau^3/3 + x\tau)\, d\tau$. Let $x < 0$ and $\tau = \sqrt{-x}\, t$ and use the method of stationary phase to show that the Airy function has the asymptotic expression $Ai(x) \approx (1/\sqrt{\pi})|x|^{-1/4}\sin\left[(2/3)|x|^{3/2} + \pi/4\right]$, as $x \to -\infty$. *Hint*: There are two stationary points.

5.7.3. Express the functions $F_+(\lambda)$ and $F_-(\lambda)$ in (5.7.10) in terms of the transforms of the initial data $u(x,0) = f(x)$ and $u_t(x,0) = g(x)$. Show that $F_+(-\lambda) = \overline{F_-(\lambda)}$

(the complex conjugate) and that the solution $u(x,t)$ in (5.7.10) is real. Obtain the asymptotic form of $u(x,t)$ by considering the contributions from the $F_+(\lambda)$ and $F_-(\lambda)$ integrals.

5.7.4. Show that $u(x,t) = J_0\left[(c/\gamma)\sqrt{\gamma^2 t^2 - x^2}\right]$ is a solution of the Klein-Gordon equation (5.7.6). Let $x = \alpha t$ in the Bessel function with $|\alpha| < \gamma$, and expand J_0 asymptotically for large x and t. Show that the result agrees with that obtained in Exercise 5.7.3 for an appropriate choice of $F_+(\lambda)$ and $F_-(\lambda)$.

5.7.5. Use the Fourier transform to obtain the solution of the Cauchy problem for the linearized Korteweg-deVries equation (see Example 9.7) $u_t(x,t) + a^2 u_x(x,t) + b^2 u_{xxx}(x,t) = 0$, $-\infty < x < \infty$, $t > 0$, with the initial condition $u(x,0) = f(x)$, $-\infty < x < \infty$. Apply the method of stationary phase to discuss the solution $u(x,t)$ for large x and t.

5.7.6. Use the Fourier transform to solve the initial value problem for the dissipative wave equation $u_{tt}(x,t) - u_{xx}(x,t) - a^2 u_{xxt}(x,t) = 0$, $-\infty < x < \infty$, $t > 0$, with the initial data $u(x,0) = f(x)$, $u_t(x,0) = a^2 f''(x)$, $-\infty < x < \infty$, $t > 0$. Apply Sirovich's method to obtain an asymptotic expression for the solution as t gets large.

5.7.7. Apply the Fourier transform to solve the boundary value problem for the elliptic equation $u_{xx}(x,t) + u_{tt}(x,t) - u_t(x,t) = 0$, $-\infty < x < \infty$, $t > 0$, assuming that $u(x,t)$ and $f(x,t)$ are suitably behaved at infinity. Apply Sirovich's method to discuss the asymptotic behavior of $u(x,t)$ for large t.

5.7.8. Verify that the solution (5.7.40) of the initial value problem for (5.7.39) has the behavior given in (5.7.42) for large t.

5.8 MAPLE METHODS

Maple's procedures for carrying out integral transforms and inverting them are contained in the package *inttrans*, which is accessed by invoking *with(inttrans)*. The output is [*addtable*, *fourier*, *fouriercos*, *fouriersin*, *hankel*, *hilbert*, *invfourier*, *invhilbert*, *invlaplace*, *invmellin*, *laplace*, *mellin*, *savetable*]. Clearly, the transforms *fourier*, *fouriercos*, *fouriersin*, *hankel*, and *laplace* are the ones we have considered in this chapter. The Fourier and Laplace transforms have the inversion formulas *invfourier* and *invlaplace*, respectively. The Fourier sine, Fourier cosine and Hankel transforms are self-inverting. We will not consider Mellin and Hilbert transforms and the addtable and savetable procedures. Maple's definitions of the Fourier and Hankel transforms differ from those given in the text. We have constructed procedures that convert Maple's version of these transforms and their inverses to ours.

Fourier Transforms

Maple's *Fourier transform* $F_M(\lambda)$ and $F_T(\lambda)$ as defined in (5.2.6) are

$$F_M(\lambda) = \int_{-\infty}^{\infty} f(x)\, e^{-i\lambda x}\, dx, \qquad F_T(\lambda) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)\, e^{i\lambda x}\, dx. \quad (5.8.1)$$

The *inverse Fourier transform* in Maple and in (5.2.7) are given as

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F_M(\lambda) e^{i\lambda x}\, d\lambda, \qquad f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F_T(\lambda) e^{-i\lambda x}\, d\lambda. \quad (5.8.2)$$

The Maple procedures *fourier*$(f(x), x, \lambda)$ and *invfourier*$(F_M(\lambda), \lambda, x)$ determine the Fourier transform $F_M(\lambda)$ of $f(x)$ and its inverse. Rather than use Maple's integration facility to evaluate integrals that arise in the Fourier transform and its inverse as defined in the text, we have constructed two procedures that use Maple's transforms and convert the outputs into forms that result from the text Fourier transforms. They are given as *Fourier*$(f(x), x, \lambda)$, which yields $F_T(\lambda)$, and *InvFourier*$(F_T(\lambda), \lambda, x)$, whose output is $f(x)$. This means that Maple's built-in tables of transforms are invoked even when the text's definition of the Fourier transform and its inverse are used.

In the following table we present some Fourier transforms and their inverses as given by Maple, with $F(\lambda) = \textit{fourier}(f(x), x, \lambda)$ [we drop the subscript notation $F_M(\lambda)$] and $f(x) = \textit{invfourier}(F(\lambda), \lambda, x)$.

$f(x)$	$F(\lambda)$
$\frac{1}{\sqrt{4\pi c^2 t}} \exp\left(-\frac{x^2}{4c^2 t}\right)$	$\exp(-\lambda^2 c^2 t)$
$\frac{1}{\sqrt{8\pi t}} \left\{\cos\left(\frac{x^2}{4t}\right) + \sin\left(\frac{x^2}{4t}\right)\right\}$	$\cos(\lambda^2 t)$
1	$2\pi\delta(\lambda)$
$H(x)$	$\pi\delta(\lambda) - i/\lambda$
$\frac{1}{2}(\delta(x+1) + \delta(x-1))$	$\cos(\lambda)$
$e^{-k\lvert x\rvert}$, $k > 0$	$2k/(k^2 + \lambda^2)$
$xf(x)$	$i\frac{\partial F(\lambda)}{\partial \lambda}$

Here $\delta(x)$ is the Dirac delta function and $H(x)$ is the Heaviside function. The inverse Fourier transform of $F(\lambda) = \exp(-\lambda^2 c^2 t)$ as defined in the text is

$$InvFourier(\exp(-\lambda^2 c^2 t), \lambda, x) = \frac{1}{\sqrt{2c^2 t}} \exp\left(-\frac{x^2}{4c^2 t}\right). \quad (5.8.3)$$

We will generally use Maple's version of the Fourier transform and its inverse in our discussion, for the purpose of simplicity.

The Fourier transform applied to the heat equation as $fourier(u_{tt}(x,t) - c^2 u_{xx}(x,t) = 0, x, \lambda)$ yields the transformed equation.

$$\frac{\partial^2}{\partial t^2} fourier(u(x,t), x, \lambda) + c^2 \lambda^2 fourier(u(x,t), x, \lambda) = 0. \tag{5.8.4}$$

This is an ODE in the Fourier transform $fourier(u(x,t), x, \lambda)$. To make use of Maple's ODE solver *dsolve*, we substitute $fourier(u(x,t), x, \lambda) = U(t)$ into (5.8.4). On assuming that the initial condition for the heat equation is $u(x,0) = f(x)$, we find that $U(0) = fourier(f(x), x, \lambda)$. The solution is

$$fourier(u(x,t), x, \lambda) = fourier(f(x), x, \lambda) e^{-c^2 \lambda^2 t}. \tag{5.8.5}$$

On inverting the transform, Maple makes use of the convolution theorem for Fourier transforms to obtain the result (5.2.38), after some simplification.

It is not possible, in general, to evaluate the Fourier integrals and obtain closed-form expressions for Fourier transforms and their inverses. The methods of Section 5.7 can be used to generate approximate results for ranges of values of the transform parameters. If the Fourier integrals contain no parameters, say they are evaluated at specific values of x or t, Maple's numerical integration facility can be used to obtain numerical approximations. In the following example we consider the numerical evaluation of an inverse transform. Even though the Fourier integral can be evaluated explicitly, we employ an approximate method to evaluate it and examine the consequences.

Example 5.12. The Fourier Transform of a Discontinuous Function and Gibbs Phenomenon. We consider the discontinuous function

$$f(x) = \begin{cases} 1, & |x| < 1, \\ 0, & |x| > 1. \end{cases} \tag{5.8.6}$$

Maple finds its Fourier transform $F(\lambda)$ to be $F(\lambda) = 2\sin(\lambda)/\lambda$. Inverting the transform reproduces the function (5.8.5) at points of continuity and has the value $\frac{1}{2}$ at the points $x = \pm 1$ of discontinuity. This is expected because the transform converges to average value of the limits of $f(x)$ at points of discontinuity.

To determine the behavior of the inverse transform near the points of discontinuity if it is evaluated approximately rather than exactly, we express it in terms of its real and imaginary parts and easily conclude that it is given as

$$f(x) = \frac{2}{\pi} \int_0^\infty \frac{\sin(\lambda)\cos(x\lambda)}{\lambda}\, d\lambda. \tag{5.8.7}$$

We obtain an approximation to the improper integral by replacing the upper limit ∞ by a finite limit s. Then Maple expresses the resulting integral as

$$f(x) \approx \frac{2}{\pi} \int_0^s \frac{\sin(\lambda)\cos(x\lambda)}{\pi}\, d\lambda = \frac{1}{\pi} \left[Si\left((x+1)\, s\right) - Si\left((x-1)\, s\right) \right]. \tag{5.8.8}$$

The Sine integral $Si(s)$ is defined as $Si(s) = \int_0^s \sin(t)/t\ dt$. By choosing ever-increasing values of s in (5.8.8), we expect to get improved approximations to the value of $f(x)$ at some fixed value of x. However, near the points of the jump discontinuity, the *Gibbs phenomenon*, which was observed in the theory of trigonometric Fourier series, comes into play. To see this we put $s = 100,000$ in (5.8.8) and evaluate the expression at several values of x. This gives for the approximate values $f_a(x)$ of $f(x)$

$$f_a(0.5) = .9999999999, \quad f_a(1) = .4999984126, \quad f_a(1.5) = -0.0000000003,$$

$$f_a(0.9999) = 1.027869704, \quad f_a(1.0001) = -0.02786703995. \tag{5.8.9}$$

We find that at the points of continuity $x = 0.5$ and $x = 1.5$ of $f(x)$ the results are extremely close to the exact values 1 and 0, respectively. At the point $x = 1$ of jump discontinuity, the approximate value is extremely close to 0.5, the average jump values. However, near the point $x = 1$ the approximate results overshoot and undershoot the exact limiting values 1 and 0, respectively, by the amount 0.0278. This is the *Gibbs phenomenon.*

Fourier Sine and Cosine Transforms

Maple's definitions of the Fourier sine and cosine transforms and their inverses agree with those given in Section 5.3. Both transforms are self-inverting in the sense that the variable of integration must be changed from x to λ to switch from the sine and cosine transforms $F_s(\lambda)$ and $F_c(\lambda)$, respectively, of $f(x)$, to their inverses. The Maple procedures *fouriersin*$(f(x), x, \lambda)$ and *fouriercos*$(f(x), x, \lambda)$ find the sine and cosine transforms of $f(x)$. Thus, the sine transforms and their inverses are given as

$$fouriersin(f(x), x, \lambda) = \sqrt{\frac{2}{\pi}} \int_0^\infty \sin(\lambda x) f(x)\ d\lambda,$$

$$fouriersin(fouriersin(f(x), x, \lambda), \lambda, x) = f(x), \tag{5.8.10}$$

and the cosine transforms and their inverses are

$$fouriercos(f(x), x, \lambda) = \sqrt{\frac{2}{\pi}} \int_0^\infty \cos(\lambda x) f(x)\ d\lambda,$$

$$fouriercos(fouriercos(f(x), x, \lambda), \lambda, x) = f(x). \tag{5.8.11}$$

It is important to note that the order of (x, λ) and (λ, x) is reversed on carrying out the direct and inverse Fourier sine and cosine transforms as shown in the above.

In the following table we present some Fourier sine transforms and their inverses as given by Maple: $F_s(\lambda) = fouriersin(f(x), x, \lambda)$, $f(x) = fouriersin(F_s(\lambda), \lambda, x)$.

$f(x)$	$F_s(\lambda)$
$\begin{cases} 1, & 0 < x < a \\ 0, & x > a \end{cases}$	$\sqrt{\frac{2}{\pi}}(1 - \cos(\lambda a))/\lambda$
1	$\sqrt{\frac{2}{\pi\lambda}}$
$e^{-kx},\ k > 0$	$\sqrt{\frac{2}{\pi}}\lambda/(k^2 + \lambda^2)$
$1/x$	$\sqrt{\frac{\pi}{2}}$
$xf(x)$	$-\partial F_c(\lambda)/\partial\lambda$

Next, we present a table of some Fourier cosine transforms and their inverses as given by Maple: $F_c(\lambda) = fouriercos(f(x), x, \lambda)$, $f(x) = fouriercos(F_c(\lambda), \lambda, x)$.

$f(x)$	$F_c(\lambda)$
$\begin{cases} 1, & 0 < x < a \\ 0, & x > a \end{cases}$	$\sqrt{\frac{2}{\pi}}(\sin(\lambda a))/\lambda$
1	$\sqrt{\frac{\pi}{2}}\,\delta(\lambda)$
$e^{-kx},\ k > 0$	$\sqrt{\frac{2}{\pi}}k/(k^2 + \lambda^2)$
e^{-kx^2}	$\sqrt{\frac{1}{2k}}\exp(-\lambda^2/4k)$
$xf(x)$	$\partial F_s(\lambda)/\partial\lambda$

Again, $\delta(x)$ is the Dirac delta function.

To demonstrate how the Fourier sine and cosine transforms are to be used when dealing with boundary value problems, we construct the solution of two boundary value problems for an ODE in the following example.

Example 5.13. A Boundary Value Problem for an Ordinary Differential Equation. We consider the ODE

$$y''(x) - k^2 y(x) = e^{-x}, \qquad x > 0, \qquad k \neq 1. \tag{5.8.12}$$

With the boundary condition at $x = 0$ given as $y(0) = a$, we apply the Fourier sine transform to (5.8.12) and obtain

$$fouriersin(y(x), x, \lambda) = \sqrt{\frac{2}{\pi}}\,\frac{\lambda\,(a + a\lambda^2 - 1)}{\lambda^2 + \lambda^4 + k^2 + k^2\lambda^2}. \tag{5.8.13}$$

Inverting the transform yields the solution

$$y(x) = \frac{(1 + ak^2 - a)e^{-kx} - e^{-x}}{k^2 - 1}. \tag{5.8.14}$$

We observe that the solution satisfies the given ODE (5.8.12), the boundary condition at $x = 0$, and vanishes at infinity.

For the case of a derivative boundary condition at $x = 0$ given as $y'(0) = b$, we apply the Fourier cosine transform to (5.8.12) and obtain

$$fouriercos(y(x), x, \lambda) = -\sqrt{\frac{2}{\pi}}\,\frac{1 + b + b\lambda^2}{\lambda^2 + \lambda^4 + k^2 + k^2\lambda^2}. \tag{5.8.15}$$

Inverting the transform yields the solution

$$y(x) = \frac{(1 + b - bk^2)e^{-k\,x} - ke^{-x}}{(-1 + k^2)k}. \tag{5.8.16}$$

The solution satisfies (5.8.12), the boundary condition at $x = 0$, and vanishes at infinity.

Higher-Dimensional Fourier Transforms

Higher-dimensional Fourier transforms must be obtained as iterated one-dimensional Fourier transforms. The same is true for the inverse transforms. (There is no Maple procedure that finds higher-dimensional Fourier transforms directly.) We again work only with Maple's definition of the one-dimensional Fourier transform, but use the transform parameters of Section 5.4. In two dimensions, Maple gives the Fourier transform of $f(x, y)$ as

$$fourier(fourier(f(x, y), x, \lambda_1), y, \lambda_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)e^{-i(x\lambda_1 + y\lambda_2)}\,dx\,dy. \tag{5.8.17}$$

We denote the transform by $F(\lambda_1, \lambda_2)$. Then the inverse transform is given as

$$invfourier(invfourier(F(\lambda_1, \lambda_2), \lambda_1, x), \lambda_2, y)$$
$$= \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F\,(\lambda_1, \lambda_2)\,e^{i(x\lambda_1 + y\lambda_2)}\,d\lambda_1\,d\lambda_2. \tag{5.8.18}$$

The extension of the formulas to more dimensions is straightforward.

As an example, we find the transform of $\exp[-(x^2 + y^2)/(4c^2t)]$. We iterate the Fourier transform as above and find that

$$fourier(fourier(e^{-(x^2+y^2)/4c^2t}, x, \lambda_1), y, \lambda_2) = 4\,c^2t\,\pi\,e^{-(\lambda_1^2c^2t - \lambda_2^2c^2t)}. \tag{5.8.19}$$

The product of Dirac delta functions $4\pi^2\delta(\lambda_1)\delta(\lambda_2)$ has the inverse Fourier transform

$$invfourier(invfourier(4\pi^2\delta(\lambda_1)\delta(\lambda_2), \lambda_1, x), \lambda_2, y) = 1. \tag{5.8.20}$$

The iterated Fourier transforms can be applied to PDEs in two or more space dimensions as in Section 5.4, but we do not consider this here.

Hankel Transforms

Maple's definition of the *Hankel transform* and its inverse differs from that given in Section 5.5 in (5.5.3) and (5.5.2). We denote the Maple Hankel transform by $F_{HM}(\lambda)$ and the definition (5.5.3) by $F_{HT}(\lambda)$. Then we have, for the Hankel transform of $f(r)$ of order n,

$$F_{HM}(\lambda) = \int_0^\infty \sqrt{\lambda r} J_n(\lambda r) f(r)\, dr, \quad F_{HT}(\lambda) = \int_0^\infty r J_n(\lambda r) f(r)\, dr. \tag{5.8.21}$$

The expressions for the inverse transforms are given as

$$f(r) = \int_0^\infty \sqrt{\lambda r} J_n(\lambda r) F_{HM}(\lambda)\, d\lambda, \; f(r) = \int_0^\infty \lambda J_n(\lambda r) F_{HT}(\lambda)\, d\lambda. \tag{5.8.22}$$

Here $J_n(x)$ represents the Bessel function of order n. [In Maple's notation it is given as $BesselJ(n, x)$.] Maple's definition of the Hankel transform is given in a self-inverting form. The Maple Hankel transform of order n of a function $f(r)$ is determined from the procedure $hankel(f(r), r, \lambda, n)$. If we denote the transform by $F(\lambda)$, the inverse Hankel transform $f(r)$ is given by $hankel(F(\lambda), \lambda, r, n)$.

$HankelTrans(f(r), r, \lambda, n)$ yields the Hankel transform $F_{HT}(\lambda)$ of order n of $f(r)$ in the form given in Section 5.5. To determine the inverse Hankel transform $f(r)$ of $F_{HT}(\lambda)$ we use $HankelTrans(F_{HT}(\lambda), \lambda, r, n)$. On applying $HankelTrans$ $(L(y(r)), r, \lambda, n)$ to the ordinary differential operator

$$L(y(r)) = y''(r) + \frac{1}{r} y'(r) - \frac{n^2}{r^2} y(r), \tag{5.8.23}$$

we obtain as the Hankel transform

$$-\lambda^2 F(\lambda) - \lim_{r \to 0^+} \left[r y'(r) J_n(\lambda r) - \lambda^2 r y(r) \left(-J_{n+1}(\lambda r) + \frac{n J_n(\lambda r)}{\lambda r} \right) \right], \tag{5.8.24}$$

where $F(\lambda)$ is the Hankel transform of order n of $f(r)$. This result agrees with the expression found in (5.5.4), if we note the relationship $J_n'(\lambda r) = -\lambda\, J_{n+1}(\lambda r) + nJ_n(\lambda r)/r$ and assume that the limit of the bracketed expression as $r \to \infty$ is zero. If we apply Maple's *hankel* procedure to (5.8.23), we obtain $-\lambda^2 F(\lambda)$ instead of (5.8.24) because Maple assumes that the limit in (5.8.24) is zero in its procedure. As shown in Example 5.9, this is not always the case. We do not apply the procedure to a specific problem but instead, present a table of some Hankel transforms with $F_{HT}(\lambda)$ as the Hankel transform of $f(r)$. (We employ the $HankelTrans$ procedure.)

$f(r)$	$F(\lambda)$
r^k	$\frac{2^{k+1}\Gamma(1/2(n+k)+1)}{\lambda^{k+2}\Gamma(1/2(n-k))}$
1	$\frac{n}{\lambda^2}$
$H(k-x),\ k>0, n=0$	$\frac{k}{\lambda} J_1(k\lambda)$
$e^{-kr^2}, n=0,$	$\frac{1}{2k}\exp(-\lambda^2/4k)$

Again, $H(x)$ is the Heaviside function.

Laplace Transforms

The built-in Maple procedure $laplace(f(t), t, \lambda)$ determines the *Laplace transform* $F(\lambda)$ of $f(t)$ as

$$F(\lambda) = \int_0^\infty f(t)e^{-\lambda t}\, dt. \tag{5.8.25}$$

The definition agrees with that given in (5.6.2). The *inverse Laplace transform* $f(t)$ is given by $invlaplace(F(\lambda), \lambda, t)$ A number of useful Laplace transforms of specific functions are presented in the exercises for Section 5.6, so we do not present a table of transforms here. Instead, we solve an initial and boundary value problem for the heat equation in a finite interval and examine the behavior of the solution for small and large values of t.

Example 5.14. The Heat Equation in a Finite Interval. We consider the initial and boundary value problem for the heat equation

$$u_t(x,t) = c^2 u_{xx}(x,t),\ 0<x<l,\ t>0, \quad u(x,0)=0,\ u(0,t)=1,\ u(l,t)=0. \tag{5.8.26}$$

We apply the Laplace transform $laplace(u_t(x,t) = c^2 u_{xx}(x,t), t, \lambda)$ to the heat equation and denote the transform by $U(x,\lambda)$. The boundary conditions must also be transformed. We obtain $laplace(1,t,\lambda) = 1/\lambda$ and $laplace(0,t,\lambda) = 0$. The transformed problem is

$$\lambda U(x,\lambda) - u(x,0) = c^2 \frac{\partial^2 U(x,\lambda)}{\partial x^2},\ 0<x<l, \quad U(0,\lambda) = \frac{1}{\lambda},\ U(l,\lambda) = 0. \tag{5.8.27}$$

Using Maple's *dsolve* procedure to solve (5.8.27) yields

$$U(x,\lambda) = \frac{\exp\left(\frac{\sqrt{\lambda}\,x}{c}\right) - \exp\left(\frac{\sqrt{\lambda}(2l-x)}{c}\right)}{\lambda\left(1-\exp\left(\frac{\sqrt{\lambda}\,2l}{c}\right)\right)}. \tag{5.8.28}$$

First, we approximate the transform for large values of λ and obtain a result that gives a useful approximation to the solution for small t according to *Abelian theory*. Let

$$r = \exp\left(\frac{-\sqrt{\lambda}\,l}{c}\right), \qquad U(x,\lambda) = \frac{\exp\left(-\frac{\sqrt{\lambda}\,x}{c}\right) - r^2 \exp\left(\frac{\sqrt{\lambda}\,x}{c}\right)}{\lambda(1-r^2)}. \quad (5.8.29)$$

Then $r < 1$ for sufficiently large λ. We expand $U(x,\lambda)$ in a Taylor series around $r = 0$ and obtain

$$U(x,\lambda) = \frac{\exp\left(-\frac{\sqrt{\lambda}\,x}{c}\right) + \sum_{n=1}^{\infty}\left[\exp\left(-\frac{\sqrt{\lambda}\,x+2n\sqrt{\lambda}\,l}{c}\right) - \exp\left(\frac{\sqrt{\lambda}\,x-2n\sqrt{\lambda}\,l}{c}\right)\right]}{\lambda}. \quad (5.8.30)$$

On using *invlaplace* to invert the transform we obtain the solution

$$u(x,t) = \operatorname{erfc}\left(\frac{x}{2c\sqrt{t}}\right) + \sum_{n=1}^{\infty}\left[\operatorname{erfc}\left(\frac{x}{2c\sqrt{t}} + \frac{ln}{c\sqrt{t}}\right) - \operatorname{erfc}\left(-\frac{x}{2c\sqrt{t}} + \frac{ln}{c\sqrt{t}}\right)\right], \quad (5.8.31)$$

where erfc(z) is the complementary error function. For large $n/\sqrt{t}$ the bracketed terms in the sum are small and (5.8.31) is well approximated by the leading terms.

To obtain an approximation of the solution $u(x,t)$ for large t, we obtain a *Tauberian result* of the form (5.6.22). We express the Laplace transform (5.8.28) in the form

$$U(x,\lambda) = \frac{\sinh\left(\frac{\sqrt{\lambda}(l-x)}{c}\right)}{\lambda \sinh\left(\frac{\sqrt{\lambda}\,l}{c}\right)}. \quad (5.8.32)$$

As in Section 5.6, we find that the zeros of $\sinh(\sqrt{\lambda}\,l/c)$ given as $\lambda_n = -(\pi nc/l)^2$ $(n = 0, 1, 2, \ldots)$ are the simple poles of $U(x,\lambda)$. The residues at these poles are easily determined by setting $G(\lambda) = \sinh(\sqrt{\lambda}(l-x)/c)/\sqrt{\lambda}$ and $H(\lambda) = \sinh(\sqrt{\lambda}\,l/c)\sqrt{\lambda}$, so that

$$\beta_0 = \frac{l-x}{l}, \quad \beta_n = -\frac{2}{\pi n}\sin\left(\frac{\pi n x}{l}\right), \; n \geq 1. \quad (5.8.33)$$

The Tauberian result is thus found to be

$$u(x,t) = \frac{l-x}{l} - \sum_{n=1}^{\infty}\frac{2}{\pi n}\sin\left(\frac{\pi n x}{l}\right)\exp\left(-\frac{\pi^2 n^2 c^2 t}{l^2}\right). \quad (5.8.34)$$

The result agrees with that obtained via the modified eigenfunction expansion of Section 4.6. The steady-state term $(l-x)/l$ is separated out, and the expansion is useful for large t.

Asymptotic Approximation Methods for Fourier Integrals

StatPhase finds the leading term(s) of an asymptotic approximation of Fourier type integrals of oscillatory type with a large parameter k of the form

$$I(k) = \int_{-\infty}^{\infty} e^{ik\phi(t)} f(t)\, dt, \tag{5.8.35}$$

where $f(t)$ is a real- or complex-valued function, referred to as the *amplitude term.* The real-valued function $\phi(t)$ is called the *phase term.* (The case of finite endpoints in the integral can also be considered.) The procedure has the form $StatPhase(\phi(t), f(t), t = a..b, k)$. It attempts to find the stationary points of the phase [i.e., the roots of $\phi'(t) = 0$]. If all the stationary points are not found within the procedure or if they do not all lie within the interval of integration, the relevant points must be supplied to the procedure by adding a fifth argument which lists the stationary points.

Example 5.15. The Asymptotic Approximation of the Bessel Function. For integral n, the *Bessel function* $J_n(x)$ can be given as

$$J_n(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-in\theta} e^{ix\sin(\theta)}\, d\theta. \tag{5.8.36}$$

We use $StatPhase$ to determine the asymptotic approximation of $J_n(x)$ as $x \to \infty$. Comparing (5.8.35) with (5.8.36), we see that $t \leftrightarrow \theta$, $k \leftrightarrow x$, $\phi(t) \leftrightarrow \sin(\theta)$, and $f(t) \leftrightarrow e^{-in\theta}/2\pi$, and we use $StatPhase(\sin(\theta), e^{-in\theta}/2\pi, \theta = -\pi..\pi, x)$. The output lists all the stationary points of the phase $\sin(\theta)$ as $\theta_k = \pi\,(k + 1/2)$ $(k = 0, \pm 1, \pm 2, \ldots)$ [i.e., the zeros of $\cos(\theta)$]. Only the points $\mp\pi/2$ lie within the interval of integration $[-\pi, \pi]$. We reapply the procedure with the added fifth argument $[-\pi/2, \pi/2]$. This yields

$$J_n(x) \approx \sqrt{\frac{2}{\pi x}} \cos\left(x - \frac{\pi n}{2} - \frac{\pi}{4}\right), \qquad x \gg 1. \tag{5.8.37}$$

The procedure *SirMeth* obtains asymptotic approximations of Fourier type integrals of dissipative type using *Sirovich's method* of Section 5.7. The integrals have the form (5.7.28)

$$I(x,t) = \int_{-\infty}^{\infty} F(\lambda)\, e^{-g(\lambda)t - i\lambda x}\, d\lambda. \tag{5.8.38}$$

The procedure $SirMeth(g(\lambda), F(\lambda), [x,t], \lambda)$ checks if the conditions given in (5.7.19) are met. If so, it gives an asymptotic approximation of the integral $I(x,t)$ valid for large t. If all the conditions are not satisfied, an error message appears.

As an example, we apply $SirMethod$ to the integral

$$I(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(\lambda) e^{-\left(a^2\lambda^2/2 - \sqrt{a^4\lambda^4 - 4\lambda^2}/2\right)t - i\lambda x}\, d\lambda, \tag{5.8.39}$$

The procedure determines that all the conditions required for the validity of Sirovich's method are satisfied and yields the result

$$I(x,t) \approx \frac{1}{|a|\sqrt{t}} F(0) \exp\left(-\frac{(x-t)^2}{2a^2 t}\right), \qquad t \gg 1. \tag{5.8.40}$$

Discrete Fourier Transform and Fast Fourier Transform

As was indicated in Section 5.2, the Fourier integral formula (5.2.5) from which the Fourier transform and its inverse were derived can be generated by considering a complex form of the Fourier series representation of $f(x)$ over the interval $[-l,\ l]$ and allowing l to tend to infinity (see Exercises 5.2.1 and 5.2.2). For the present discussion, we consider the function $f(x)$ to be given over the interval $[0,\ 2l]$ rather than the interval $[-l,\ l]$. Then the *complex Fourier series representation* of $f(x)$ is given as

$$f(x) = \sum_{k=-\infty}^{\infty} \hat{c}_k \exp\left(\frac{i\pi kx}{l}\right), \tag{5.8.41}$$

and the *Fourier coefficients* $\hat{c}_k$ are given as

$$\hat{c}_k = \frac{1}{2l} \int_0^{2l} f(t) \exp\left(-\frac{i\pi kt}{l}\right) dt. \tag{5.8.42}$$

We have modified the forms of the complex Fourier series and the Fourier coefficients in Exercise 5.2.1, by replacing i by $-i$.

In general, the integrals in (5.8.42) for the Fourier coefficients cannot be evaluated exactly and numerical methods must be used. We use the *trapezoidal rule* to do so. The interval $[0, 2l]$ is divided into N equal parts to yield the points $x_j = 2l(j-1)/N$ with $j = 1, 2, \ldots, N$. The values $f(x_j)$ are given at each of the x_j. [It will be assumed that $f(2l) = f(0)$ because of periodicity.] We use the procedure *AproximateInt* from Maple's *Student Calculus1* package, with the arguments $method = trapezoid$ and $partition = N$, and obtain after some simplification the *approximate Fourier coefficients*

$$\hat{c}_k \approx \frac{1}{N} \sum_{j=0}^{N-1} f(x_j) \exp\left(\frac{-2\,ikj\pi}{N}\right). \tag{5.8.43}$$

We observe that $\hat{c}_{k+N} = \hat{c}_k$, so that only N different $\hat{c}_k$ are determined from (5.8.43). These terms are entered into a truncated form of the Fourier series (5.8.41), given as

$$f(x) \approx f_N(x) = \sum_{k=0}^{N-1} \hat{c}_k \exp\left(\frac{i\pi kx}{l}\right). \tag{5.8.44}$$

As indicated below, this approximation is referred to as a *trigonometric interpolation polynomial.* Rather than discuss this now, we reexamine the foregoing from a slightly different perspective.

In general, the function $f(x)$ may not be defined over the entire interval $[0, 2l]$. Instead, a set of numbers f_j (they may be real or complex) are prescribed at the points x_j, and we associate these numbers with the values of a function $f(x)$ [i.e., $f(x_j) = f_j$]. The set of points x_j is defined as above, and the set of prescribed values at the points will be denoted by $f(x_j)$ whether $f(x)$ is known or not. Discrete Fourier transforms and their inverses are given in terms of the values.

We use the Maple package *DiscreteTransforms*, which contains the two procedures *FourierTransform* and *InverseFourierTransform* to determine discrete Fourier transforms and their inverses. (Therefore, the notation follows that of Maple rather than that of Exercises 5.2.1 and 5.2.2.) Then the discrete representation of $f(x)$, given in terms of complex exponentials, is

$$f(x_j) = \frac{1}{\sqrt{N}} \sum_{k=1}^{N} c_k \exp\left(\frac{2\pi i(k-1)(j-1)}{N}\right), \qquad j = 1, \ldots, N. \quad (5.8.45)$$

The c_k are the *Fourier coefficients* and are given as

$$c_k = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} f(x_j) \exp\left(\frac{-2\pi i(k-1)(j-1)}{N}\right), \qquad k = 1, \ldots, N. \quad (5.8.46)$$

The set of values $\{c_k\}$ represent the *discrete Fourier transform* of the set of input values $\{f(x_j)\}$. The set of values $\{f(x_j)\}$ is the *inverse discrete Fourier transform* of the $\{c_k\}$.

It may be noted that $f(x_1) = f(x_{N+1})$ with $x_1 = 0$ and $x_{N+1} = 2l$, so that $f(0) = f(2l)$. The validity of the representation of the $\{c_k\}$ is based on the following orthogonality and normalization conditions.

$$\sum_{k=1}^{N} e^{\left(\frac{2\pi i(k-1)(n-1)}{N}\right)} e^{\left(\frac{-2\pi i(k-1)(m-1)}{N}\right)} = 0, \quad n \neq m, \quad (5.8.47)$$

$$\sum_{k=1}^{N} e^{\left(\frac{2\pi i(k-1)(n-1)}{N}\right)} e^{\left(\frac{-2\pi i(k-1)(n-1)}{N}\right)} = N. \quad (5.8.48)$$

If $f(x)$ is known, the foregoing can be used to determine a *trigonometric interpolation polynomial* $f_N(x)$ for $f(x)$ in terms of a discrete set of N interpolation values $x_j = 2l(j-1)/N$ given over the interval $[0, 2l]$. It has the form

$$f_N(x) = \frac{1}{\sqrt{N}} \sum_{k=1}^{N} c_k \exp\left(\frac{\pi i(k-1)x}{l}\right), \quad (5.8.49)$$

with the *Fourier coefficients* given as in (5.8.46). The trigonometric polynomial $f_N(x)$ clearly assumes the values $f(x_j)$ at the interpolation points x_j. It agrees with (5.8.44) except for a shift in the summation index and a change in the normalization of the Fourier coefficient that gives rise to the $1/\sqrt{N}$ term.

Example 5.16. A Discrete Fourier Transform. We consider the function $f(x) = \exp(x/4)\cos(5x)$ over the interval $[0, 2\pi]$ with $N = 4$. The interpolation points are $x_1 = 0$, $x_2 = \pi/2$, $x_3 = \pi$, $x_4 = 3\pi/2$. The values of $f(x)$ at the interpolation points are $f(x_1) = 1$, $f(x_2) = 0$, $f(x_3) = -\exp(\pi/4)$, $f(x_4) = 0$.

The trigonometric interpolation polynomial (5.8.49) becomes

$$f_4(x) = \frac{1}{2}\sum_{k=1}^{4} c_k \exp[i(k-1)x], \tag{5.8.50}$$

with the Fourier coefficients c_k determined from (5.8.46) with $N = 4$ to be

$$\begin{cases} c_1 = (1 - \exp(\pi/4))/2 \approx -.59664, \\ c_2 = (1 + \exp(\pi/4))/2 \approx 1.59664, \\ c_3 = (1 - \exp(\pi/4))/2 \approx -.59664, \\ c_4 = (1 + \exp(\pi/4))/2 \approx 1.59664. \end{cases} \tag{5.8.51}$$

Then

$$\exp(x/4)\cos(5x) \approx \frac{1}{2}\sum_{k=1}^{4} c_k \exp[i(k-1)x], \qquad x \in [0, 2\pi]. \tag{5.8.52}$$

At the four interpolation points there is exact equality in (5.8.52).

The c_k represent the discrete Fourier transform of the values $f(x_j)$ and they can be determined by invoking the Maple procedure *FourierTransform*(V), where the vector V is given as $V = [1,\ 0,\ -\exp(\pi/4),\ 0]$. Its entries are the values of $f(x)$ at the four interpolation points. Maple's output is a four-component vector $[-0.596640 + 0.0i, 1.59664 + 0.0i, -0.596640 + 0.0i, 1.59664 + 0.0i]$. The numbers $0.0i$ signify that the imaginary part of each of the c_k equals zero. The result agrees with that given above.

If we apply the procedure *InverseFourierTransform*(W), where W represents the foregoing vector, we obtain the four-component vector $[1.0 + 0.0i, 0.0 + 0.0i, -2.19328 + 0.0i, 0.0 + 0.0i]$. Its entries are the values of $f(x)$ at the four interpolation points, given in floating-point form. Since $-\exp(\pi/4) \approx -2.19328$, the results agree with the above.

The process of determining the discrete Fourier transform and its inverse may involve an exceedingly large number of operations, a single one of which is taken to comprise a (complex) addition and multiplication. The *fast Fourier transform* is a method for determining the Fourier coefficients c_k of the discrete Fourier transform with a greatly reduced number of calculations. The number of operations required to determine the N Fourier coefficients c_k straightforwardly, using the foregoing representation, is found to be N^2. However, if we use the fast Fourier transform to determine the c_k, the number of operations is $2NM$ if $N = 2^M$. This is a tremendous

reduction for large N. For example, if $M = 10$, we have $N = 1024$, $N^2 = 1048576$, and $2NM = 20480$. (In Example 5.16 we had $N = 4$, so that $M = 2$, and no advantage is gained by using the fast Fourier transform.)

We do not exhibit the basis of the fast Fourier transform method, which involves a decomposition of (5.8.42). The procedures FFT and $iFFT$ use the fast Fourier transform and its inverse to determine the c_k and the $f(x_j)$. The procedures can be applied only if N can be expressed as $N = 2^M$ for some integer M. The more recent Maple procedures *FourierTransform* and *InverseFourierTransform* invoke the fast Fourier transform in more general cases.

Apart from its use in trigonometric interpolation and related Fourier series, the discrete Fourier transform can also be used to approximate Fourier transforms numerically. The infinite limits in the Fourier transform integral are replaced by finite negative and positive limits, and the integral is then transformed into (two) integrals, each having zero as the lower limit and a positive upper limit. On approximating each integral using Riemann sums, we obtain expressions that have the form of discrete Fourier transforms or their inverses, each of which will generally have a large value of N. Then the fast Fourier transform method can be used to evaluate the discrete Fourier transforms and thereby approximate the given Fourier transform. We do not present any examples here.

Exercises 5.8

5.8.1. Verify that Maple's definition of the Fourier transform and its inverse are as given in (5.8.1) and (5.8.2).

5.8.2. Use Maple to confirm the results given in the table of Fourier transforms.

5.8.3. Generate a plot of the function in (5.8.8) that exhibits the Gibbs phenomenon.

5.8.4. Use Maple to confirm the results given in the table of Fourier sine transforms.

5.8.5. Use Maple to confirm the results given in the table of Fourier cosine transforms.

5.8.6. Confirm the results in Example 5.13.

5.8.7. Use Maple to confirm the results given in the table of Hankel transforms.

5.8.8. Obtain the solutions of Exercises 5.6.1–5.6.10 using Maple's *laplace* procedure.

5.8.9. Carry out the steps presented in Example 5.14.

5.8.10. Apply the *StatPhase* procedure to obtain (5.8.37).

5.8.11. Apply the procedure *SirMethod* to obtain (5.8.40).

5.8.12. Reproduce the results of Example 5.16.

CHAPTER 6

INTEGRAL RELATIONS

The problems considered in the preceding chapters dealt mostly with partial differential equations, whose coefficients, inhomogeneous terms, and initial and/or boundary data were smooth functions. Consequently, the solutions were expected to be smooth functions as well. It is often the case, however, that the medium for which the differential equation models some physical property is heterogeneous and some of its characteristics, such as density or conductivity, change discontinuously across some region. For example, this situation arises when we consider the longitudinal vibration of a composite rod composed of two rods of different constant densities joined at some point. Then the coefficients in the wave equation that describes the motion of the vibrating rod may be singular at the point where the rods are attached.

In addition, it is of interest to consider problems where the inhomogeneous term in the equation, which may represent a forcing term or a source (or sink), is concentrated over some lower-dimensional region such as a curve or a point. Also, the data for the problem may have discontinuities or singularities. As a result, the solutions of these problems are no longer expected to be smooth functions in general. Thus it becomes necessary to attach a meaning to *solutions* of differential equations that are not differentiable as often as required by the equation. In some cases these solutions

Partial Differential Equations of Appplied Mathematics, Third Edition. By Erich Zauderer

are not even continuous everywhere. (Some of these matters were considered in our discussion of first order equations in Chapter 2.)

This chapter deals with the foregoing questions by showing how the given partial differential equations can be replaced by equivalent *integral relations*. These relations will be derived directly from the partial differential equations and may involve derivatives of the unknown function. In any case, fewer derivatives than are required for the solution of the partial differential equation are needed for the solution of the equivalent integral relation. Thereby, we generalize or weaken, for some problems, the concept of solution of a differential equation and the results obtained are called *generalized* or *weak solutions*.

The wave equation is discussed in some detail and an equivalent *integral wave equation* is derived. We have seen in Chapter 3 that solutions of hyperbolic equations, in contrast to those for elliptic and parabolic equations, do not smooth out discontinuities or singularities in the data. The wave equation is a prototype equation of hyperbolic type, so we concentrate our discussion on it. (*Singular solutions* are considered in Section 10.2 as well.)

Finally, we introduce the concept of *energy integrals* for partial differential equations of each of the three types. We show how they can be used to prove *uniqueness* and *continuous dependence* of the solutions on the data. Our discussion is restricted to second order partial differential equations of the form considered in Chapter 4. However, the ideas and methods introduced can be carried over to other equations.

6.1 INTRODUCTION

In Section 4.1 the parabolic equation (4.1.4) was derived by applying a *balance law* over a certain region. Assuming that all terms occurring in the balance law are continuous (or, more generally, smooth), it is possible to derive the parabolic equation (4.1.4). If the smoothness requirement is not met throughout the region, say there is some subregion across which the properties of the medium undergo a sharp discontinuous change, the arguments leading to (4.1.4) may fail. Then, the balance law can be used to derive *matching conditions* for the solutions of the partial differential equation (4.1.4) that remain valid on both sides of the region of discontinuity. These conditions show how to connect solutions across the region of discontinuity. A similar approach may be applied if the inhomogeneous term of the equation is singular in some region.

Generally speaking, if the PDE is derived from some physical principle applied over an arbitrary region as in Section 4.1, it is possible to take into account the foregoing special situations and incorporate the appropriate results into the derivation. We have derived various equations in Chapter 1 from a different point of view, one from which it is not obvious how to take into account discontinuities or other singularities in our derivation. Therefore, we adopt the following approach. We assume that the partial differential equation is given and convert it into an equivalent *integral relation* (thereby reversing the commonly used method of derivation of the differential equation from the integral relation). On using this approach in Chapter 2 to construct

generalized solutions of quasilinear equations, it was found that different integral relations (each of which is equivalent to the given differential equation when the solutions are smooth) gave rise to different generalized solutions. This nonuniqueness in the determination of generalized solutions does not occur for linear equations, which we deal with exclusively in this chapter. A detailed discussion is given for the hyperbolic equation (4.1.10) and the appropriate results are given without derivation for the parabolic and elliptic equations (4.1.8) and (4.1.9), respectively. No other equations are discussed.

Integral Relation: Hyperbolic PDE

We consider the *hyperbolic equation* (4.1.10) in *two* or *three space dimensions*; that is,

$$\rho(\mathbf{x})\frac{\partial^2 u(\mathbf{x},t)}{\partial t^2} - \nabla\cdot(p(\mathbf{x})\nabla u(\mathbf{x},t)) = \rho(\mathbf{x})F(\mathbf{x},t) - q(\mathbf{x})u(\mathbf{x},t), \qquad (6.1.1)$$

where $\rho(\mathbf{x})$, $p(\mathbf{x})$, and $q(\mathbf{x})$ have the properties given in Chapter 4. ($\mathbf{x}$ is a two- or three-dimensional variable.) Let R be a closed and bounded region in $(\mathbf{x},t)$-space with ∂R as its boundary. We integrate (6.1.1) over R to obtain

$$\iint_R (\rho u_{tt} - \nabla\cdot(p\nabla u))\ dv = \iint_R (\rho F - qu)\ dv. \qquad (6.1.2)$$

Let the gradient operator in $(\mathbf{x},t)$-space be defined as $\tilde{\nabla} = [\nabla, \partial/\partial t]$, so that (6.1.2) may be written as [since $\rho = \rho(\mathbf{x})$]

$$\iint_R \tilde{\nabla}\cdot[p\nabla u, -\rho u_t]\ dv = \iint_R (qu - \rho F)\ dv, \qquad (6.1.3)$$

where $[p\nabla u, -\rho u_t]$ is a three- or four-component vector. Applying the divergence theorem to the first integral in (6.1.3) gives

$$\int_{\partial R} [p\nabla u, -\rho u_t]\cdot\mathbf{n}\ ds = \iint_R (qu - \rho F)\ dv, \qquad (6.1.4)$$

where $\mathbf{n}$ is the exterior unit normal vector to ∂R. Equation (6.1.4) is the equivalent *integral relation* (or, more precisely, the integrodifferential equation) we are seeking.

We now choose a special form for the region R that is commonly specified when (time-dependent) PDEs are derived using a balance law or a similar physical principle. Fixing t, we choose a closed and bounded region $R_{\mathbf{x}}$ in $\mathbf{x}$-space. Then we construct the region R in $(\mathbf{x},t)$-space whose points $(\mathbf{x},t)$ satisfy the conditions $\mathbf{x}\in R_{\mathbf{x}}$ and $t_0 \le t \le t_1$. (The region R is shown in Figure 6.1.) (The region R can also be characterized as a direct product in the form $R = R_{\mathbf{x}} \times [t_0,\ t_1]$.) Let $\partial R_{\mathbf{x}}$ be the boundary of $R_{\mathbf{x}}$, and let $\partial\tilde{R}_{\mathbf{x}}$ be the lateral boundary of R. Also, denote the upper and lower caps of R by ∂R_1 and ∂R_0, respectively (note that $\partial R_0 \equiv R_{\mathbf{x}}$). The exterior unit normal vector $\mathbf{n}$ on these surfaces assumes the following form. With $\mathbf{n}_{\mathbf{x}}$ defined to be the exterior unit normal vector to $\partial R_{\mathbf{x}}$, we find that on $\partial\tilde{R}_{\mathbf{x}}$, $\mathbf{n} = [\mathbf{n}_{\mathbf{x}},\ 0]$, on ∂R_1, $\mathbf{n} = [\mathbf{0},\ 1]$, while on ∂R_0, $\mathbf{n} = [\mathbf{0},\ -1]$.

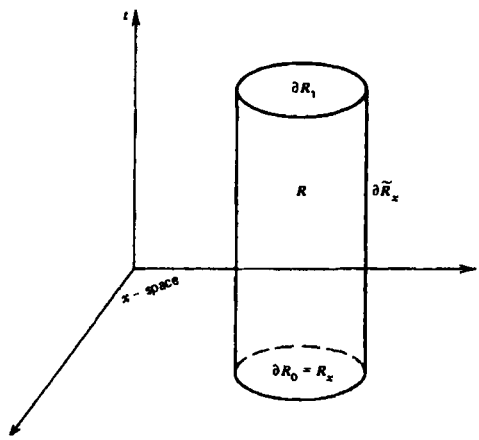

Figure 6.1 The region R.

Then (6.1.4) takes the following form in *two- or three-space dimensions*,

$$\int_{\partial R_0} \rho u_t \, ds - \int_{\partial R_1} \rho u_t \, ds + \int_{\partial \tilde{R}_x} (p \nabla u) \cdot \mathbf{n}_x \, ds = \iint_R (qu - \rho F) \, dv. \quad (6.1.5)$$

The integration over ∂R_0 and ∂R_1 in (6.1.5) is effectively taken over the region R_x. On the lateral surface $\partial \tilde{R}_x$ we have $ds = ds_x \, dt$, where ds_x is the surface differential over ∂R_x, and in the region R we have $dv = d\mathbf{x} \, dt$, where $d\mathbf{x}$ is an area or volume element in two or three dimensions, respectively. We can write (6.1.5) as

$$\int_{R_x} \rho(\mathbf{x})[u_t(\mathbf{x}, t_1) - u_t(\mathbf{x}, t_0)] \, d\mathbf{x} = \int_{t_0}^{t_1} \int_{\partial R_x} p(\mathbf{x}) \frac{\partial u(\mathbf{x}, t)}{\partial n_x} \, ds_x \, dt$$

$$- \int_{t_0}^{t_1} \int_{R_x} (q(\mathbf{x}) u(\mathbf{x}, t) - \rho(\mathbf{x}) F(\mathbf{x}, t)) \, d\mathbf{x} \, dt, \quad (6.1.6)$$

where $\partial u / \partial n_x$ is the exterior normal derivative of u on the surface ∂R_x. In *one space dimension* when R_x is the interval (x_0, x_1), (6.1.6) has the form

$$\int_{x_0}^{x_1} \rho(x)[u_t(x, t_1) - u_t(x, t_0)] \, dx = \int_{t_0}^{t_1} [p(x_1) u_x(x_1, t) - p(x_0) u_x(x_0, t)] \, dt$$

$$- \int_{t_0}^{t_1} \int_{x_0}^{x_1} [q(x) u(x, t) - \rho(x) F(x, t)] \, dx \, dt. \quad (6.1.7)$$

To consider a concrete example, we assume that (6.1.1) describes the *longitudinal vibration of a rod* and that the problem is one-dimensional. Then the integral relation (6.1.7) characterizes the change in momentum of a segment of a rod (x_0, x_1) in the time interval (t_0, t_1). The displacement of the rod at the time t is given by $u(x, t)$, the density is $\rho(x)$, and the tension is $T(x, t) = p(x) u_x(x, t)$—as required

by *Hooke's law*—where $p(x)$ equals *Young's modulus* at the point x. The momentum per unit length equals $\rho(x)u_t(x,t)$ and we put $q(x) = 0$ in (6.1.7). Also, $\rho(x)F(x,t)$ represents the external force density. Thus (6.1.7) equates the change in momentum to the forces acting on the segment $(x_0,\ x_1)$ of the rod. In the limit as $x_0 \to x_1$ and $t_0 \to t_1$, we obtain the partial differential equation (6.1.1) with $q = 0$ and ∇ replaced by $\partial/\partial x$. For the limiting procedure to be valid, $u(x,t)$ must have continuous second derivatives and $\rho(x)$, $p'(x)$, and $F(x,t)$ must be nonsingular. This enables the mean value theorems for derivatives and integrals to be applied to (6.1.7). At points where the limit process is not valid, (6.1.7) does yield *matching conditions*, as will be shown.

Integral Relation: Parabolic and Elliptic PDEs

For completeness we present the appropriate integral relations for the *parabolic* and *elliptic equations* of Section 4.1. Their derivations are similar to those given in the hyperbolic case and are not presented.

For the *parabolic equation* (4.1.4) we have

$$\int_{\partial R} [p\nabla u, -\rho u] \cdot \mathbf{n}\, ds = \iint_R (qu - \rho F)\ dv, \tag{6.1.8}$$

and in a form equivalent to that in (6.1.6),

$$\int_{R_{\mathbf{x}}} \rho(\mathbf{x})[u(\mathbf{x},t_1) - u(\mathbf{x},t_0)]\ d\mathbf{x} = \int_{t_0}^{t_1} \int_{\partial R_{\mathbf{x}}} p(\mathbf{x}) \frac{\partial u(\mathbf{x},t)}{\partial n_{\mathbf{x}}}\ ds_{\mathbf{x}}\ dt$$

$$- \int_{t_0}^{t_1} \int_{R_{\mathbf{x}}} (q(\mathbf{x})u(\mathbf{x},t) - \rho(\mathbf{x})F(\mathbf{x},t))\ d\mathbf{x}\ dt. \tag{6.1.9}$$

For the *elliptic case*, given the equation

$$-\nabla \cdot (p(\mathbf{x})\nabla u(\mathbf{x})) + q(\mathbf{x})u(\mathbf{x}) = \rho(\mathbf{x})F(\mathbf{x}), \tag{6.1.10}$$

we have the equivalent integral relation

$$\int_{\partial R} p(\mathbf{x}) \frac{\partial u(\mathbf{x})}{\partial n} ds = \iint_R (q(\mathbf{x})u(\mathbf{x}) - \rho(\mathbf{x})F(\mathbf{x}))\, dv. \tag{6.1.11}$$

The regions R and $R_{\mathbf{x}}$ and their boundaries are defined as in the foregoing hyperbolic problem for the parabolic case, while $R = R_{\mathbf{x}}$ in the elliptic case. We do not write down the one-dimensional forms of (6.1.8)–(6.1.11).

In each of the foregoing integral relations in two- or three-dimensional $\mathbf{x}$-space, if we choose $R_{\mathbf{x}}$ to be a rectangular region (i.e., $x_0 < x < x_1$, $y_0 < y < y_1$, $z_0 < z < z_1$, in three dimensions), it is an easy matter to retrieve the PDEs from the integral relations in the limit as $R_{\mathbf{x}}$ shrinks to a point and $t_0 \to t_1$. However, the function u and the coefficients, as well as F, must be sufficiently smooth for this limiting process to be valid. This implies the equivalence of the differential and integral representations of the given equations in regions where everything is smooth. In the following sections we show how to use the integral relations in cases where the foregoing limit process breaks down in some region, so that the given differential equation is not valid there.

Exercises 6.1

6.1.1. Derive the one-dimensional form of (6.1.1) from the integral relation (6.1.7) by using mean value theorems for derivatives and integrals, assuming that the functions are sufficiently smooth.

6.1.2. Obtain the integral relation (6.1.8) from the parabolic equation (4.1.4).

6.1.3. Obtain the integral relation (6.1.9) from the parabolic equation (4.1.4).

6.1.4. Specialize (6.1.9) to the case of one space dimension and derive the (one-dimensional) parabolic equation (4.1.5) from it using appropriate mean value theorems.

6.1.5. Derive the integral relation (6.1.11) from the elliptic equation (6.1.10).

6.1.6. Consider a two-dimensional form of (6.1.11) and let R be the rectangle $x_0 < x < x_1$ and $y_0 < y < y_1$. Assuming that the functions are smooth, derive (6.1.10) from (6.1.11) by using appropriate mean value theorems.

6.2 COMPOSITE MEDIA: DISCONTINUOUS COEFFICIENTS

We consider any one of the three *integral relations* (6.1.6), (6.1.9), or (6.1.11) and assume that a Cauchy, an initial and boundary value, or a strict boundary value problem, whichever is appropriate for the equation, is given over a region G in $\mathbf{x}$-space. (The region G may be bounded or unbounded.) Let G be divided into two subregions G_1 and G_2 with S_0 as the lower-dimensional boundary region separating G_1 and G_2. We assume that the properties of the given medium vary discontinuously across S_0 (for all time t if the problem has a time dependence), so that one or more of the functions $\rho(\mathbf{x})$, $p(\mathbf{x})$, or $q(\mathbf{x})$ is discontinuous across S_0 (they are permitted to have, at most, jump discontinuities). As a one-dimensional example, if two strings of different densities are attached at a point x_0 (this point corresponds to S_0) and we consider the equation for the vibration of the composite string, the density $\rho(x)$ will be discontinuous at the point x_0.

At points $\mathbf{x}$ or $(\mathbf{x}, t)$ that are not in the region of discontinuity and where the coefficients and the solution are taken to be smooth functions, we may go to the limit in the integral relations and derive the appropriate PDE. Thus at points $\mathbf{x}$ or $(\mathbf{x}, t)$ such that $\mathbf{x}$ is interior to G_1 or G_2, the PDE is valid. The solution in G_1 is denoted by u_1 and that in G_2 by u_2. We require that the solution of the problem be continuous across S_0. If u is a temperature distribution or is the displacement of a string, we expect it to be continuous in the interior of the region G. Therefore, the first *matching condition* is

$$u_1\big|_{S_0} = u_2\big|_{S_0}. \tag{6.2.1}$$

To obtain a second matching condition, we apply the appropriate integral relation over a region $R_{\mathbf{x}}$ that contains a portion of S_0 in its interior, as pictured in Figure 6.2. (In the hyperbolic and parabolic cases, the region R considered above, is given as the

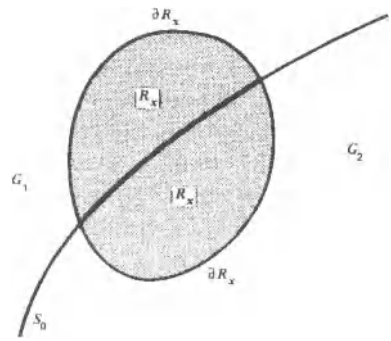

Figure 6.2 The composite region.

direct product $R = R_{\mathbf{x}} \times [t_0,\ t_1]$. In the elliptic case, we put $R = R_{\mathbf{x}}$.) We consider the integral relations (6.1.6), (6.1.9), and (6.1.11). Let the region $R_{\mathbf{x}}$ collapse onto S_0. Since u (as well as u_t, which is an interior derivative in S_0) is continuous across S_0 and ρ, p, q, and possibly, F have, at most, jump discontinuities across S_0, the only contribution that results in the limit comes from the normal derivative terms, as they may have different values on both sides of S_0. Because this result is valid over any portion of S_0, we conclude on the basis of the *duBois-Reymond lemma* (see Exercise 8.1.9) that the second *matching condition* is

$$p_1 \left.\frac{\partial u_1}{\partial n}\right|_{S_0} = p_2 \left.\frac{\partial u_2}{\partial n}\right|_{S_0}, \tag{6.2.2}$$

where $p_1(\mathbf{x})$ and $p_2(\mathbf{x})$ represent the limiting values of $p(\mathbf{x})$ (which may be discontinuous across S_0) as $\mathbf{x}$ approaches a point in S_0 from the subregions G_1 and G_2, respectively. [Both normal derivatives on S_0 in (6.2.2) are taken in the same direction.]

We remark that (6.2.1)–(6.2.2) state that u and $p\,\partial u/\partial n$ are continuous across S_0, and these conditions are valid for all time if u is time dependent. If p is continuous across S_0 but one or more of the other terms in the equation are discontinuous, so that a matching condition is required, then u and $\partial u/\partial n$ are continuous across S_0. In that case, apart from the change in the matching condition (6.2.2), where $p_1 = p_2$, the problem is solved in the same way.

In the one-dimensional case, if S_0 corresponds to the point x_0, and G_1 and G_2 correspond to values of x less than and greater than x_0, respectively, (6.2.1)–(6.2.2) are replaced by

$$u_1\big|_{x_0} = u_2\big|_{x_0}, \quad p_1 \left.\frac{\partial u_1}{\partial x}\right|_{x_0} = p_2 \left.\frac{\partial u_2}{\partial x}\right|_{x_0}. \tag{6.2.3}$$

Cauchy and Initial and Boundary Value Problems

When solving initial and boundary value problems or Cauchy problems in composite media, the matching conditions derived above supplement the initial and boundary conditions that these problems must satisfy.

For the case of an initial and boundary value problem for a *hyperbolic equation* in a *composite medium* with S_0 as the region of discontinuity and G given as the union of G_1, S_0, and G_2 (see Figure 6.3), we have the PDEs

$$\begin{cases} \rho_1 \partial^2 u_1/\partial t^2 - \nabla \cdot (p_1 \nabla u_1) + q u_1 = \rho_1 F_1, & \mathbf{x} \in G_1,\ t > 0, \\ \rho_2 \partial^2 u_2/\partial t^2 - \nabla \cdot (p_2 \nabla u_2) + q u_2 = \rho_2 F_2, & \mathbf{x} \in G_2,\ t > 0. \end{cases} \tag{6.2.4}$$

The initial conditions are

$$\begin{cases} u_1(\mathbf{x}, 0) = f_1(\mathbf{x}), & \partial u_1(\mathbf{x}, 0)/\partial t = g_1(\mathbf{x}), \quad \mathbf{x} \in G_1, \\ u_2(\mathbf{x}, 0) = f_2(\mathbf{x}), & \partial u_2(\mathbf{x}, 0)/\partial t = g_2(\mathbf{x}), \quad \mathbf{x} \in G_2. \end{cases} \tag{6.2.5}$$

The boundary conditions, given on the boundary S of the region G, are

$$\begin{cases} \alpha_1 u_1 + \beta_1 \partial u_1/\partial n\big|_S = B_1(\mathbf{x}, t), & \mathbf{x} \in \partial G_1, \\ \alpha_2 u_2 + \beta_2 \partial u_2/\partial n\big|_S = B_2(\mathbf{x}, t), & \mathbf{x} \in \partial G_2, \end{cases} \tag{6.2.6}$$

and the matching conditions are

$$u_1\big|_{S_0} = u_2\big|_{S_0}, \quad p_1 \left.\frac{\partial u_1}{\partial n}\right|_{S_0} = p_2 \left.\frac{\partial u_2}{\partial n}\right|_{S_0}. \tag{6.2.7}$$

The subscripts 1 and 2 signify that the functions are evaluated in the regions G_1 and G_2, respectively, or on their boundaries.

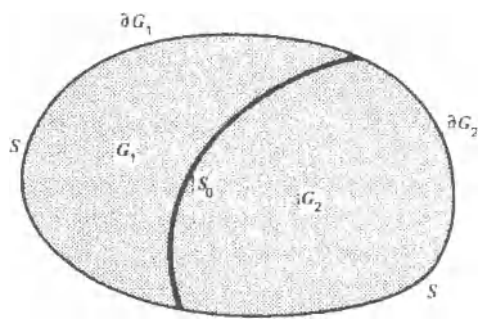

Figure 6.3 The region G.

If a Cauchy problem is considered for the *hyperbolic equation*, the region G is of infinite extent, so the boundary conditions (6.2.6) are dropped but the other equations are retained. In either case, we can solve the problem by considering each of the equations in (6.2.4) for the regions G_1 and G_2 separately. We apply the initial and

boundary conditions for each subregion with the additional conditions $u_1 = \hat{u}_1$ and $u_2 = \hat{u}_2$ on S_0 where $\hat{u}_1$ and $\hat{u}_2$ are as yet unknown. Once u_1 and u_2 are determined, $\hat{u}_1$ and $\hat{u}_2$ are found by applying the matching conditions (6.2.7).

For the case of an initial and boundary value problem for a *parabolic equation* in a *composite medium* with S_0 as the region of discontinuity, we replace $\partial^2 u_i/\partial t^2$, $i = 1, 2$, by $\partial u_i/\partial t$, $i = 1, 2$, in (6.2.4) and drop the second initial condition $\partial u_i(\mathbf{x}, 0)/\partial t = g_i(\mathbf{x})$, $i = 1, 2$, in (6.2.5). For the case of a boundary value problem for an *elliptic equation* in a *composite medium* with S_0 as the region of discontinuity, the time derivatives in (6.2.4) and the initial conditions in (6.2.5) are omitted. The other conditions are unchanged for both cases.

Example 6.1. The Longitudinal Vibration of an Infinite Composite Rod. We consider the longitudinal vibration of a composite rod of infinite extent. The rod is made up of two homogeneous rods joined at $x = 0$, each of which has a constant density ρ and a constant Young's modulus p (see the foregoing discussion). Here the region G is the infinite interval $-\infty < x < \infty$ with G_1 given as $-\infty < x < 0$ and G_2 given as $0 < x < \infty$. The discontinuity region S_0 is the point $x = 0$. We equate q and F to zero in (6.2.4) and replace ∇ by $\partial/\partial x$. Also, ρ_1, ρ_2 and p_1, p_2 are prescribed constants for the problem.

The appropriate *wave equations* for $u(x, t)$ in the regions G_1 and G_2 are

$$\frac{\partial^2 u_i(x,t)}{\partial t^2} - c_i^2 \frac{\partial^2 u_i(x,t)}{\partial x^2} = 0, \qquad i = 1, 2, \tag{6.2.8}$$

where $c_i^2 = p_i/\rho_i$, $i = 1, 2$. Thus there are different speeds of wave propagation c_1 and c_2 in the two regions G_1 and G_2, respectively.

We assume that a wave $f(t - x/c_1)$ is approaching the junction point $x = 0$, as t approaches zero from negative values, and we want to determine the resulting reflected and transmitted waves. To do so, we formulate this as a *Cauchy problem* with the initial data

$$u_1(x, 0) = f\left(-\frac{x}{c_1}\right), \qquad \frac{\partial u_1(x,0)}{\partial t} = f'\left(-\frac{x}{c_1}\right), \qquad x < 0, \tag{6.2.9}$$

$$u_2(x, 0) = 0, \qquad \frac{\partial u_2(x,0)}{\partial t} = 0, \qquad x > 0. \tag{6.2.10}$$

Note that a solution of (6.2.8) with $i = 1$ and the initial data (6.2.9) is just $u_1(x, t) = f(t - x/c_1)$—that is, a wave traveling to the right with speed c_1. For small t, this is assumed to be the only disturbance in the region G_1. [This is the case if we assume that for some $x_1 < 0$, we have $f(-x/c_1) = 0$ for $x > x_1$.] As t increases, the wave reaches the junction point $x = 0$ and gives rise to a reflected and transmitted wave. In addition to (6.2.9)–(6.2.10), the solutions u_1 and u_2 must satisfy the matching conditions (6.2.3), where $x_0 = 0$.

The general solutions of the wave equations (6.2.8) are

$$u_i(x, t) = f_i\left(t - \frac{x}{c_i}\right) + g_i\left(t + \frac{x}{c_i}\right), \qquad i = 1, 2. \tag{6.2.11}$$

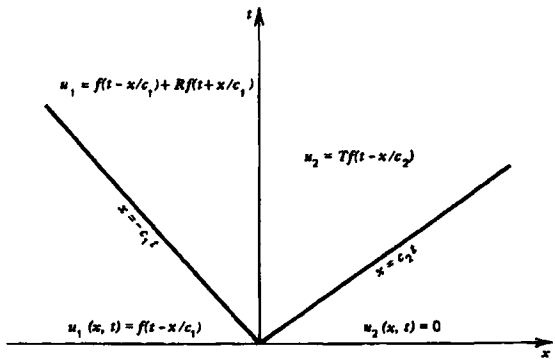

Figure 6.4 Vibration of a composite rod.

From (6.2.9)–(6.2.10) we obtain four equations for f_1, f_2, g_1, and g_2. Differentiating these equations as needed, we easily conclude that $g_1(z) = 0$ for $z < 0$, $g_2(z) = 0$ for $z > 0$, $f_1(z) = f(z)$ for $z > 0$ and $f_2(z) = 0$ for $z < 0$.

The matching conditions yield a simultaneous system of equations for $f_2(t)$ and $g_1(t)$ with $t > 0$. Solving these equations and introducing the results into (6.2.11) yields the solution for $t > 0$, as

$$u_1(x,t) = f\left(t - \frac{x}{c_1}\right) + Rf\left(t + \frac{x}{c_1}\right),\ x < 0,\ u_2(x,t) = Tf\left(t - \frac{x}{c_2}\right),\ x > 0, \tag{6.2.12}$$

where the *reflection coefficient* R and the *transmission coefficient* T are

$$R = \frac{\sqrt{\rho_1 p_1} - \sqrt{\rho_2 p_2}}{\sqrt{\rho_1 p_1} + \sqrt{\rho_2 p_2}}, \qquad T = \frac{2\sqrt{\rho_1 p_1}}{\sqrt{\rho_1 p_1} + \sqrt{\rho_2 p_2}}. \tag{6.2.13}$$

For small t the solution is $u_1 = f(t - x/c_1)$ and $u_2 = 0$. The *incident wave* $f(t - x/c_1)$ travels to the right (i.e., toward $x = 0$) with speed c_1. When it reaches the interface $x = 0$, two additional waves arise: the *reflected wave* $Rf(t + x/c_1)$ which travels to the left with speed c_1, and the transmitted wave $Tf(t - x/c_2)$, which travels to the right with speed c_2. Apart from a constant factor, each wave has the same waveform $f(x)$. The *reflection* and *transmission coefficients* R and T each modify the basic waveform $f(x)$. We note that $1 + R = T$, so that, in a sense, there is a conservation of the total waveform over the entire length of the rod. If $\rho_1 p_1 = \rho_2 p_2$, we have $R = 0$ and $T = 1$, so that the incident wave is transmitted without undergoing reflection. If $\rho_2 p_2 \gg \rho_1 p_1$, we obtain $R \approx -1$ and $T \approx 0$. This represents the case of *total reflection*. No wave is transmitted to the right of the junction point $x = 0$. Only a reflected wave is generated. The solution is described in Figure 6.4.

Eigenvalue Problems and Eigenfunction Expansions

It was indicated above that to solve Cauchy or initial and boundary value problems in composite media, one should first solve two subproblems in the two subregions G_1 and G_2 with arbitrary, unspecified boundary data on the discontinuity region S_0. The solutions of the full problem are then determined by applying the matching conditions.

There is an alternative approach that we now discuss for problems in bounded spatial regions. In the absence of a discontinuity region, such problems were solved in Chapter 4 with the use of *eigenfunction expansions*. The eigenfunctions were determined from the eigenvalue problem associated with the related homogeneous equation and homogeneous boundary conditions. We now show that it is possible to construct eigenvalue problems for composite media. The results can be used to solve boundary and initial and boundary value problems for composite media by means of eigenfunction expansions.

Again considering the region G with S_0 as a discontinuity region for the coefficients, we introduce the following *eigenvalue problem* for $v(\mathbf{x})$,

$$L[v(\mathbf{x})] = -\nabla \cdot (p(\mathbf{x})\nabla v(\mathbf{x})) + q(\mathbf{x})v(\mathbf{x}) = \lambda\rho(\mathbf{x})v(\mathbf{x}),\ \mathbf{x} \in G,\ \mathbf{x} \notin S_0, \quad (6.2.14)$$

$$\alpha(\mathbf{x})v(\mathbf{x}) + \beta(\mathbf{x}) \left.\frac{\partial v}{\partial n}\right|_S = 0, \quad (6.2.15)$$

$$v(\mathbf{x}) \text{ and } p\frac{\partial v(\mathbf{x})}{\partial n} \text{ are continuous across } S_0. \quad (6.2.16)$$

The coefficients $p(\mathbf{x})$, $q(\mathbf{x})$, and $\rho(\mathbf{x})$ may have jump discontinuities across S_0. All the significant properties of eigenvalues and eigenfunctions given in Chapter 4 remain valid for this eigenvalue problem.

We demonstrate the *orthogonality* of eigenfunctions corresponding to different eigenvalues in the one-dimensional case, with G given as the interval $0 < x < l$. With L defined as in (4.3.1) and v_i and v_j as eigenfunctions corresponding to different eigenvalues λ_i and λ_j, $(\lambda_i \neq \lambda_j)$, (4.3.16) yields

$$\int_0^{x_0} \frac{d}{dx}\left(pv_jv_i' - pv_iv_j'\right) dx + \int_{x_0}^{l} \frac{d}{dx}\left(pv_jv_i' - pv_iv_j'\right) dx = 0, \quad (6.2.17)$$

where x_0 is the point of discontinuity (it corresponds to S_0) and $0 < x_0 < l$. This follows because the contribution to each integral vanishes at the boundary points $x = 0$ and $x = l$, as before. At the point $x = x_0$, the functions pv_iv_j' and pv_jv_i' both occur with opposite signs and their contribution vanishes since v_i, v_j, pv_i', and pv_j' are assumed to be continuous at $x = x_0$.

The inner product and the norm for this eigenvalue problem are defined as

$$(v_i, v_j) = \int_0^{x_0} \rho v_i v_j\, dx + \int_{x_0}^{l} \rho v_i v_j\, dx, \quad \|v_i\| = \sqrt{(v_i, v_i)}. \quad (6.2.18)$$

Then, on using (4.3.17) and (6.2.17), we conclude that $v_i(x)$ and $v_j(x)$ are orthogonal. The other properties listed and proven in Section 4.3 for the Sturm-Liouville problem can be proven for the present problem as well.

Next we solve the problem of heat conduction in an insulated composite, finite rod with the ends kept at zero temperature. To do so, we use the eigenvalue problem that results if the problem is solved by separation of variables.

Example 6.2. Heat Conduction in a Finite Composite Rod. We consider heat conduction in an insulated piecewise homogeneous rod of length l, with the ends kept at zero temperature for all time. With the rod positioned in the interval $0 \le x \le l$ and the junction of the two homogeneous portions occurring at the interior point $x = x_0$, we have the following initial and boundary value problem for the temperature $u(x,t)$ of the rod, with $u = u_1(x,t)$ for $x < x_0$ and $u = u_2(x,t)$ for $x > x_0$,

$$\begin{cases} \rho_1 \, \partial u_1(x,t)/\partial t - p_1 \, \partial^2 u_1(x,t)/\partial x^2 = 0, & 0 < x < x_0, \; t > 0, \\ \rho_2 \, \partial u_2(x,t)/\partial t - p_2 \, \partial^2 u_2(x,t)/\partial x^2 = 0, & x_0 < x < l, \; t > 0. \end{cases} \tag{6.2.19}$$

The coefficients ρ_1, ρ_2, p_1, and p_2 are all constant. (See Section 4.1 for their description.) The boundary and initial conditions are

$$u_1(0,t) = 0, \; u_2(l,t) = 0, \; t > 0, \quad u(x,0) = f(x), \; 0 < x < l. \tag{6.2.20}$$

The matching conditions at $x = x_0$ are

$$u_1(x_0,t) = u_2(x_0,t), \; p_1 \, \partial u_1(x_0,t)/\partial x = p_2 \, \partial u_2(x_0,t)/\partial x, \quad t > 0. \tag{6.2.21}$$

Using separation of variables we are led to the following eigenvalue problem for the function $v(x)$ with $v = v_1(x)$ for $x < x_0$ and $v = v_2(x)$ for $x > x_0$:

$$p_1 v_1''(x) + \lambda \rho_1 v_1(x), \; 0 < x < x_0, \quad p_2 v_2''(x) + \lambda \rho_2 v_2(x), \; x_0 < x < l. \tag{6.2.22}$$

The homogeneous boundary conditions are

$$v_1(0) = 0, \qquad v_2(l) = 0, \tag{6.2.23}$$

and the matching conditions at $x = x_0$ are

$$v_1(x_0) = v_2(x_0), \quad p_1 \, v_1'(x_0) = p_2 \, v_2'(x_0). \tag{6.2.24}$$

To determine the eigenfunctions, we solve (6.2.22) for $v_1(x)$ and $v_2(x)$ and apply the boundary conditions (6.2.23). This yields

$$v_1(x) = A \, \sin\left(\sqrt{\frac{\lambda \rho_1}{p_1}} \, x\right), \qquad v_2(x) = B \, \sin\left(\sqrt{\frac{\lambda \rho_2}{p_2}} \, (x - l)\right), \tag{6.2.25}$$

where A and B are arbitrary constants. The matching conditions yield a simultaneous homogeneous linear system for A and B. As we require a nonzero solution, the determinant of the coefficients must vanish. This yields

$$c_1 \rho_1 \, \cot\left(\frac{\sqrt{\lambda}}{c_1} x_0\right) + c_2 \rho_2 \, \cot\left(\frac{\sqrt{\lambda}}{c_2} (l - x_0)\right) = 0, \tag{6.2.26}$$

where $c_1 = \sqrt{p_1/\rho_1}$ and $c_2 = \sqrt{p_2/\rho_2}$. This equation determines the eigenvalues λ_k $(k = 1, 2, \ldots)$. We do not solve this equation but note that the λ_k are real, countably infinite, and positive according to the general theory.

The eigenfunctions $v^{(k)}(x)$ corresponding to the λ_k, for $k = 1, 2, \ldots$, are

$$v^{(k)}(x) = \frac{\sin(\sqrt{\lambda_k}\, x/c_1)}{\sin(\sqrt{\lambda_k}\, x_0/c_1)},\ x < x_0,\ v^{(k)}(x) = \frac{\sin(\sqrt{\lambda_k}(l-x)/c_2)}{\sin(\sqrt{\lambda_k}(l-x_0)/c_2)},\ x_0 < x. \tag{6.2.27}$$

For the norm of $v^{(k)}(x)$, we have

$$||v^{(k)}(x)||^2 = \frac{\rho_1 x_0}{2\sin^2(\sqrt{\lambda_k}\, x_0/c_1)} + \frac{\rho_2(l-x_0)}{2\sin^2(\sqrt{\lambda_k}(l-x_0)/c_2)}. \tag{6.2.28}$$

The orthogonality of the set of eigenfunctions $v^{(k)}(x)$ may be verified directly, and using (6.2.28), the set may be orthonormalized. The completeness property of the eigenfunctions asserts that a function $\phi(x)$ under suitable conditions can be expanded in a series of eigenfunctions.

Applying these results to the initial and boundary value problem for the heat equation (6.2.19)–(6.2.21), separation of variables, or finite Fourier transforms, yields the series representation for the solution $u(x,t)$ in the form

$$u(x,t) = \sum_{k=1}^{\infty} \alpha_k e^{-\lambda_k t} v^{(k)}(x), \tag{6.2.29}$$

where the λ_k and $v^{(k)}(x)$ are the eigenvalues and eigenfunctions of the problem (6.2.22)–(6.2.24). The α_k are the Fourier coefficients of the initial temperature distribution $f(x)$; that is, $u(x,0) = \sum_{k=1}^{\infty} \alpha_k v^{(k)}(x) = f(x)$, and they are given as

$$\alpha_k = \frac{1}{||v^{(k)}(x)||^2}\left[\rho_1 \int_0^{x_0} f(x) v_1^{(k)}(x)\, dx + \rho_2 \int_{x_0}^{l} f(x) v_2^{(k)}(x)\, dx\right] \tag{6.2.30}$$

for $k = 1, 2, \ldots$. The $v_1^{(k)}(x)$ and $v_2^{(k)}(x)$ are given in (6.2.27), and $||v^{(k)}(x)||^2$ is given in (6.2.28).

The series (6.2.29) represents the formal solution of the problem. The nonhomogeneous version of the foregoing heat conduction problem can be solved using the eigenfunctions $v^{(k)}(x)$ and the finite Fourier transform. Similarly, related problems for hyperbolic and elliptic equations also can be solved in terms of these eigenfunctions.

Exercises 6.2

6.2.1. Formulate, in the manner of (6.2.4)–(6.2.7), an initial and boundary value problem for the parabolic equations (4.1.4) and (6.1.8) if S_0 is a discontinuity region.

6.2.2. Formulate, in the manner of (6.2.4)-(6.2.7), a boundary value problem for the elliptic equations (6.1.10)–(6.1.11) if S_0 is a discontinuity region.

6.2.3. (a) Show that if $p_2 \approx 0$ in Example 6.1, then $R \approx 1$ and $T \approx 2$ in (6.2.13). Demonstrate that the (approximate) solution to this problem (i.e., with $p_2 \approx 0$) can be obtained by solving for $u_1(x,t)$ in $x < 0$ with the boundary condition $\partial u_1(0,t)/\partial x = 0$, and then obtaining $u_2(x,t)$ in $x > 0$, by using the boundary condition $u_2(0,t) = u_1(0,t)$. (b) Show that if $p_2 \to \infty$, we have $R \to -1$ and $T \to 0$, so that there is no transmitted wave. Introduce the assumption that $p_2 \gg p_1$ into the matching condition for Example 6.1, set up (approximate) boundary values for the two regions $x < 0$ and $x > 0$, and determine that their solutions are consistent with the results obtained from (6.2.12) for large p_2.

6.2.4. Consider the problem of heat conduction in an infinite composite rod composed of two homogeneous rods connected at $x = 0$. Put $u = u_1(x,t)$ for $x < 0$ and $u = u_2(x,t)$ for $x > 0$ and assume that they satisfy the equations $\rho_1\, \partial u_1(x,t)/\partial t - p_1\, \partial^2 u_1(x,t)/\partial x^2 = 0$, $-\infty < x < 0$, $t > 0$, $\rho_2\, \partial u_2(x,t)/\partial t - p_2\, \partial^2 u_2(x,t)/\partial x^2 = 0$, $0 < x < \infty$, $t > 0$. Let $u_1(x,0) = A$ and $u_2(x,0) = B$, where A and B are constants and apply the matching conditions (6.2.21) at $x_0 = 0$. Determine the temperature $u(x,t)$ for $t > 0$. *Hint*: Let $u_1(x,t) = \alpha_1 + \beta_1 \operatorname{erf}(-x/2c_1\sqrt{t})$, $u_2(x,t) = \alpha_2 + \beta_2 \operatorname{erf}(x/2c_2\sqrt{t})$, where α_1 and β_1 are constants and $c_i^2 = p_i/\rho_i$ for $i = 1, 2$. The function $\operatorname{erf}(z)$ is the error function integral defined in (5.3.26).

6.2.5. Reconsider the problem of Exercise 6.2.4 and replace the constant initial data by the arbitrary initial conditions $u_1(x,0) = f_1(x)$ and $u_2(x,0) = f_2(x)$. Solve the respective problems for $u_1(x,t)$ and $u_2(x,t)$ by assuming that $p_1\, \partial u_1(0,t)/\partial x = p_2\, \partial u_2(0,t)/\partial x = g(t)$ is a known function. The solution of the initial and boundary value problems for $u_1(x,t)$ and $u_2(x,t)$ may then be obtained from the result (5.3.37). Use the matching condition $u_1(0,t) = u_2(0,t)$ to show that $g(t)$ is the solution of the *Abel integral equation* $G(t) = \int_0^t g(\tau)/\sqrt{t-\tau}\, d\tau$, where $G(t)$ is a known function. The solution of this equation is $g(t) = (1/\pi)\, d/dt \left(\int_0^t G(\tau)/\sqrt{t-\tau}\, d\tau \right)$.

6.2.6. Show that the solution of the problem in Exercise 6.2.4 may be obtained by applying the Laplace transform in the time variable.

6.2.7. Adapt the discussion of Section 4.2 to show that the operator $(1/\rho)L$ associated with the eigenvalue problem (6.2.14)–(6.2.16) is a positive operator. Conclude that the eigenvalues for the problem are nonnegative.

6.2.8. Discuss the eigenvalue problem (6.2.22)–(6.2.24) if the boundary conditions (6.2.23) are replaced by $v_1'(0) = v_2'(l) = 0$. Obtain an equation for the determination of the eigenvalues and obtain an orthonormal set of eigenfunctions. Show that $\lambda = 0$ is an eigenvalue and $v(x) = 1$ is an eigenfunction for this problem.

6.2.9. Expand the function $f(x) = 1$ in a series of the eigenfunctions $v^{(k)}(x)$ given in (6.2.27) and use this result to obtain a solution of the problem (6.2.19)–(6.2.21) if $u(x,0) = 1$.

6.2.10. Consider the initial and boundary value problem for the hyperbolic equations $\rho_1 \partial^2 u_1(x,t)/\partial t^2 - p_1 \partial^2 u_1(x,t)/\partial x^2 = 0, 0 < x < x_0, t > 0, \rho_2 \partial^2 u_2(x,t)/\partial t^2 - p_2 \partial^2 u_2(x,t)/\partial x^2 = 0, x_0 < x < l, t > 0$, in a composite medium. Let $u(x,0) = f(x), u_t(x,0) = g(x), 0 < x < l$, where $u = u_1(x,t)$ for $x < x_0$ and $u = u_2(x,t)$ for $x > x_0$. Assuming that $u_1(x,t)$ and $u_2(x,t)$ satisfy the boundary conditions (6.2.20) and the matching conditions (6.2.21), apply separation of variables and show how the solution can be expressed in terms of the eigenfunctions obtained in Example 6.2.

6.2.11. Develop a finite Fourier transform approach for the eigenfunctions obtained in Example 6.2 for the purpose of solving the problem (6.2.19)–(6.2.21) if the equation and the boundary conditions are inhomogeneous.

6.2.12. Show how Laplace transforms can be used to solve the problem (6.2.19)–(6.2.21).

6.2.13. Let $u_1(x,y)$ and $u_2(x,y)$ satisfy the Helmholtz equations $\nabla^2 u_1(x,y) + k_1^2 u_1(x,y) = 0, -\infty < y < \infty, x < 0, \nabla^2 u_2(x,y) + k_2^2 u_2(x,y) = 0, -\infty < y < \infty, x > 0$, with k_1 and k_2 as constants, and let $u_1(x,y)$ and $u_2(x,y)$ be related across $x = 0$ by the matching conditions $u_1(0,y) = u_2(0,y), \partial u_1(0,y)/\partial x = \partial u_2(0,y)/\partial x$. Represent $u_1(x,y)$ and $u_2(x,y)$ as $u_1(x,y) = \exp[ik_1(x\cos\theta + y\sin\theta)] + R\exp[ik_1(x\cos\hat{\theta} + y\sin\hat{\theta})], u_2(x,y) = T\exp[ik_2(x\cos\phi + y\sin\phi)]$, where $-\pi/2 < \theta, \phi < \pi/2$, and $\pi/2 < \hat{\theta} < 3\pi/2$. The angle θ is assumed to be specified. Determine the constants $\hat{\theta}$, ϕ, R, and T such that the Helmholtz equations and the matching conditions are satisfied. This problem characterizes the scattering of a plane wave at an interface. A general discussion of scattering problems is given in Chapter 10. The restrictions on the angles $\hat{\theta}$ and ϕ are required to guarantee a unique solution to this problem and to correspond to physically motivated *radiation conditions*. The relationships between the angle of incidence θ and the angles of reflection and refraction $\hat{\theta}$ and ϕ, respectively, are known as *Snell's laws*.

6.3 SOLUTIONS WITH DISCONTINUOUS FIRST DERIVATIVES

It has been shown in Section 3.2 that discontinuities in second derivatives for second order partial differential equations must occur across characteristic curves or surfaces. A large class of second order partial differential equations were shown in Section 6.1 to have equivalent integral relations in the case where the coefficients and the unknown functions were smooth. Since the integral relations contain, at most, first derivatives of the unknown functions, they can be used to attach a meaning to solutions of differential equations that are continuous and have only piecewise continuous first derivatives.

We begin by considering the integral form (6.1.4) of the *hyperbolic equation* (6.1.1) in three dimensions. Assume that a solution $u(\mathbf{x},t)$ of (6.1.4) is continuous across a surface $\phi(\mathbf{x},t) = 0$ but has jump discontinuities in its first derivatives across that surface. [Here $\mathbf{x}$ represents (x,y,z).] Denote the surface by S_0 and assume that it divides the given region R into the subregions R_1 and R_2. The solution $u(\mathbf{x},t)$ is

assumed to be smooth in R_1 and R_2, and the functions ρ, p, q, and F in (6.1.1) and (6.1.4) are assumed to be smooth throughout R. A unit normal vector $\mathbf{n}$ to S_0 can be given in terms of the space-time gradient vector $\tilde{\nabla}\phi$ as

$$\mathbf{n} = \frac{\tilde{\nabla}\phi}{|\tilde{\nabla}\phi|} = \frac{[\nabla\phi,\ \phi_t]}{\sqrt{(\nabla\phi)^2 + \phi_t^2}}. \tag{6.3.1}$$

We apply the integral relation (6.1.4) over R_1 and R_2 and allow these regions to collapse onto S_0. With the exception of u_t and ∇u, all functions in the integral relation (6.1.4) including $\mathbf{n}(\mathbf{x}, t)$ are continuous across S_0. Therefore, we easily obtain

$$\int_{S_0} \frac{p[\nabla u] \cdot \nabla\phi - \rho[u_t]\phi_t}{|\tilde{\nabla}\phi|}\, ds = 0, \tag{6.3.2}$$

where $[V]$ denotes the jump in V across S_0. The same expression can be obtained over any arbitrary subregion of S_0, so that the integrand itself must vanish in view of the duBois-Reymond lemma (see Exercise 8.1.9). Thus

$$p[\nabla u] \cdot \nabla\phi - \rho[u_t]\phi_t = 0. \tag{6.3.3}$$

Now $u(\mathbf{x}, t)$ is continuous across S_0 and is continuously differentiable in the interior of S_0—that is, directional derivatives of $u(\mathbf{x}, t)$ in tangential directions on S_0 are continuous. As a result, we can obtain additional equations for the jumps in u and ∇u across S_0. (We note that there are four derivatives to be determined in the three-dimensional case, so that four equations are needed.) We observe that the vectors $\mathbf{n_1} = [\phi_t, 0, 0, -\phi_x]$, $\mathbf{n_2} = [0, \phi_t, 0, -\phi_y]$, and $\mathbf{n_3} = [0, 0, \phi_t, -\phi_z]$ are linearly independent and are all orthogonal to the normal vector $\mathbf{n}$. Thus the scalar product of the space-time gradient vector $\tilde{\nabla}u$ into any of the vectors $\mathbf{n}_1$, $\mathbf{n}_2$, and $\mathbf{n}_3$ yields an interior derivative in S_0. Denoting $u(\mathbf{x}, t)$ by $u_1(\mathbf{x}, t)$ in R_1 and by $u_2(\mathbf{x}, t)$ in R_2, we differentiate u_1 and u_2 in each of the directions $\mathbf{n}_1$, $\mathbf{n}_2$, and $\mathbf{n}_3$. The resulting interior derivatives of u are continuous across S_0 so that the difference of the results for u_1 and u_2 must vanish. This yields the following three additional equations:

$$[u_x]\phi_t - [u_t]\phi_x = 0,\ [u_y]\phi_t - [u_t]\phi_y = 0,\ [u_z]\phi_t - [u_t]\phi_z = 0 \tag{6.3.4}$$

for the jumps in ∇u and u_t across S_0.

The four equations (6.3.3)–(6.3.4) are a homogeneous linear system for the jumps $[u_x]$, $[u_y]$, $[u_z]$, and $[u_t]$ across S_0. As we assume that one or more of these jumps is nonzero, the determinant of the coefficients of this system must vanish to guarantee a nonzero solution. The value of the determinant is

$$\phi_t^2 \left[\rho\phi_t^2 - p(\nabla\phi)^2\right] = 0. \tag{6.3.5}$$

Equating the bracketed term in (6.3.5) to zero yields the result

$$\rho\phi_t^2 - p(\nabla\phi)^2 = 0. \tag{6.3.6}$$

On comparing with (3.3.17) we find that (6.3.6) is the *characteristic equation* for the hyperbolic equation (6.1.1). Thus $\phi(\mathbf{x}, t) = 0$ must be a *characteristic surface*, and discontinuities in first derivatives can occur only across characteristic surfaces. It is shown in the exercises that $\phi_t = 0$, which is also a consequence of (6.3.5), does not lead to a surface of discontinuity. We conclude that a continuous, piecewise continuously differentiable solution of (6.1.1) or, more precisely, of (6.1.4), satisfies (6.1.1) in regions where it is smooth and the jumps in the first derivatives must satisfy (6.3.3) on all characteristic discontinuity surfaces.

Equation (6.3.6) can be expressed (apart from a nonzero factor) as a dot product of the normal vector $\mathbf{n}$ [i.e., (6.3.1)] to the characteristic surface $\phi = 0$, and the vector $\mathbf{t} = [p\nabla\phi, \ -\rho\phi_t]$. Since $\mathbf{n} \cdot \mathbf{t} = 0$, the two vectors are orthogonal and $\mathbf{t}$ is tangent to the characteristic surface. But (6.3.3) can be written as $[\tilde{\nabla} u] \cdot \mathbf{t} = 0$, and this states that the (interior) derivative of $[u]$ in the direction of $\mathbf{t}$ is zero on the characteristic, consistent with the fact that u is continuous there. Consequently, (6.3.3) is satisfied automatically if the solution u is continuous across the characteristic but has jumps in the first derivatives.

For the *parabolic case* (6.1.8), where u_t does not appear in the integral relation, the foregoing analysis yields (6.3.3) with the u_t term absent. The further calculations are the same as before, and instead of (6.3.5), we obtain

$$p\phi_t^2(\nabla\phi)^2 = 0, \tag{6.3.7}$$

which characterizes the surface S_0 across which u_t and ∇u can have jump discontinuities. Again $\phi_t = 0$ does not yield a surface of discontinuity (see the exercises). This leaves $(\nabla\phi)^2 = 0$, which is the characteristic equation for the parabolic equation (4.1.4). We conclude that $\phi = \phi(t)$, so that the surfaces t = constant are the characteristics. As a result, ∇u is continuous across the characteristics, since it represents interior differentiation on the surfaces t = constant. Thus it appears that only u_t can have a jump across the characteristics. However, by considering the jump in the parabolic equation (4.1.4) across t = constant (in the manner of Section 3.2), we obtain $\rho[u_t] = 0$, since u, its interior derivatives, and all other terms in the equation are assumed to be continuous across t = constant. Consequently, u_t must also be continuous across t = constant.

For the *elliptic case* (6.1.11) we obtain the result (6.3.7) with the ϕ_t^2 term absent. Since the surface S_0 is given as $\phi(x, y, z) = 0$ in this case (i.e., there is no t-dependence), we again conclude that there are no real characteristics or discontinuity surfaces for the elliptic equation (6.1.10).

Example 6.3. The Cauchy Problem for the Wave Equation. We consider a Cauchy problem for the one-dimensional *wave equation* with discontinuous initial data. Let $u(x, t)$ satisfy

$$u_{tt}(x, t) - c^2 u_{xx}(x, t) = 0, \qquad -\infty < x < \infty, \ t > 0, \tag{6.3.8}$$

with the initial conditions

$$u(x,0)=0, \qquad -\infty<x<\infty, \quad u_t(x,0)=\begin{cases}0, & x<0\\ 1, & x\geq 0.\end{cases} \tag{6.3.9}$$

Since $u(x,t)$ is initially continuous and $u_t(x,0)$ has a jump discontinuity, we expect the solution to have discontinuous first derivatives on the characteristics that issue from the point $(x,t)=(0,0)$.

A formal application of *d'Alembert's solution* (2.2.21) yields

$$u(x,t)=\frac{1}{2c}\int_{x-ct}^{x+ct} u_t(s,0)\,ds, \tag{6.3.10}$$

with $u_t(x,0)$ given as in (6.3.9). We divide the half-plane $t>0$ into three sectors as shown in Figure 6.5. The sectors I, II, and III are separated by the characteristic curves $x=\pm ct$ that issue from the initial discontinuity point $(0,0)$ of the velocity $u_t(x,t)$. The integration in (6.3.10) is easily carried out and yields

$$u(x,t)=\begin{cases}t, & x-ct>0, & \text{I}\\ (t+x/c)/2, & x-ct\leq 0\leq x+ct, & \text{II}\\ 0, & x+ct<0, & \text{III.}\end{cases} \tag{6.3.11}$$

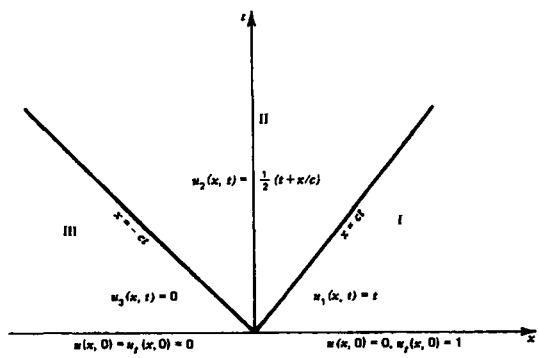

Figure 6.5 The solution of the Cauchy problem.

Clearly, $u(x,t)$ satisfies the wave equation (6.3.8) in the interior of each sector and is continuous across the characteristics $x=\pm ct$. In sector I, $\partial u/\partial x=0$ and $\partial u/\partial t=1$. In sector II, $\partial u/\partial x=1/2c$ and $\partial u/\partial t=1/2$, while in sector III, $\partial u/\partial x=\partial u/\partial t=0$. Thus the derivatives of $u(x,t)$ have jump discontinuities across the characteristics $x=\pm ct$. Therefore, $u(x,t)$ is not a strict solution of (6.3.8) since it lacks the required number of derivatives.

To show that (6.3.11) satisfies the integral form (6.1.4) of the wave equation, we show that u_1, u_2, and u_3 (the values of u in the three sectors) satisfy the one-dimensional form of the jump condition (6.3.3) across the characteristics $x=\pm ct$. We have already seen that $u(x,t)$ is a solution of the wave equation (6.3.8) away from

these characteristics. For the characteristic $x = ct$, we have $\phi(x,t) = x - ct = 0$, so that $\phi_x = 1$ and $\phi_t = -c$. Also, $\rho = 1$ and $p = c^2$, so that (6.3.3) becomes

$$c^2[u_x] + c[u_t] = c^2\left\{\frac{\partial u_2}{\partial x} - \frac{\partial u_1}{\partial x}\right\} + c\left\{\frac{\partial u_2}{\partial t} - \frac{\partial u_1}{\partial t}\right\} = \frac{1}{2}c - \frac{1}{2}c = 0. \quad (6.3.12)$$

On the characteristic $x = -ct$, we have $\phi = x + ct = 0$ with $\phi_x = 1$ and $\phi_t = c$. Thus (6.3.3) becomes

$$c^2[u_x] - c[u_t] = c^2\left\{\frac{\partial u_3}{\partial x} - \frac{\partial u_2}{\partial x}\right\} - c\left\{\frac{\partial u_3}{\partial t} - \frac{\partial u_2}{\partial t}\right\} = -\frac{1}{2}c + \frac{1}{2}c = 0. \quad (6.3.13)$$

We have demonstrated that $u(x,t)$ is a twice-differentiable solution of the wave equation (6.3.8) away from the characteristics and that it satisfies the jump condition (6.3.3) across the characteristics $x = \pm ct$. Thus (6.3.11) satisfies the integral form (6.1.4) of the wave equation.

Exercises 6.3

6.3.1. Show that for the one-dimensional wave equation (6.3.8) the jump conditions (6.3.3)–(6.3.4) reduce to the single condition $c[u_x] \pm [u_t] = 0$ across the characteristics $\phi(x,t) = x \mp ct$ = constant.

6.3.2. Determine that the function $u(x,t) = \begin{cases} f(x-ct), & x - ct < 0, \\ 0 & x - ct > 0, \end{cases}$ with $f(0) = 0$ and $f'(0) \neq 0$, satisfies the jump condition across the characteristic $\phi(x,t) = x - ct = 0$ for the wave equation given in Exercise 6.3.1.

6.3.3. Show that $u(x,t) = \begin{cases} e^{-t}f(x-t), & x - t < 0, \\ 0 & x - t > 0, \end{cases}$ where $f(x)$ is a smooth function with $f(0) = 0$ and $f'(0) \neq 0$, is a solution of the hyperbolic equation $u_{tt}(x,t) - u_{xx}(x,t) + 2u_t(x,t) + u(x,t) = 0$, with discontinuous first derivatives across $x - t = 0$. Derive the jump conditions for this equation and verify that the jump condition across the characteristic $x - t = 0$ is satisfied for $u(x,t)$.

6.3.4. Determine that the function $u(x,y,t) = \begin{cases} \sin[\phi(x,y,t)], & \phi(x,y,t) < 0, \\ 0 & \phi(x,y,t) > 0, \end{cases}$ where $\phi(x,y,t) = x\cos(\theta) + y\sin(\theta) - ct$ with a constant θ, is a solution of the two-dimensional wave equation $u_{tt}(x,y,t) - c^2[u_{xx}(x,y,t) + u_{yy}(x,y,t)] = 0$, whose derivatives are discontinuous across the characteristic $\phi(x,y,t) = 0$. Show that the jump conditions at $\phi(x,y,t) = 0$ are satisfied.

6.3.5. Show that $u(x,y,z,t) = \begin{cases} \sinh(r-ct)/r, & r - ct < 0, \\ 0 & r - ct > 0, \end{cases}$ where $r^2 = x^2 + y^2 + z^2$, satisfies the three-dimensional wave equation $u_{tt}(x,y,z,t) - \nabla^2 u(x,y,z,t) = 0$, and that its first derivatives satisfy the jump conditions across the characteristic $r - ct = 0$.

6.3.6. Show that if we have $\phi_t = 0$ in (6.3.5), we conclude from (6.3.4) that $[u_t] = 0$ across the surface $\phi(\mathbf{x})$ = constant, and find from (6.3.3) that the normal derivative of $u(\mathbf{x}, t)$ must be continuous across $\phi(\mathbf{x})$ = constant. Therefore, both $u(\mathbf{x}, t)$ and its first derivatives must be continuous across $\phi(\mathbf{x})$ = constant.

6.3.7. Obtain (6.3.7) from the integral relation (6.1.8) in the parabolic case.

6.3.8. Show that $\phi_t = 0$ in (6.3.7) does not yield a surface of discontinuity for the first derivatives of solutions of the corresponding parabolic equation.

6.3.9. Show that the solution of the initial and boundary value problem for the heat equation, $u_t(x,t) - c^2 u_{xx}(x,t)$, $0 < x < \infty$, $t > 0$, $u(x,0) = 0$, $x > 0$, $u(0,t) = \begin{cases} 0, & 0 < t < t_0, \\ 1 & t_0 < t, \end{cases}$ is given in the following form

$$u(x,t) = \begin{cases} 0, & 0 < x < \infty,\ 0 < t < t_0, \\ \operatorname{erfc}(x/2c\sqrt{t - t_0}), & 0 < x < \infty,\ t_0 < t, \end{cases}$$

where $\operatorname{erfc}(z)$ is the complementary error function defined in (5.3.33). Verify that even though $u(0,t)$ has a discontinuity at $t = t_0$, $u(x,t)$, $u_x(x,t)$, and $u_t(x,t)$ have a zero jump across $t = t_0$ for $x > 0$. Note that $t = t_0$ is a characteristic for the heat equation.

6.4 WEAK SOLUTIONS

The results of the preceding section have extended the concept of the solution of second order partial differential equations to the case of discontinuous first derivatives. Had we introduced initial displacements $u(x,0)$ with jump discontinuities in Example 6.3, a formal application of d'Alembert's solution would have shown that the solution has discontinuities across the characteristics $x = \pm ct$ for the wave equation. Such *solutions* cannot be discussed on the basis of the foregoing methods where the continuity of solutions was assumed. To deal with such problems, we weaken the concept of solution even further and obtain a new integral expression for each of the hyperbolic, parabolic, and elliptic equations considered. Again, this expression is equivalent to the given differential equation when the solutions are smooth. However, it remains valid even if the solution has jump discontinuities.

Initial and Boundary Value Problems for Hyperbolic Equations

We discuss the *hyperbolic equation* (6.1.1) in detail and then state the results for the *parabolic* and *elliptic equations* of Section 4.1. Given the region R, we consider the smooth function $v(\mathbf{x}, t)$ that is assumed to vanish identically on and near the boundary ∂R of R, if R is of finite extent. If R is of infinite extent, we assume that $v(\mathbf{x}, t)$ vanishes outside some bounded region, as well as on and near any finite boundary of R. We define the operator $\tilde{L}$ as

$$\tilde{L}[u] = \rho u_{tt} - \nabla \cdot (p \nabla u) + qu, \tag{6.4.1}$$

where we are again considering two- or three-dimensional spatial regions. Then we obtain, since the operator $\tilde{L}$ is formally self-adjoint (see Section 3.6),

$$v\tilde{L}[u] - u\tilde{L}[v] = \rho[vu_{tt} - uv_{tt}] - [v\nabla \cdot (p\nabla u) - u\nabla \cdot (p\nabla v)]$$

$$= \frac{\partial}{\partial t}[\rho v u_t - \rho u v_t] - \nabla \cdot [pv\nabla u - pu\nabla v], \tag{6.4.2}$$

since $\rho = \rho(\mathbf{x})$. Using the space-time gradient vector $\tilde{\nabla} = [\nabla, \partial/\partial t]$ we have $v\tilde{L}[u] - u\tilde{L}[v] = -\tilde{\nabla} \cdot [pv\nabla u - pu\nabla v,\ -\rho v u_t + \rho u v_t]$. Integrating over the region R, we obtain

$$\iint_R \left\{ v(\tilde{L}[u] - \rho F) - (u\tilde{L}[v] - v\rho F) \right\} dV$$

$$= -\iint_R \tilde{\nabla} \cdot [pv\nabla u - pu\nabla v,\ -\rho v u_t + \rho u v_t]\, dV$$

$$= -\int_{\partial R} [pv\nabla u - pu\nabla v,\ -\rho v u_t + \rho u v_t] \cdot \mathbf{n}\, ds = 0 \tag{6.4.3}$$

where $\mathbf{n}$ is the exterior unit normal to ∂R. The last integral in (6.4.3) results on using the divergence theorem, and it equals zero since v vanishes identically near ∂R. This is so even if R is unbounded, since $v(\mathbf{x}, t)$ vanishes for sufficiently large $\mathbf{x}$ and t. The divergence theorem is then applied to a region bounded by ∂R and by portions of circles or spheres whose radius is subsequently allowed to tend to infinity.

We conclude from (6.4.3) that

$$\iint_R v(\tilde{L}[u] - \rho F)\, dV = \iint_R (u\tilde{L}[v] - v\rho F))\, dV. \tag{6.4.4}$$

If $u(\mathbf{x}, t)$ is a smooth solution of (6.1.1), the integral on the left in (6.4.4) vanishes and we conclude that for any admissible $v(\mathbf{x}, t)$ we have

$$\iint_R (u\tilde{L}[v] - v\rho F))\, dV = 0. \tag{6.4.5}$$

Conversely, if $u(\mathbf{x}, t)$ satisfies the integral expression (6.4.5) for all admissible $v(\mathbf{x}, t)$ and $u(\mathbf{x}, t)$ is twice continuously differentiable, we obtain from the equation (6.4.4)

$$\iint_R v(\tilde{L}[u] - \rho F)\, dV = 0. \tag{6.4.6}$$

Using the *fundamental lemma of the calculus of variations* (see Exercise 8.1.8), we conclude from the arbitrariness of $v(\mathbf{x}, t)$ that $\tilde{L}[u] - \rho F = 0$, so that $u(\mathbf{x}, t)$ satisfies (6.1.1).

If $u(\mathbf{x}, t)$ is a twice continuously differentiable solution of the hyperbolic equation (6.1.1), we say that $u(\mathbf{x}, t)$ is a *classical solution* of the differential equation. If $u(\mathbf{x}, t)$ is a solution of (6.4.5) for all admissible $v(\mathbf{x}, t)$ but does not have the required number of derivatives to be a classical solution of the differential equation (6.1.1), we say that $u(\mathbf{x}, t)$ is a *weak* or *generalized solution* of the differential equation. Note that the solutions of the integral relation (6.1.4) may also be termed weak solutions if they are not twice differentiable.

Initial Value Problems for Hyperbolic Equations

Weak solutions of initial value problems for the *hyperbolic equation* (6.1.1) are defined as follows. We proceed as before but assume that a portion of ∂R, say ∂R_0, lies on the initial region $t = 0$, where $u(\mathbf{x}, 0) = f(\mathbf{x})$ and $u_t(\mathbf{x}, 0) = g(\mathbf{x})$ are specified. The function $v(\mathbf{x}, t)$ need not vanish on ∂R_0 but it must vanish on and near the remainder of ∂R. As a result, the integral over ∂R in (6.4.3) is nonzero when taken over the subset ∂R_0 of ∂R. Then, if

$$\iint_R (u\tilde{L}[v] - v\rho F))\, dV + \int_{\partial R_0} (\rho f v_t - \rho g v)\, ds = 0 \tag{6.4.7}$$

for all admissible $v(\mathbf{x}, t)$, $u(\mathbf{x}, t)$ is said to be a *weak solution* of the aforementioned initial value problem for (6.1.1).

It can again be shown that if $u(\mathbf{x}, t)$ is a classical solution of (6.1.1) in two subregions R_1 and R_2 of R and has a jump discontinuity across the surface S_0 that separates R_1 and R_2, then S_0 must be a characteristic surface. This is done by applying (6.4.4) to the two regions R_1 and R_2 and considering the nonzero contributions (i.e., the integral over S_0). We represent S_0 by $\phi(\mathbf{x}, t) = 0$. The integral is then given as

$$\int_{S_0} \{v\,[[\nabla u], [u_t]] \cdot [p\nabla\phi,\ -\rho\phi_t] - [u][\nabla v, v_t] \cdot [p\nabla\phi,\ -\rho\phi_t]\}\ ds = 0, \tag{6.4.8}$$

where $[V]$ denotes the jump in V across S_0. Since v is arbitrary we can put $v = 0$ on S_0. We then conclude on the basis of the duBois-Reymond lemma, as in Section 6.3, that $[u][\nabla v, v_t] \cdot [p\nabla\phi,\ -\rho\phi_t] = 0$. Now if the vector $[p\nabla\phi, -\rho\phi_t]$ is not tangential to the surface S_0, then $[\nabla v, v_t] \cdot [p\nabla\phi, -\rho\phi_t]$ is effectively a transverse (i.e., nontangential) derivative of v on S_0 and cannot be assumed to vanish throughout S_0 even if v is zero there. As a result, we must put $[u] = 0$ on S_0, and this contradicts our assumption that u has a nonzero jump across S_0. Therefore, the vector must be tangential and its dot product with normal vector $[\nabla\phi, \phi_t]$ to S_0 must vanish. This yields the characteristic equation $\rho\phi_t^2 - p(\nabla\phi)^2 = 0$, so that S_0 is a *characteristic surface*.

Weak Solutions of Parabolic and Elliptic Equations

We now define *weak* or *generalized solutions* for the *parabolic* and *elliptic equations* of Section 4.1. We say that $u(\mathbf{x}, t)$ is a *classical solution* of the *parabolic equation*

$$\hat{L}[u] = \rho\, u_t - \nabla \cdot (p\nabla u) + qu = \rho F \tag{6.4.9}$$

if $u(\mathbf{x}, t)$ has a continuous time derivative and two continuous derivatives in the spatial variables. With the region R and the smooth function $v(\mathbf{x}, t)$ defined as before, we easily conclude by adapting the foregoing procedure for classical solutions $u(\mathbf{x}, t)$ that

$$\iint_R [u\{-\rho v_t - \nabla \cdot (p\nabla v) + qv\} - v\rho F]\, dV = 0. \tag{6.4.10}$$

The operator $\hat{L}^*[v] = -\rho v_t - \nabla \cdot (p\nabla v) + qv$ in the braces is the formal adjoint of the operator $\hat{L}[u]$ of (6.4.9) (see Section 3.6). If $u(\mathbf{x}, t)$ satisfies (6.4.9) for all admissible $v(\mathbf{x}, t)$ and is not a classical solution of (6.4.9), we say that it is a *weak* or *generalized solution* of the differential equation (6.4.9). Weak solutions of initial value problems for (6.4.9) are considered in the exercises.

Similarly, if $u(\mathbf{x})$ is a twice continuously differentiable solution of the *elliptic equation*

$$L[u] = -\nabla \cdot (p\nabla u) + qu = \rho F, \tag{6.4.11}$$

it is a *classical solution* of (6.4.11), and it is easily shown to satisfy the integral relation

$$\iint_R [u\{\nabla \cdot (p\nabla v) - qv\} + v\rho F]\, dV = 0 \tag{6.4.12}$$

for all admissible $v(\mathbf{x})$, since L is a formally self-adjoint operator. If $u(\mathbf{x})$ satisfies (6.4.12) but is not a classical solution, it is said to be a *weak* or *generalized solution* of (6.4.11). In view of the smoothness properties of solutions of elliptic equations, there is generally no distinction between weak and classical solutions in the elliptic case.

Examples

We now consider some examples. In the first example we consider a weak solution of an initial value problem for the wave equation (6.3.8). In the second example we examine the Fourier series solutions obtained in Chapter 4 and show how they may be interpreted as weak solutions.

Example 6.4. Weak Solutions of the Wave Equation. The Cauchy problem for the wave equation

$$u_{tt}(x,t) = c^2 u_{xx}(x,t),\ |x| < \infty,\ t > 0, \quad u(x,0) = \begin{cases} 0, & x < 0, \\ 1, & x \geq 0, \end{cases} \quad u_t(x,0) = 0, \tag{6.4.13}$$

has the formal *d'Alembert's solution* for $t \geq 0$

$$u(x,t) = \begin{cases} 0, & x < -ct, \\ 1/2, & -ct < x < ct, \\ 1, & x > ct. \end{cases} \tag{6.4.14}$$

Since $u(x,t)$ is discontinuous across the characteristics $x = \pm ct$, we must interpret the solution in a generalized sense.

We consider the rectangular region R displayed in Figure 6.6 and assume that the smooth function $v(x,t)$ vanishes near the boundary ∂R of R except for the part of ∂R that coincides with the x-axis. The region R is divided into sectors I, II, and III, and we show that (6.4.14) satisfies (6.4.7) for all admissible $v(x,t)$ so that it is a *weak solution* of (6.4.13). We have

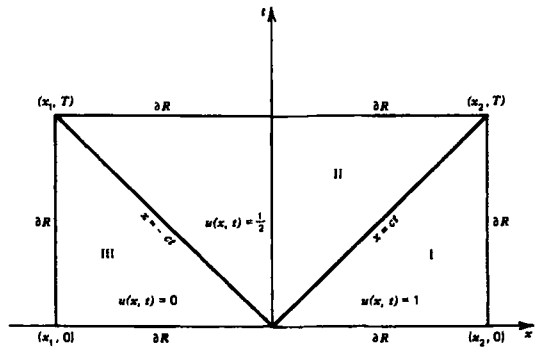

Figure 6.6 The region R.

$$\iint_R u[v_{tt} - c^2 v_{xx}]\, dV = \iint_{\mathrm{I}} 1[v_{tt} - c^2 v_{xx}]\, dx\, dt$$

$$+ \iint_{\mathrm{II}} \frac{1}{2}[v_{tt} - c^2 v_{xx}]\, dx\, dt + \iint_{\mathrm{III}} 0[v_{tt} - c^2 v_{xx}]\, dx\, dt. \tag{6.4.15}$$

For the integral over the region A, where A represents I or II, we have

$$\iint_A [v_{tt} - c^2 v_{xx}]\, dx\, dt = -\iint_A \left[\frac{\partial}{\partial x}, \frac{\partial}{\partial t}\right] \cdot [c^2 v_x, -v_t]\, dx\, dt.$$

$$= -\int_{\partial A} [c^2 v_x, -v_t] \cdot \mathbf{n}\, ds = -\int_{\partial A} v_t\, dx + c^2 v_x\, dt, \tag{6.4.16}$$

on using the divergence theorem and the fact that $\mathbf{n}\, ds = [dt, -dx]$. The integration in the line integral over ∂A is taken in the positive direction.

Since $v(x,t)$ vanishes on ∂R if $t > 0$, and on $x \mp ct = 0$ we have $v_t\, dx + c^2 v_x\, dt = \pm c\, dv$, we obtain

$$\iint_R u[v_{tt} - c^2 v_{xx}]\, dx\, dt = -\int_0^{x_2} v_t\, dx - \int_{(x_2,T)}^{(0,0)} c\, dv + \int_{(x_1,T)}^{(0,0)} \frac{c}{2}\, dv$$

$$-\int_{(0,0)}^{(x_2,T)} \frac{c}{2}\, dv = -\int_0^{x_2} v_t\, dx - cv(0,0) + \frac{c}{2}v(0,0) + \frac{c}{2}v(0,0) = -\int_0^{x_2} v_t\, dx. \tag{6.4.17}$$

Since T is arbitrary, we have shown that $u(x,t)$ satisfies (6.4.7) (appropriately modified for this example) for all x and for $t > 0$, and it is therefore a *weak solution* of the Cauchy problem (6.4.13).

Example 6.5. Eigenfunction Expansions and Weak Solutions. The series solutions given in terms of eigenfunctions obtained in Chapter 4 by the method of separation of variables or by finite Fourier transforms, were shown to require strong conditions on the data for the given problems to render them classical solutions. This was especially the case for problems that involved hyperbolic equations. It was indicated in Chapter 4 how the series may be understood as generalized solutions in case they are not classical solutions, and we now use the preceding results to show in a more precise manner how the series solutions are to be interpreted as *weak* or *generalized solutions.*

Our discussion is restricted to the hyperbolic equation (4.1.10) with the homogeneous boundary conditions (4.2.2) and the initial conditions (4.2.4). Proceeding as in Section 4.6, we expand $u(\mathbf{x}, t)$ in a series of eigenfunctions $u(\mathbf{x}, t) = \sum_{k=1}^{\infty} N_k(t) M_k(\mathbf{x})$, with $N_k(t)$ defined as in (4.6.14) except that $B_k = 0$ for all k since the boundary conditions are homogeneous. We assume that each of the eigenfunctions $M_k(\mathbf{x})$ is twice continuously differentiable and so are the $N_k(t)$. The partial sum $u_m(\mathbf{x}, t) = \sum_{k=1}^{m} N_k(t) M_k(\mathbf{x})$, is easily seen to be a classical solution of the equation

$$\rho \frac{\partial^2 u_m}{\partial t^2} - \nabla \cdot (p \nabla u_m) + q u_m = \rho F_m, \tag{6.4.18}$$

where $F_m(\mathbf{x}, t)$ is the mth partial sum of the Fourier series of $F(\mathbf{x}, t)$. The initial and boundary conditions are $u_m(\mathbf{x}, 0) = \sum_{k=1}^{m} (f, M_k) M_k(\mathbf{x})$, $\partial u_m(\mathbf{x}, 0)/\partial t = \sum_{k=1}^{m} (g, M_k) M_k(\mathbf{x})$, $\alpha u_m + \beta \partial u_m / \partial n \big|_{\partial G} = 0$.

We consider the time interval $0 < t < T$ and the region G on which the initial data are assigned together with its boundary ∂G. Define the region by R in $(\mathbf{x}, t)$-space as the direct product $G \times [0, T]$. Let $v(\mathbf{x}, t)$ be a smooth function that vanishes on and near ∂R. Then, since each u_m is a classical solution of (6.4.18), we have from (6.4.5)

$$\iint_R u_m \{\rho v_{tt} - \nabla \cdot (p \nabla v) + q v\} \, dV = \iint_R \rho F_m v \, dV. \tag{6.4.19}$$

We now assume that as $m \to \infty$ we have $u_m(\mathbf{x}, 0) \to f(\mathbf{x})$, $\partial u_m(\mathbf{x}, 0)/\partial t \to g(\mathbf{x})$, $\nabla u_m(\mathbf{x}, 0) \to \nabla f(\mathbf{x})$, and $F_m(\mathbf{x}, t) \to F(\mathbf{x}, t)$, where f and g are the initial data and ρF is the inhomogeneous term of the hyperbolic equation. The convergence of $u_m(\mathbf{x}, 0)$ to $f(\mathbf{x})$ is assumed to be uniform, while all other cases represent mean square convergence. It can then be shown that the partial sums $u_m(\mathbf{x}, t)$ converge in the mean square sense to a function $u(\mathbf{x}, t)$ in the region R. The function $u(\mathbf{x}, t)$ need not be a solution in the classical sense.

To proceed, we consider the integral

$$\iint_R (u - u_m) \{\rho v_{tt} - \nabla \cdot (p \nabla v) + q v\} - \rho (F - F_m) v \, dV. \tag{6.4.20}$$

Let $\tilde{L}[v]$ represent the expression in the braces in (6.4.20), and apply the Cauchy-Schwarz inequality (see Exercise 4.2.12) to the absolute value of the integral. This

gives us

$$\left| \iint_R (u - u_m)\tilde{L}[v] - \rho(F - F_m)v \, dV \right| \leq ||u - u_m|| \, ||\tilde{L}[v]/\rho|| + ||F - F_m|| \, ||v||. \tag{6.4.21}$$

On proceeding to the limit as $m \to \infty$ in (6.4.21) we conclude that the right side tends to zero, since $\lim_{m\to\infty} ||u - u_m|| = 0$ and $\lim_{m\to\infty} ||F - F_m|| = 0$ by assumption. Consequently,

$$\iint_R u\{\rho v_{tt} - \nabla \cdot (p\nabla v) + qv\} \, dV = \iint_R \rho F v \, dV, \tag{6.4.22}$$

and we have shown that $u(\mathbf{x}, t)$ is a *weak solution* of 1.10). [The manner in which we have shown that $u_m(\mathbf{x}, t)$ converges to a solution $u(\mathbf{x}, t)$ is often called *weak convergence*.] On using (6.4.7) instead of (6.4.5) to define a weak solution, the foregoing analysis shows that $u(\mathbf{x}, t)$ is a weak solution of the given initial and boundary value problem.

We have not taken into account the boundary conditions assigned for the problems considered in this section because $v(\mathbf{x}, t)$ was chosen to vanish on the boundary. An alternative weak formulation of initial and boundary value problems for more general forms of (second order) equations than those considered in this section is given in the exercises. Boundary terms do play a role in the resulting weak solutions. This formulation is used in the *Galerkin method* for obtaining approximate solutions of the given initial and boundary value problems (see Section 8.2 and Chapter 12).

Exercises 6.4

6.4.1. Show that $u(x, y, t) = \begin{cases} 0, & x - ct < 0, \\ 1, & x - ct > 0, \end{cases}$ is a weak solution of the two-dimensional wave equation $u_{tt}(x, y, t) - c^2[u_{xx}(x, y, t) + u_{yy}(x, y, t)] = 0$. *Hint*: Use (6.4.3) and apply it to regions on both sides of the characteristic surface $\phi(x, y, t) = x - ct = 0$. Show that the boundary integral over the surface $\phi(x, y, t) = 0$, which results from the integration over the region where $\phi(x, y, t) > 0$, vanishes by showing that the terms in v can be expressed as a directional derivative.

6.4.2. Verify that $u(x, t) = \begin{cases} \cos(x + ct), & x + ct < 0, \\ \sin(x + ct), & x + ct > 0, \end{cases}$ is a weak solution of the one-dimensional wave equation.

6.4.3. Obtain the integral (6.4.9) that characterizes weak solutions for the parabolic equation (6.4.8).

6.4.4. Obtain an integral relation [similar to (6.4.7)] that characterizes weak solutions of initial value problems for the parabolic equation (6.4.8).

6.4.5. Using the convergence properties of Fourier sine series discussed in Chapter 4, set up an initial and boundary value problem for the homogeneous wave equation in the finite interval $0 < x < l$ with $u(0,t) = u(l,t) = 0$ such that its separation of variables result is a weak solution and not a classical solution. That is, choose the initial data $f(x)$ and $g(x)$ such that the conditions given in Example 6.5 are met, but the series solution is not twice differentiable term by term.

6.4.6. Solve the initial and boundary value problem for the heat equation given in Exercise 6.3.9, with the initial condition $u(x,0) = 1$ and the boundary condition $u(0,t) = 1$ for $0 < t < t_0$ and $u(0,t) = 0$ for $t > t_0$. Show that the solution is continuous across $t = t_0$.

6.4.7. Use separation of variables to show that a formal solution of the steady-state problem for the heat equation $u_t(x,t) - c^2 u_{xx}(x,t) = 0$, $-\infty < x < \infty$, $t > 0$, $u(0,t) = \begin{cases} \sin(\omega t), & t > 0, \\ 0 & t < 0, \end{cases}$ where ω is a constant, is given as $u(x,t) = \begin{cases} \exp(-\sqrt{\omega/2c^2}\, x)\sin(-\sqrt{\omega/2c^2}\, x + \omega t), & t > 0, \\ 0, & t < 0, \end{cases}$ if we require that $u(x,t) \to 0$ as $x \to \infty$. Determine that $u(x,t)$ is not a weak solution of the heat equation. Show that (5.3.31) with $\hat{g} = u(0,t)$ (as defined in this exercise) is a weak solution of the heat equation.

6.4.8. Consider the (nonself-adjoint) elliptic equation $-\nabla\cdot(p(\mathbf{x})\nabla u(\mathbf{x})) + q(\mathbf{x})u(\mathbf{x}) + \mathbf{b}(\mathbf{x})\cdot\nabla u(\mathbf{x}) = \rho(\mathbf{x})F(\mathbf{x})$ in the bounded region G, with one of the (four) boundary conditions given in Section 4.1 assigned on ∂G. Multiply across by a function $v(\mathbf{x})$ and integrate the equation over the region G. Use Green's theorem to obtain the integral relation

$$\iint_G [p\nabla u\cdot\nabla v + quv + v\mathbf{b}\cdot\nabla u]\, dV = \iint_G \rho v F\, dV + \int_{\partial G} pv\frac{\partial u}{\partial n}\, ds.$$

Determine the form of the integral relation for each of the four boundary conditions given in Section 4.1. If $u(\mathbf{x})$ is specified on a part of the boundary, put $v(\mathbf{x}) = 0$ there. For each case this yields an integral relation of the Galerkin form associated with the boundary value problem for the elliptic equation. The function $u(\mathbf{x})$ is said to be a weak solution of the given boundary value problem, in the Galerkin sense, if it satisfies the corresponding integral relation for all admissible $v(\mathbf{x})$. Only in the case of Dirichlet data on all or part of the boundary must $u(\mathbf{x})$ be chosen to satisfy the boundary condition and $v(\mathbf{x})$ be equated to zero there. Otherwise, $u(\mathbf{x})$ and $v(\mathbf{x})$ are unrestricted. One of the forms of the Galerkin method for obtaining approximate solutions of boundary value problems is based on the foregoing results. An integral relation of Galerkin form can also be associated with initial and boundary value problems for hyperbolic and parabolic equations. (See Chapter 12 for a general discussion.)

6.5 THE INTEGRAL WAVE EQUATION

The integral relation (6.1.4) that corresponds to the one-dimensional *wave equation* (6.4.12) leads to a number of interesting expressions that can be used to solve various initial and boundary value problems for the wave equation as we now demonstrate. If $u(x,t)$ is a solution of the one-dimensional wave equation (6.4.12), then (6.1.4) takes the form of a line integral and we obtain

$$\int_{\partial R} [c^2 u_x,\ -u_t] \cdot \mathbf{n}\, ds = \int_{\partial R} u_t\, dx + c^2 u_x\, dt = 0, \tag{6.5.1}$$

since $\mathbf{n}\, ds = [dt, -dx]$. When the wave equation (6.4.12) has a nonhomogeneous term F, this term must be integrated over the region R. We call (6.5.1) the *integral wave equation*, and now choose a variety of regions R and integrate (6.5.1) over their boundaries.

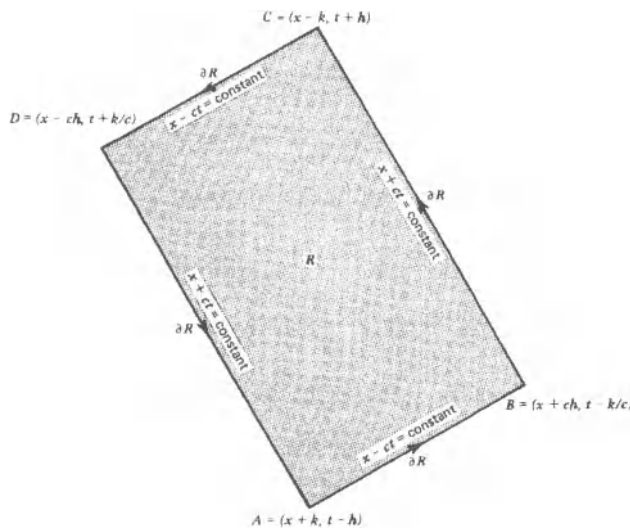

Figure 6.7 The characteristic quadrilateral.

Characteristic Quadrilaterals and Triangles

First we select R be a region in (x,t)-space bounded on four sides by characteristic lines $x \pm ct$ = constant as pictured in Figure 6.7. As a result, ∂R is a *characteristic quadrilateral*. On the lines $x = \pm ct$ = constant we have $dx = \mp c\, dt$ and $u_t\, dx + c^2 u_x\, dt = \mp c\, du$. Then the integral (6.5.1) takes the form

$$\int_{\partial R} u_t\,dx + c^2 u_x\,dt = \int_A^B c\,du - \int_B^C c\,du + \int_C^D c\,du - \int_D^A c\,du$$
$$= 2c\,[u(B) + u(D) - u(A) - u(C)] = 0, \qquad (6.5.2)$$

with $u(C)$, for example, given as $u(C) = u(x-k, t+h)$, where k and h are prescribed positive constants. This gives the difference equation

$$u(x-k, t+h) + u(x+k, t-h) = u\left(x + ch, t - \frac{k}{c}\right) + u\left(x - ch, t + \frac{k}{c}\right). \qquad (6.5.3)$$

Using Taylor's expansion it can be shown that any twice continuously differentiable solution of the difference equation must satisfy the wave equation, in view of the arbitrariness of h and k (see Exercises 6.5.1 and 6.5.2).

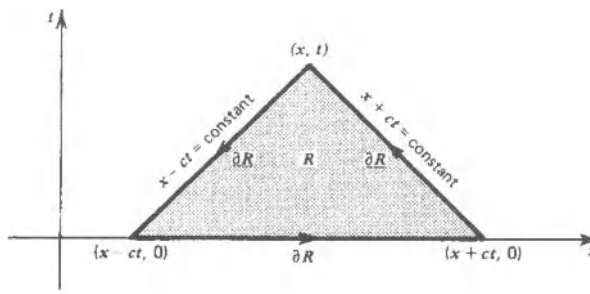

Figure 6.8 The characteristic triangle.

As our next choice for the region R, we choose a *characteristic triangle* with base on the x-axis, whose two other sides are characteristic line segments, as shown in Figure 6.8. Then ∂R is the boundary of the characteristic triangle, and using the foregoing results, we obtain

$$\int_{\partial R} u_t\,dx + c^2 u_x\,dt = -\int_{(x+ct,0)}^{(x,t)} c\,du + \int_{(x,t)}^{(x-ct,0)} c\,du - \int_{(x-ct,0)}^{(x+ct,0)} u_t(x,0)\,dx$$
$$= -2cu(x,t) + cu(x+ct,0) + cu(x-ct,0) + \int_{x-ct}^{x+ct} u_t(x,0)\,dx = 0. \qquad (6.5.4)$$

Given the Cauchy problem for the wave equation (6.4.12) with initial data $u(x,0) = f(x)$, $u_t(x,0) = g(x)$, we obtain from (6.5.4),

$$u(x,t) = \frac{1}{2}[f(x-ct) + f(x+ct)] + \frac{1}{2c}\int_{x-ct}^{x+ct} g(x)\,dx, \qquad (6.5.5)$$

which is just *d'Alembert's solution*. For the inhomogeneous wave equation with a term F on the right of (6.4.12), we must add the integral of $F(x,t)$ over the characteristic triangle R to the solution (6.5.5).

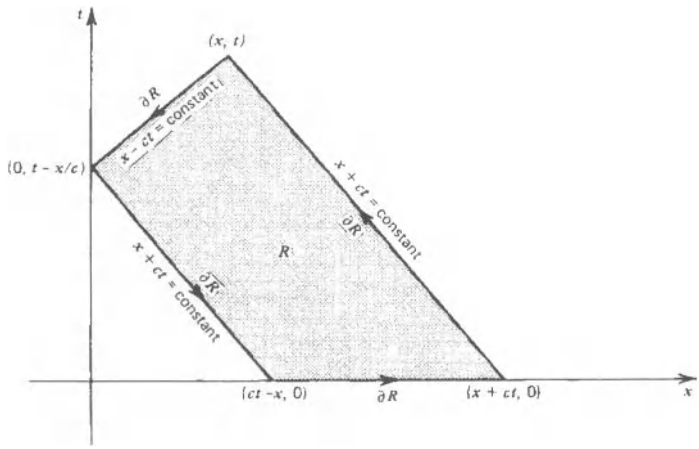

Figure 6.9 The quadrilateral.

A third choice for the region of integration R in (6.5.1) is the quadrilateral pictured in Figure 6.9 with the boundary ∂R, three of whose sides are characteristics and whose fourth side is an interval on the x-axis. We apply (6.5.1) over the boundary of the region R to obtain

$$\int_{\partial R} u_t\,dx + c^2 u_x\,dt = -\int_{(x+ct,0)}^{(x,t)} c\,du + \int_{(x,t)}^{(0,t-x/c)} c\,du$$

$$-\int_{(0,t-x/c)}^{(ct-x,0)} c\,du + \int_{(ct-x,0)}^{(x+ct,0)} u_t(x,0)\,dx = -2cu(x,t) + cu(x+ct,0)$$

$$+2cu(0,t-x/c) - cu(ct-x,0) + \int_{(ct-x,0)}^{(x+ct,0)} u_t(x,0)\,dx = 0. \quad (6.5.6)$$

With the initial conditions $u(x,0) = f(x)$, $u_t(x,0) = g(x)$ and the boundary condition $u(0,t) = h(t)$, we obtain

$$u(x,t) = \frac{1}{2}[f(x-ct) - f(ct-x)] + \frac{1}{2c}\int_{ct-x}^{ct+x} g(x)\,dx + h\left(t - \frac{x}{c}\right). \quad (6.5.7)$$

Differentiating (6.5.7) with respect to x and evaluating the result at $x = 0$ gives

$$u_x(0,t) = f'(ct) + \frac{1}{c}g(ct) - \frac{1}{c}h'(t) = u_x(ct,0) + \frac{1}{c}u_t(ct,0) - \frac{1}{c}u_t(0,t). \quad (6.5.8)$$

Using these results, any initial and boundary value problem for the wave equation over a finite, semi-infinite, or infinite interval can be solved and domains of dependence for the solutions can be established, as we now show.

Examples

Example 6.6. The Wave Equation in a Semi-Infinite Interval. We consider the initial and boundary value problem for the inhomogeneous wave equation in a semi-infinite interval; that is,

$$u_{tt}(x,t) - c^2 u_{xx}(x,t) = F(x,t), \qquad 0 < x < \infty,\ t > 0, \tag{6.5.9}$$

with initial and boundary conditions

$$u(x,0) = f(x),\ u_t(x,0) = g(x),\ x > 0,\ \ \alpha\, u(0,t) - \beta\, u_x(0,t) = B(t),\ t > 0, \tag{6.5.10}$$

where $\alpha \geq 0$, $\beta \geq 0$ and $\alpha + \beta > 0$ (with constant α and β) and F, f, g, and B are given functions.

To solve this problem we break up the first quadrant in the (x,t)-plane into two regions. In region I, $x > ct$, and in region II, $x < ct$. These regions are separated by the characteristic curve $x = ct$ which issues from the origin, as shown in Figure 6.10.

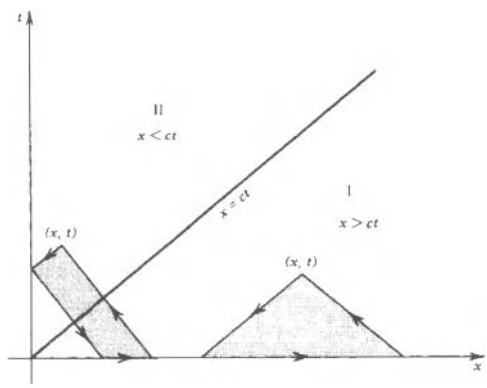

Figure 6.10 The first quadrant.

For a point (x,t) in region I, we integrate over a characteristic triangle and obtain d'Alembert's solution (6.5.5) plus an integral of $F(x,t)$ taken over the characteristic triangle,

$$u(x,t) = \text{d'Alembert's solution} + \frac{1}{2c}\int_0^t \int_{x-c(t-\tau)}^{x+c(t-\tau)} F(\sigma,\tau)\, d\sigma\, d\tau. \tag{6.5.11}$$

If the point (x,t) is in region II, we integrate over the quadrilateral pictured in Figure 6.10, three of whose sides are characteristics, and obtain from (6.5.7)

$$u(x,t) = \frac{1}{2}[f(x+ct) - f(ct-x)] + \frac{1}{2c}\int_{ct-x}^{ct+x} g(x)\,dx + h\left(t - \frac{x}{c}\right) + \frac{1}{2c}\iint_Q F(\sigma,\tau)\,d\sigma\,d\tau, \tag{6.5.12}$$

where $h(t) = u(0,t)$, which is as yet unspecified if $\beta \neq 0$ in (6.5.10) and $F(x,t)$ is integrated over the interior of the quadrilateral in Figure 6.10, which we denote by Q.

If $F = 0$, then $u_x(0,t)$ is given as in (6.5.8). Substituting (6.5.8) into the boundary condition (6.5.10) gives

$$\alpha h(t) - \beta\left[f'(ct) + \frac{1}{c}\,g(ct) - \frac{1}{c}\,h'(t)\right] = B(t), \tag{6.5.13}$$

which is a first order ordinary differential equation for $u(0,t) = h(t)$. The initial condition is $h(0) = \lim_{t\to 0} u(0,t) = u(0,0)$, assuming that $u(x,t)$ is continuous at $(0,0)$. In that case $u(0,0) = \lim_{x\to 0} u(x,0) = \lim_{x\to 0} f(x) = f(0)$.

Now if $\beta = 0$ and $\alpha = 1$ in (6.5.10), then $u(0,t) = h(t) = B(t)$ so that (6.5.12) is already completely specified. However, if $B(0) \neq f(0)$, the solution (6.5.11)–(6.5.12) is discontinuous across the characteristic $x = ct$. As the characteristic is approached from points in region I, the limit of $u(x,t)$ is $f(0) + C$ and the limit of $u(x,t)$ as $x = ct$ is approached from region II is $B(0) + C$, where the constant C is identical for both regions. We remark that even if $B(0) = f(0)$, the derivatives u_x and u_t may be discontinuous across the characteristic $x = ct$, say, if $B'(0) = \lim_{t\to 0} u_t(0,t) \neq g(0) = \lim_{x\to 0} u_x(x,0)$. In these cases, the solution of the initial and boundary value problem must be interpreted in the *generalized sense*. If $\beta \neq 0$ in (6.5.10), $u(0,t)$ is unspecified and we may put $u(0,0) = f(0)$ to make the solution continuous across the characteristic line $x = ct$. The solution of (6.5.13) with $h(0) = f(0)$ is

$$h(t) = f(ct) + \int_0^t \exp\left[\frac{c\alpha}{\beta}(\tau - t)\right]\left\{\frac{c}{\beta}B(\tau) - \frac{c\alpha}{\beta}f(c\tau) + g(c\tau)\right\}d\tau, \tag{6.5.14}$$

when $\beta \neq 0$. If $F \neq 0$, it contributes an additional term to $h(t)$.

In the special case where $f = g = F = 0$ we obtain the following solution of the initial and boundary value problem (6.5.9)–(6.5.10):

$$u(x,t) = \begin{cases} 0, & x > ct, \\ c/\beta \int_0^{t-x/c} \exp\left[(c\alpha/\beta)(\tau - t + x/c)\right] B(\tau)\,d\tau, & 0 < x < ct, \end{cases} \tag{6.5.15}$$

if $\beta \neq 0$. If $\alpha = 0$ and $\beta = 1$, we have

$$u(x,t) = \begin{cases} 0, & x > ct, \\ c\int_0^{t-x/c} B(\tau)\,d\tau, & 0 < x < ct, \end{cases} \tag{6.5.16}$$

while if $\alpha = 1$ and $\beta = 0$, we obtain

$$u(x,t) = \begin{cases} 0, & x > ct, \\ B(t - x/c), & 0 < x < ct. \end{cases} \tag{6.5.17}$$

In each of the foregoing cases, the boundary condition (6.5.10) gives rise to a wave of the general form $h(t - x/c)$ that travels to the right with speed c. For this reason the problem is often referred to as a *signaling problem.*

The domain of dependence of a point (x, t) in the first quadrant is given by the characteristic triangle in region I and by the quadrilateral in region II. This follows from the solutions (6.5.11)–(6.5.12).

Example 6.7. The Wave Equation in a Finite Interval. The initial and boundary value problem for the wave equation in a finite interval can be solved by separation of variables and by the use of finite transform methods in terms of standing waves, as was shown in Chapter 4. We now use the foregoing results to construct a solution that yields a useful small time description in terms of propagating waves (see Example 5.11). Given the interval $0 < x < l$, we consider the problem

$$u_{tt}(x,t) - c^2 u_{xx}(x,t) = F(x,t), \qquad 0 < x < l,\ t > 0, \tag{6.5.18}$$

with initial conditions

$$u(x,0) = f(x), \qquad u_t(x,0) = g(x), \qquad 0 < x < l, \tag{6.5.19}$$

and the boundary data

$$\alpha\, u(0,t) - \beta\, u_x(0,t) = B_1(t),\ \alpha\, u(l,t) + \beta\, u_x(l,t) = B_2(t), \quad t > 0, \tag{6.5.20}$$

where $\alpha \geq 0$, $\beta \geq 0$, and $\alpha + \beta > 0$ (with constant α and β) and F, f, g, and B are given functions.

To solve the problem we break up the strip $0 \leq x \leq l$, $t \geq 0$ into a collection of regions R_1, R_2, $R_3, \ldots$, as shown in Figure 6.11. The regions R_k $(k = 1, 2, \ldots)$ are bounded by portions of the initial line $t = 0$, the boundary lines $x = 0$, l, and portions of characteristic lines $x \pm ct$ = constant. The problem can be solved successively by starting with the solutions in R_1, R_2, and R_3. In R_1, $u(x,t)$ is given by d'Alembert's solution. In R_2 and R_3 the solution is obtained as in Example 6.6 from (6.5.12) or a slight modification thereof. In R_4 we find $u(x,t)$ by using a characteristic quadrilateral and (6.5.3), which gives $u(x,t)$ in terms of its (known) values in R_1, R_2, and R_3. Similarly, it is possible to construct quadrilaterals and to use (6.5.3) or (6.5.12) to determine the solution in all the regions R_k with $k > 4$, on proceeding step by step from R_5 to R_6 to R_7, and so on. The values of $u(x,t)$ on the characteristic lines that separate the regions R_k depend on the compatibility of the data at $(0,0)$ and $(1,0)$, which determines whether the solution is continuous.

The hatched strips in Figure 6.11 indicate how the solution $u(x,t)$ propagates if the initial or boundary data are concentrated in a small x or a small t interval. Two cases are depicted. In one case we assume that $u(x,0) = f(x)$ vanishes outside the interval (x_0, x_1) whereas the other data $g(x)$,$F(x,t)$,$B_1(t)$, and $B_2(t)$ all vanish. This results in two waves with sharply defined wave fronts if $f(x)$ is discontinuous at

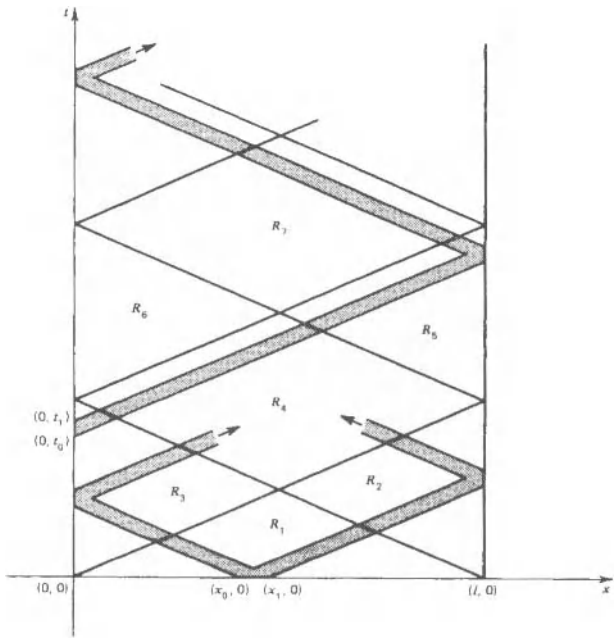

Figure 6.11 The regions R_k.

x_0 and x_1, traveling to the right and to the left with speed c. As they hit the boundaries $x = 0$ and $x = l$, they are reflected and move in the opposite direction, as shown, again with sharp wave fronts. This process continues indefinitely.

In the second case, we assume that $B_1(t) = 0$ everywhere outside the interval (t_0, t_1) and that all the other data vanish. This yields a wave traveling to the right with speed c until it reaches $x = l$. It is then reflected and reverses direction until it is again reflected from the boundary $x = 0$, and so on. In both cases, if $\beta \neq 0$ in (6.5.20), the waves do not have sharp trailing edges.

For the general case, it may be possible to detect at the initial stages of the wave motion (if the data are concentrated in small intervals or regions), propagating waves resulting from the effect of the data. However, as time increases, the interference from all waves makes it impossible to distinguish individual waves, and the representation of the solution in terms of standing waves given in Chapter 4 becomes more useful.

Example 6.8. Moving Boundaries. We consider an initial and boundary value problem for the *wave equation* in a region initially given as $x \geq 0$. As t increases from zero, the boundary point $x = 0$ begins to move according to the equation $x = h(t)$, with $h(0) = 0$. Thus, $u(x, t)$ satisfies the wave equation

$$u_{tt}(x,t) - c^2 u_{xx}(x,t) = 0, \qquad h(t) < x < \infty, \; t > 0, \tag{6.5.21}$$

with the initial and boundary conditions

$$u(x,0) = f(x),\ u_t(x,0) = g(x),\ x > 0, \quad u(h(t),t) = B(t),\ t > 0, \quad (6.5.22)$$

with prescribed $f(x)$, $g(x)$, and $B(t)$.

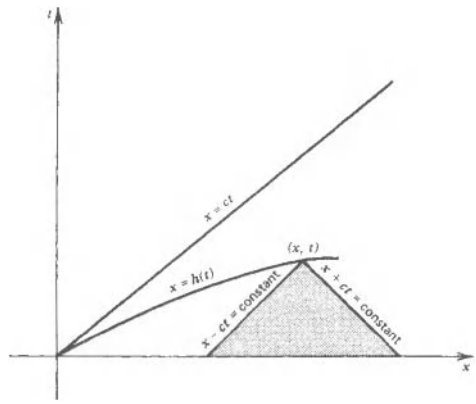

Figure 6.12 The case when $h'(t) > c$.

If the speed of the moving boundary $|dx/dt| = |h'(t)|$ exceeds the speed c of wave propagation for (6.5.21), the problem (6.5.21)–(6.5.22) is *not well posed.* This can be seen by examining Figure 6.12, where the case $h'(t) > 0$ is considered. (We concentrate our discussion on this case.) The solution $u(x,t)$ at a point $(h(t),t)$ on the boundary curve is uniquely determined (by means of d'Alembert's solution) from the initial data $f(x)$ and $g(x)$ alone. Unless the boundary condition $u(h(t),t) = B(t)$ is such that the values of u equal those resulting from the initial data, a solution does not exist. Consequently, the problem is not well posed in this case. [If $h'(t) < -c$, the boundary condition (6.5.22) does not determine the solution $u(x,t)$ uniquely, as will be shown in the exercises. As a result, the problem is again not well posed.]

When the boundary speed $|dx/dt| = |h'(t)|$ is less than the speed c of wave propagation, the problem (6.5.21)–(6.5.22) is *well posed* and the solution can be obtained in the manner indicated in Figure 6.13, where the case $h'(t) > 0$ is considered. [However, the result is also valid for $h'(t) < 0$.] We consider the regions R_1 and R_2 separated by the characteristic line $x = ct$. At a point (x,t) in R_1, d'Alembert's solution is valid and it expresses $u(x,t)$ in terms of the initial data. At a point (x,t) in R_2, we use the quadrilateral displayed in Figure 6.13 to obtain the result,

$$u(x,t) = u(\sigma,\tau) + \frac{1}{2}[f(x+ct) - f(\sigma+c\tau)] + \frac{1}{2c}\int_{\sigma+c\tau}^{x+ct} g(s)\,ds. \quad (6.5.23)$$

The point (σ,τ) is where the characteristic $x - ct$ = constant through (x,t) intersects the curve $x = h(t)$. It is determined from the equations $\sigma - c\tau = x - ct$, $h(\tau) - \sigma = 0$.

It follows that $x = h(\tau) + c(t - \tau)$, so that τ is a function of $x - ct$. Thus, $u(\sigma, \tau) = B(\tau)$ is a right-traveling wave. Once (σ, τ) is determined, we may express $u(x, t)$ in R_2 as

$$u(x,t) = B(\tau) + \frac{1}{2}[f(x + ct) - f(\sigma + c\tau)] + \frac{1}{2c}\int_{\sigma+c\tau}^{x+ct} g(s)\, ds. \qquad (6.5.24)$$

As $(x, t) \to (0, 0)$, we also have $(\sigma, \tau) \to (0, 0)$, so that with $\lim_{t\to 0} u(h(t), t) = \lim_{t\to 0} B(t) = B(0)$, we find that unless $\lim_{x\to 0} u(x, 0) = \lim_{x\to 0} f(x) = f(0) = B(0)$, the solution is discontinuous across the characteristic $x = ct$ and must be interpreted in a generalized sense.

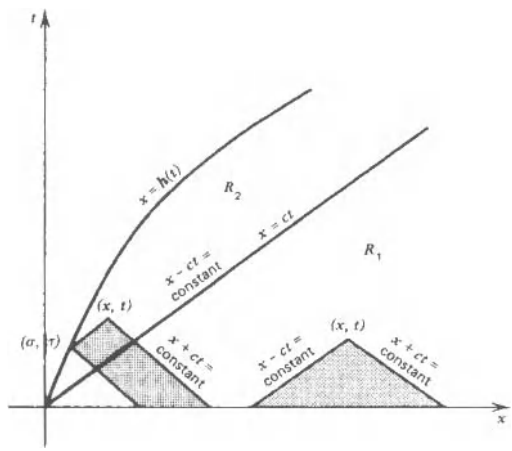

Figure 6.13 The case when $0 < h'(t) < c$.

An instructive special case occurs if we consider the boundary curve $x = h(t) = c_0 t$, $|c_0| < c$, and the periodic boundary condition $u(h(t), t) = B(t) = u(c_0 t, t) = A\cos(\omega t)$, where A and ω are prescribed constants. Thus, a boundary point moves with the constant velocity c_0. We assume that $f(x) = g(x) = 0$. Then $\tau = (ct - x)/(c - c_0)$ and the solution is

$$u(x,t) = \begin{cases} A\cos\left[(c\omega/(c - c_0))\,(t - x/c)\right], & c_0 t < x < ct \\ 0, & ct < x. \end{cases} \qquad (6.5.25)$$

This result has the following interesting interpretation. The wave $u(x, t) = A\cos[(c\omega/(c - c_0))(t - x/c)]$, which results from the effect of the boundary motion—which may be thought to represent a moving energy source—has a frequency of oscillation $\hat{\omega} = [c/(c - c_0)]\omega$, which exceeds the frequency of oscillation ω of the source term [i.e., $A\cos(\omega t)$] if $c_0 > 0$. The converse is true if $c_0 < 0$. In both cases the wave $u(x, t)$ moves to the right. Thus, if the boundary source moves in the direction of the resulting wave motion (i.e., $c_0 > 0$) the frequency $\hat{\omega}$ increases, whereas if it moves

in the opposite direction (i.e., $c_0 < 0$), the frequency $\hat{\omega}$ decreases. This phenomenon is known as the *Doppler effect.*

Spacelike and Timelike Curves

A curve $x = h(t)$ in the (x, t)-plane for which $|h'(t)| < c$, where c is the characteristic or wave speed of the wave equation $u_{tt} = c^2 u_{xx}$, is called a *timelike curve.* If $|h'(t)| > c$, it is called a *spacelike curve.* Any spacelike curve can be used as an initial curve for the wave equation on which u and $\partial u/\partial n$, the normal derivative on the curve, must be specified. Any timelike curve can play the role of a boundary curve on which one condition on $u(x, t)$ can be assigned, as is the case for the time axis $x = 0$. We now demonstrate how the foregoing results can be used to solve an initial and boundary value problem with initial data given on a spacelike curve and boundary data prescribed on a timelike curve.

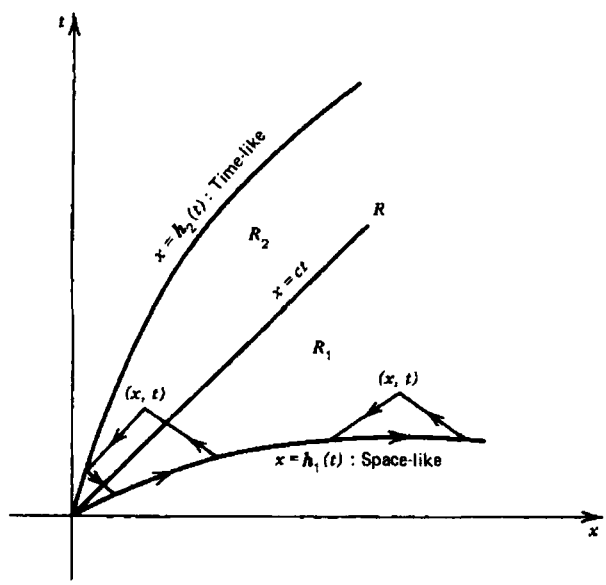

Figure 6.14 Timelike and spacelike curves.

Let $u(x, t)$ be a solution of the wave equation (6.5.21) and suppose that the spacelike curve $x = h_1(t)$ and the timelike curve $x = h_2(t)$ bound the region R depicted in Figure 6.14. Let $u(x, t)$ and $\partial u(x, t)/\partial n$, the normal derivative, be specified on $x = h_1(t)$ as

$$u(h_1(t), t) = f(t), \qquad \frac{\partial u(h_1(t), t)}{\partial n} = g(t), \tag{6.5.26}$$

and let $u(x, t)$ be prescribed on $x = h_2(t)$ as $u(h_2(t), t) = B(t)$.

It follows from (6.5.26) that

$$\begin{cases} \partial u(h_1(t),t)/\partial t = h_1'(t)u_x + u_t = f'(t), \\ \partial u(h_1(t),t)/\partial n = (-u_x + h_1'(t)u_t)\,/\sqrt{1+h_1'(t)^2} = g(t), \end{cases} \tag{6.5.27}$$

since $\partial u(h_1(t),t)/\partial n = ([-1, h_1'(t)] \cdot \nabla u)\,/\sqrt{1+h_1'(t)^2}$. It is possible to solve for u_x and u_t on $x = h_1(t)$ in terms of $f(t)$ and $g(t)$ from (6.5.27). In the region R_1 of Figure 6.14 the solution at an arbitrary point (x,t) can be specified in terms of $f(t)$ and $g(t)$ by applying the formula (6.5.1) to the characteristic triangle shown in the figure. In the region R_2 of Figure 6.14, the solution at a point (x,t) can be specified in terms of $f(t)$, $g(t)$, and $B(t)$ by using (6.5.1) and integrating over the quadrilateral shown in the figure. We observe from the figure that of the two backward characteristics issuing from the point (x,t), only one intersects the timelike curve, whereas both intersect the spacelike curve. This signifies that only one wave is generated by the data on a timelike curve, whereas two waves are generated by data on a spacelike curve.

Characteristic Initial Value Problem

Another problem that is of great interest for the wave equation is the *characteristic initial value problem.* In this problem $u(x,t)$ is prescribed on the two forward characteristics issuing from the point $(x_0, 0)$ and we look for a solution of the wave equation in the interior of the sector bounded by the characteristics. This a problem with moving boundaries, where each boundary moves with the speed of wave propagation. We consider the equation

$$u_{tt}(x,t) - c^2 u_{xx}(x,t) = F(x,t), \qquad x_0 - ct < x < x_0 + ct,\ t > 0, \tag{6.5.28}$$

with the characteristic data, for $t > 0$,

$$u(x,t)|_{x-ct=x_0} = B_1(t),\ u(x,t)|_{x+ct=x_0} = B_2(t), \tag{6.5.29}$$

and the compatibility condition $B_1(0) = B_2(0)$. Using the difference equation (6.5.3) we obtain the solution $u(x,t)$ as

$$u(x,t) = B_1\left[\frac{1}{2}\left(t + \frac{x-x_0}{c}\right)\right] + B_2\left[\frac{1}{2}\left(t + \frac{x_0-x}{c}\right)\right]$$
$$- B_1(0) + \frac{1}{2c}\iint_{CQ} F(\sigma,\tau)\,d\sigma\,d\tau, \tag{6.5.30}$$

where the integration is carried out over the characteristic quadrilateral (denoted as CQ) depicted in Figure 6.15.

It may be noted that in contrast to the results for first order equations, where characteristic initial data resulted in nonunique solutions, the solution (6.5.30) is unique. This can be attributed to the fact that the data are given on two intersecting

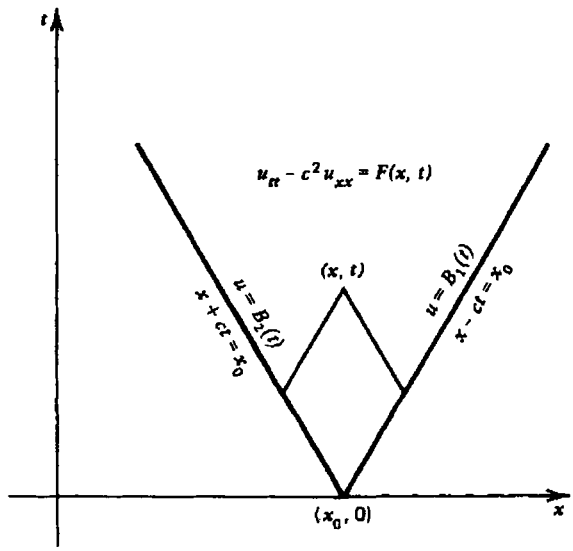

Figure 6.15 The characteristic initial value problem.

characteristics and are compatible at the point of intersection. If $u(x,t)$ is specified on only one characteristic line, the solution is not unique.

The characteristic initial value problem plays a role in *Riemann's method* of solution of the Cauchy problem for the second order hyperbolic equation in two variables discussed in Section 8.3. The equation is assumed to be in the canonical form for hyperbolic equations (apart from the fact that $c \neq 1$),

$$u_{tt}(x,t) - c^2 u_{xx}(x,t) = a u_x(x,t) + b u_t(x,t) + d u(x,t) + F(x,t), \quad (6.5.31)$$

where a, b, and d may be functions of x and t. We assume that the characteristic initial data for $u(x,t)$ are given as in (6.5.29) and are compatible. By treating the right side of (6.5.31) as an inhomogeneous term, (6.5.30) yields the solution in the form of an integral equation

$$u(x,t) = B_1\left[\frac{1}{2}\left(t + \frac{x - x_0}{c}\right)\right] + B_2\left[\frac{1}{2}\left(t + \frac{x_0 - x}{c}\right)\right] - B_1(0)$$

$$+ \frac{1}{2c}\iint_{CQ} F(\sigma,\tau)\, d\sigma\, d\tau + \frac{1}{2c}\iint_{CQ} (a u_\sigma + b u_\tau + d u)\, d\sigma\, d\tau. \quad (6.5.32)$$

We put $u(x,t) = R(x,t) + T[u, u_x, u_t]$, where $R(x,t)$ represents the terms that involve B_1, B_2, and F, and $T[u, u_x, u_t]$ represents the integral terms in u_x, u_t, and u.

The integral equation (6.5.32) can be solved by iteration. We select an arbitrary initial approximation $u^{(0)}(x,t)$ [we can set $u^{(0)} = 0$ in this case] and insert $u^{(0)}$ into the right side of (6.5.32). In the next approximation we have

$$u^{(1)}(x,t) = R(x,t) + T[u^{(0)}, u_x^{(0)}, u_t^{(0)}]. \tag{6.5.33}$$

Iterating this process gives

$$u^{(n)}(x,t) = R(x,t) + T[u^{(n-1)}, u_x^{(n-1)}, u_t^{(n-1)}]. \tag{6.5.34}$$

It can be shown that if the data and the coefficients in (6.5.31) are smooth, the sequence of functions $u^{(n)}(x,t)$ converges to a unique solution of the characteristic initial value problem; that is, $u^{(n)}(x,t) \to u(x,t)$ as $n \to \infty$. We also have $\partial u^{(n)}/\partial x \to \partial u/\partial x$ and $\partial u^{(n)}/\partial t \to \partial u/\partial t$, which implies that $u(x,t)$ satisfies (6.5.32). The details of the proof are not presented here.

Exercises 6.5

6.5.1. Use Taylor's series to show that a solution of the difference equation (6.5.3) must be a solution of the wave equation (6.4.12) in view of the arbitrariness of h and k.

6.5.2. Show that the general solution $u(x,t) = F(x-ct) + G(x+ct)$ of the wave equation (6.4.12), is also a solution of the difference equation (6.5.3).

6.5.3. Obtain the solution $u(x,t)$ for each of the formulas (6.5.15)–(6.5.17) if $B(t) = \sin(\omega t)$. Verify that in each case the solution obtained satisfies the boundary condition.

6.5.4. Let $f(x) = 1$ and $g = \beta = B_1 = B_2 = F = 0$ in the problem (6.5.18)–(6.5.20). Use the method of Example 6.7 to obtain the solution $u(x,t)$ of this problem in the regions R_1, $R2, \ldots, R_7$ indicated in Figure 6.11.

6.5.5. Show that the problem $u_{tt}(x,t) - u_{xx}(x,t) = 0$, $2t < x < \infty$, $t > 0$, $u(x,0) = x$, $u_t(x,0) = 0$, $0 < x < \infty$, $u(2t,t) = 1$, $t > 0$ has no solution.

6.5.6. Solve the initial and (moving) boundary value problem (6.5.21)–(6.5.22) if $f(x) = g(x) = 0$, $h(t) = -c_0 t$, $0 < c_0 < c$, and $B(t) = A\cos(\omega)t$. Obtain the Doppler effect and show that the solution oscillates at a frequency below the input frequency ω.

6.5.7. Solve the problem (6.5.21)–(6.5.22) if $f(x) = 1$, $g(x) = 0$, $h(t) = -c_0 t$, $c_0 > c$, and $B(t) = 1$. Show that the problem has a nonunique solution.

6.5.8. Solve the characteristic initial value problem (6.5.28)–(6.5.29) if $F(x,t) = \cos\omega t$, $B_1(t) = e^{-t}$, and $B_2(t) = 1$.

6.5.9. Let $x = h_1(t) = 2t$ be a spacelike curve on which the data (6.5.26) where $f(t) = \sin t$ and $g(t) = 1$ are given. Solve the Cauchy problem for the wave equation $u_{tt}(x,t) = u_{xx}(x,t)$ in the region $2t - x > 0$ with the given Cauchy data on $x = 2t$.

6.5.10. Solve the *Goursat problem* for the wave equation $u_{tt}(x,t) = c^2 u_{xx}(x,t)$ in the region $0 < x < ct$, $t > 0$ with the data $u(0,t) = g(t)$, $u(ct,t) = f(t)$, $t > 0$, assuming that $f(0) = g(0)$. (A problem in which the data are given on a characteristic curve and on an intersecting timelike curve is called a *Goursat problem.*)

6.5.11. Solve the modified Goursat problem for the wave equation $u_{tt}(x,t) = c^2 u_{xx}(x,t)$ in the region $ct < x < \infty$, $t > 0$ with the data $u(ct,t) = g(t)$, $u(x,0) = f(x)$, assuming that $f(0) = g(0)$. [Here the data are given on a characteristic line and the (spacelike) x-axis.]

6.5.12. Solve the characteristic initial value problem for the hyperbolic equation $u_{tt}(x,t) - u_{xx}(x,t) + 2u_t(x,t) + u(x,t) = 0$, $-t < x < t$, $t > 0$, with the data $u(t,t) = e^{-t}$, $u(-t,t) = e^{t}$, $t > 0$, by using the iteration procedure given in (6.5.32)–(6.5.34). *Hint*: The solution is $u(x,t) = e^{-x}$.

6.5.13. Show that the initial and (moving) boundary value problem for the heat equation $u_t(x,t) - c^2 u_{xx}(x,t) = 0$, $c_0 t < x < \infty$, $t > 0$, $u(x,0) = f(x)$, $0 < x < \infty$, $u(c_0 t, t) = g(t)$, $t > 0$ can be solved by introducing the moving coordinate system $\sigma = x - c_0 t$, $\tau = t$. In the new variables σ and τ we obtain a problem for $\hat{u}(\sigma,\tau)$ in the region $\sigma > 0$, $\tau > 0$ with data given at $\sigma = 0$ and $\tau = 0$. Introduce a change of the dependent variable that yields an initial and boundary value problem for a heat equation with data given at $\sigma = 0$ and $\tau = 0$.

6.6 CONCENTRATED SOURCE OR FORCE TERMS

It has been generally assumed in the preceding sections and chapters that the inhomogeneous term ρF, which occurs in the equations of Section 4.1 and which represents a source or force term, is fairly smoothly distributed throughout the region where the problem is to be considered. It often happens, however, that the source or force term is effectively concentrated near some lower-dimensional region, such as a curve for a problem in two-dimensional space-time. In such cases it is convenient to idealize the inhomogeneous term as having infinite density in the region of concentration and vanishing outside that region. Thus ρF must be singular in the region where it is concentrated, and the given differential equation must be interpreted in terms of the integral relations given in Section 6.1. We assume that although ρF is singular, the integral of ρF over the region of its concentration is finite, and we now show how to characterize the effect of the inhomogeneous term on the solution.

Hyperbolic Equations

We consider the *hyperbolic equation* (6.1.1) and its *integral relation* (6.1.5). If the equation is defined over $(n+1)$-dimensional space-time, the inhomogeneous term ρF is assumed to be concentrated over an n-dimensional region S_0 and to have the property that $\iint_R \rho F \, dv = \int_{S_0} \rho F_0 \, ds$, where ds is a surface or line element. This result is a consequence of the fact that $F(\mathbf{x},t) = 0$ for $(\mathbf{x},t) \notin S_0$, but $F(\mathbf{x},t)$ is concentrated on S_0 so that the integral over R collapses to an integral over S_0. The

function $F_0(\mathbf{x}, t)$ represents the surface distribution of the force or source term over S_0. [$F(\mathbf{x}, t)$ is also known as a *single-layer distribution.*] The solution $u(\mathbf{x}, t)$ of (6.1.4) is assumed to be continuous across S_0. We do not consider *double-layer distributions* of $F(\mathbf{x}, t)$ over S_0. In that case, the solution $u(\mathbf{x}, t)$ is discontinuous across S_0. (See Exercises 7.2.23 and 7.2.24, where some mathematical aspects of the definitions of these distributions are presented.)

To determine the variation of u across S_0, we let the region R collapse onto S_0 in (6.1.4), so that in the limit ∂R coincides with S_0 and, in fact, covers it twice. The limiting normal vectors $\mathbf{n}$ of ∂R [in the integral relation (6.1.4)] have opposite direction at common points on S_0 so that they differ only in sign. This implies that $\int_{S_0} [[p\nabla u, -\rho u_t]]_{S_0} \cdot \mathbf{n}\, ds = -\int_{S_0} \rho F_0\, ds$ since ρ, p, and q are assumed to be smooth functions, so that the contribution from $\iint_R qu\, dv$ vanishes in the limit. The bracket $[\cdots]_{S_0}$ in the integral represents the jump in the quantity across S_0. Since the result is also valid if the integrals are taken over an arbitrary portion of S_0, we conclude from the duBois-Reymond lemma (see Exercise 8.1.9) that the jump condition,

$$[[p\nabla u, -\rho u_t]]_{S_0} \cdot \mathbf{n} = -\rho F_0, \tag{6.6.1}$$

is valid at any point $(\mathbf{x}, t)$ on S_0. The continuity of $u(\mathbf{x}, t)$ across S_0 yields a second jump condition,

$$[u]_{S_0} = 0. \tag{6.6.2}$$

If $\phi(\mathbf{x}, t) = 0$ represents the surface S_0, the left side of (6.6.1) has the form of the left side of (6.3.3). Applying interior differentiation to (6.6.2) in S_0 yields the additional equations (in the three-dimensional case) (6.3.6) for the jumps in the first derivatives across S_0. The jumps in the first derivatives of u can be determined from this system of four equations. However, if S_0 is a characteristic surface, the system has no solution for nonzero F_0. For in that case, the homogeneous version of this system [i.e., (6.3.3)–(6.3.4)] has a nontrivial solution, as we have seen. The difficulty arises because $u(\mathbf{x}, t)$ itself must be discontinuous across S_0, if S_0 is characteristic. Since the (source) surface S_0 is moving at a characteristic speed, it cannot generate a disturbance that travels ahead of it, as that would require a speed greater than the (maximum) characteristic speed. It does, however, leave a disturbance in its wake, and as a result, $u(\mathbf{x}, t)$ has a jump across S_0.

In the special case that S_0 is a cylindrical surface in $(\mathbf{x}, t)$-space with generators parallel to the t-axis, we conclude that $[u_t]_{S_0} = 0$ since $[u]_{S_0} = 0$. Then we can express (6.6.1) in terms of the jump in the normal derivative of $u(\mathbf{x}, t)$ across S_0 as $[\partial u/\partial n]_{S_0} = -\rho F_0/p$, with ρ and p evaluated on S_0.

We only consider problems for (6.1.1) in which S_0 divides the region R, over which the problem is defined, into two disjoint regions. Away from S_0, $u(\mathbf{x}, t)$ satisfies a homogeneous version of (6.1.1) since $F(\mathbf{x}, t) = 0$ there. To solve the problem, we can find $u(\mathbf{x}, t)$ in each of the two regions using the given data for the problem and matching these solutions across S_0 using (6.6.1) – (6.6.2). This approach is used in Example 6.10, with a more direct approach presented in Example 6.9. With appropriate modifications the foregoing matching conditions are valid for the elliptic and parabolic equations of Section 4.2.

One-Dimensional Hyperbolic Equations: Stationary Concentrated Forces

In the one-dimensional version of (6.1.1), if ρF is concentrated at the point $x = x_0$ for all time (i.e., it is stationary), we have instead of (6.6.1)–(6.6.2), $[u]_{x_0} = 0$, $[\partial u/\partial x]_{x_0} = -\rho(x_0)F_0(t)/p(x_0)$. The following example yields an interesting application of the foregoing results.

Example 6.9. The Vibration of a Loaded String. We consider the vibration of a finite string with fixed endpoints and with concentrated masses m_i $(i = 1, 2, \ldots, n)$ placed at points x_i $(i = 1, \ldots, n)$ on the string, where $0 < x_i < l$. As the string vibrates, the masses m_i exert concentrated forces $F_i(t)$ at the points x_i. Newton's second law of motion requires that

$$F_i(t) = -m_i u_{tt}(x_i, t), \qquad i = 1, 2, \ldots, n, \ t > 0, \tag{6.6.3}$$

where $u(x, t)$ is the displacement of the string. The result (6.6.3) follows since the acceleration of the point mass m_i is just the acceleration of the point x_i on the string and this is given by $u_{tt}(x_i, t)$.

The problem to be solved for an arbitrary, not necessarily homogeneous string requires that $u(x, t)$ satisfy

$$\rho(x)\frac{\partial^2 u(x,t)}{\partial t^2} = \frac{\partial}{\partial x}\left(p(x)\frac{\partial u(x,t)}{\partial x}\right), \qquad 0 < x < l, \ t > 0, \ x \neq x_i \tag{6.6.4}$$

[generally, $p(x)$ represents the constant tension T], with the conditions $u(0, t) = 0$, $u(l, t) = 0$, $t > 0$, at the boundary, and the initial conditions $u(x, 0) = f(x)$, $u_t(x, 0) = g(x)$, $0 < x < l$. The jump conditions at $x = x_i$ are

$$[u]_{x_i} = 0, \quad p(x_i)\left[\frac{\partial u}{\partial x}\right]_{x_i} = -F_i(t) = m_i u_{tt}(x_i, t), \qquad i = 1, \ldots, n, \tag{6.6.5}$$

where (6.6.3) was used and ρF is replaced by F in this case.

The initial and boundary value problem (6.6.4)–(6.6.5) can be solved as follows. Let $u(x, t) = M(x)N(t)$, and we obtain, on separating variables, with λ as the separation constant,

$$N''(t) + \lambda N(t) = 0, \qquad t > 0, \tag{6.6.6}$$

$$(p(x)M'(x))' + \lambda\rho(x)M(x) = 0, \qquad 0 < x < l, \ x \neq x_i, \tag{6.6.7}$$

with the boundary conditions $M(0) = 0$, $M(l) = 0$, and the jump or matching conditions $[M(x)]_{x_i} = 0$, $m_i M(x_i)N''(t) = p(x_i)[M'(x)]_{x_i}N(t)$, with $i = 1, \ldots, n$. Using (6.6.6), we can replace the second matching condition by

$$p(x_i)[M'(x)]_{x_i} = \lambda m_i M(x_i) = 0, \qquad i = 1, 2, \ldots, n. \tag{6.6.8}$$

The foregoing constitutes an *eigenvalue problem* for $M(x)$, with the interesting aspect that λ, the eigenvalue parameter, enters not only in the differential equation (6.6.6) but also in the jump conditions. It can be shown that the eigenvalues and eigenfunctions for this problem have essentially the same properties as those occurring in the Sturm-Liouville problem (see the exercises). A vibrating string problem with a load (or mass) placed at one end of the string is considered in Exercises 6.6.4 and 6.6.5.

Once the eigenfunctions are determined, the solution of the initial and boundary value problem is found in the usual manner by using eigenfunction expansions.

One-Dimensional Hyperbolic Equations: Moving Concentrated Forces

Let us consider the one-dimensional versions of the hyperbolic (or parabolic) equations of Section 4.1 and assume that the force or source term moves but is concentrated along the curve $x = h(t)$, which plays the role of S_0 in the foregoing discussion. In the hyperbolic case we use the one-dimensional form of the matching or jump conditions (6.6.1)–(6.6.2) on $x = h(t)$. Since $u(x,t)$ is assumed to be continuous across $x = h(t)$, a relationship between the jumps $[u_x]$ and $[u_t]$ across the curve can be established by differentiating $[u]$ along $x = h(t)$. Also, an expression relating the jumps $[u_x]$ and $[u_t]$ is determined from (6.6.1). This yields

$$\left[\frac{\partial u}{\partial x}\right]_{x=h(t)} = -\frac{\rho\sqrt{1+h'(t)^2}F_0(t)}{p-\rho h'(t)^2}, \tag{6.6.9}$$

where ρ and p are evaluated on $x = h(t)$. [In the parabolic case it is easily seen that the jump condition corresponding to (6.6.1) does not involve a time derivative of u.] The matching conditions for the hyperbolic and parabolic cases can be given in terms of a known function $f_0(t)$ [expressed in terms of $F_0(t)$ as in (6.6.9) for the hyperbolic case] as $[u]_{x=h(t)} = 0$, $[\partial u/\partial x]_{x=h(t)} = -f_0(t)$.

Example 6.10. The Wave Equation with a Moving Concentrated Force. A concentrated force term for the wave equation moves with velocity c_0, and it is assumed that $|c_0| < c$, where c is the speed of wave propagation. Thus the force term moves along the line $x = c_0 t$ in space-time. We begin by assuming that the force term is oscillatory and obtain a solution similar to that for the problem considered in Example 6.8.

The initial value or Cauchy problem is formulated as follows. The function $u(x,t)$ satisfies the wave equation away from the line $x = c_0 t$; that is,

$$u_{tt}(x,t) = c^2 u_{xx}(x,t),\ -\infty < x < c_0 t,\ c_0 t < x < \infty,\ t > 0,\ c_0 \geq 0, \tag{6.6.10}$$

where

$$u(x,t) = \begin{cases} u_1(x,t), & -\infty < x < c_0 t, \\ u_2(x,t), & c_0 t < x < \infty, \end{cases} \qquad t > 0. \tag{6.6.11}$$

The matching conditions are

$$\begin{cases} u_2(c_0t,t) - u_1(c_0t,t) = 0, \\ \partial u_2(c_0t,t)/\partial x - \partial u_1(c_0t,t)/\partial x = -f_0(t) = -A\cos(\omega t), \end{cases} \tag{6.6.12}$$

where A and ω are constants. The initial conditions are

$$\begin{cases} u_1(x,0) = \partial u_1(x,0)/\partial t = 0, & -\infty < x < 0, \\ u_2(x,0) = \partial u_2(x,0)/\partial t = 0, & 0 < x < \infty. \end{cases} \tag{6.6.13}$$

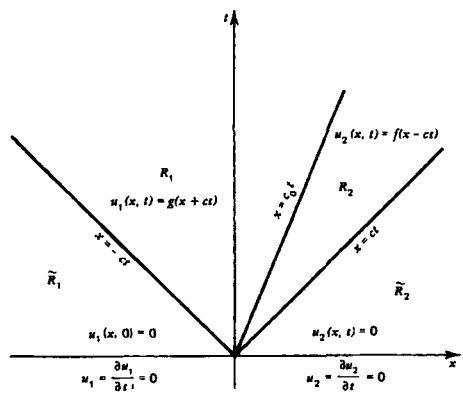

Figure 6.16 The moving force problem.

In the regions $\tilde{R}_1$ and $\tilde{R}_2$ of Figure 6.16 we have $u_1 = u_2 = 0$ since the domains of dependence of points in these regions contain only initial data that are all zero. In R_1 we look for a solution in the form $u_1(x,t) = g(x+ct)$ and in R_2 in the form $u_2(x,t) = f(x-ct)$ since the moving force generates waves moving to the right and to the left. This is so because the curve along which the force moves is timelike and characteristics issue from either side of the curve. (A more systematic approach based on the integral wave equation of Section 6.5 can also be used.) Applying the matching conditions (6.6.12), we readily obtain

$$\begin{cases} u_1(x,t) = (c^2 - c_0^2)/(2\omega c)A\sin\left[\omega(x+ct)/(c+c_0)\right], & -ct < x < c_0t, \\ u_2(x,t) = (c_0^2 - c^2)/(2\omega c)A\sin\left[\omega(x-ct)/(c-c_0)\right], & c_0t < x < ct. \end{cases} \tag{6.6.14}$$

We again observe the *Doppler effect* in (6.6.14). The wave $u_1(x,t)$, which moves in a direction opposite to that of the force if $c_0 > 0$, has the frequency $\hat{\omega} = [c/(c+c_0)]\omega$ which is smaller than the frequency ω of the force term. The wave $u_2(x,t)$, which travels in the same direction as the force has the frequency $\tilde{\omega} = [c/(c-c_0)]\omega$, which exceeds the frequency ω of the force term.

If the force speed c_0 exceeds the characteristic speed c, the foregoing results are not valid. The problem is formulated as in (6.6.10)–(6.6.13), except that we consider

an arbitrary force term $f_0(t)$ (i.e., it need not be oscillatory) and assume that $c_0 > c$. Thus, the line $x = c_0 t$ lies between the characteristic line $x = ct$ and the positive x-axis and is not situated as shown in Figure 6.16. Because the initial data (6.6.13) are zero, we find that $u_1(x,t) = 0$ in the sector between the negative x-axis and the line $x = -ct$. Similarly, $u_2(x,t) = 0$ in the sector between the positive x-axis and the line $x = c_0 t$. Then the matching conditions (6.6.12) determine the (Cauchy) data on the spacelike line $x = c_0 t$ to be

$$u_1(c_0 t, t) = 0, \qquad \frac{\partial u_1(c_0 t, t)}{\partial x} = f_0(t). \tag{6.6.15}$$

To solve for $u_1(x,t)$ in the sector bounded by $x = c_0 t$ and $x = ct$, we construct a characteristic triangle whose base is a segment of the line $x = c_0 t$. Using the integral wave equation (6.5.1) gives

$$u_1(x,t) = \frac{c^2 - c_0^2}{2c} \int_{(x-ct)/(c_0-c)}^{(x+ct)/(c_0+c)} f_0(s)\, ds, \qquad ct \le x \le c_0 t. \tag{6.6.16}$$

Now $u(x,t)$ is continuous across the characteristics $x = \pm ct$ since the initial data are continuous at the origin. Consequently, $u_1(x,t)$ at $x = ct$ equals (6.6.16) evaluated on that line, and $u_1(x,t) = 0$ on $x = -ct$ since $u_1(x,t)$ vanishes to the left of $x = -ct$. To complete the determination of $u(x,t)$, we must solve the characteristic initial value problem in the sector bounded by the lines $x = \pm ct$ with the foregoing data. From (6.5.30) we obtain

$$u_1(x,t) = \frac{c^2 - c_0^2}{2c} \int_0^{(x+ct)/(c_0+c)} f_0(s)\, ds, \qquad -ct \le x \le ct. \tag{6.6.17}$$

Thereby, the solution is completely determined.

In the problem considered previously where the force speed is less than the characteristic speed, the moving force point gives rise to two waves, one of which precedes the force. In the present problem, the force point again gives rise to two waves, one moving to the right and the other to the left. However, no wave precedes the force point, since the wave speed is less than the force speed.

If the force speed $c_0 = c$, the characteristic speed, the solution $u(x,t)$ must be discontinuous across the path of the force, as was noted previously. In that case the concentrated force gives rise to only one wave that travels in its wake. Ahead of the moving force point the solution is zero if the initial data are zero. As a result, if the solution were continuous on the path of the force, this wave would have zero height on the force path. Since the wave is constant on each characteristic that intersects the path of force, it must have zero height everywhere, and this is not possible.

Exercises 6.6

6.6.1. Adapt the method presented in (4.3.16)–(4.3.17) to determine that the orthogonality condition for the eigenvalue problem given for (6.6.7) is $\int_0^l M_j(x)M_k(x)\rho(x)\,dx + \sum_{i=1}^n m_i M_j(x_i)M_k(x_i) = 0$, where $M_j(x)$ and $M_k(x)$ correspond to the eigenvalues λ_j and λ_k with $\lambda_j \neq \lambda_k$. Conclude that the appropriate inner product associated with this eigenvalue problem is $(\phi(x), \psi(x)) = \int_0^l \phi(x)\psi(x)\rho(x)\,dx + \sum_{i=1}^n m_i\phi(x_i)\psi(x_i)$. The (induced) norm is given as $||\phi(x)||^2 = (\phi(x), \phi(x)) = \int_0^l \phi^2(x)\rho(x)\,dx + \sum_{i=1}^n m_i[\phi(x_i)]^2$. Show that the eigenvalues are real and positive. Determine a formula for the Fourier coefficients c_k in an expansion of a function $f(x)$ in a series of the eigenfunctions $M_k(x)$.

6.6.2. Let $\rho(x) = p(x) = 1$, $l = \pi$, $n = 1$, and $x_1 = 1$ in the problem for (6.6.7). Determine an equation for the eigenvalues and obtain the corresponding eigenfunctions.

6.6.3. Expand $f(x) = 1$ in a series of eigenfunctions found in Exercise 6.6.2.

6.6.4. Assume that the load in the vibrating string problem of Example 6.9 is placed at the end of the string at $x = l$. The boundary condition $u(l, t) = 0$ is then replaced by $m_1 u_{tt}(l, t) = -p(l)u_x(l, t)$. Show why this follows from (6.6.5). Use separation of variables to obtain the eigenvalue problem for this case that replaces the one given in Example 6.9.

6.6.5. Let $\rho = p = 1$ in the problem of Exercise 6.6.4. Show that the eigenvalues for the resulting eigenvalue problem are determined from the equation $\cot(l\sqrt{\lambda}) = m_1\sqrt{\lambda}$. Demonstrate graphically that there are infinitely many eigenvalues $\lambda = \lambda_k$. Show that the eigenfunctions $M_k(x)$ are given as $M_k(x) = \sin(\sqrt{\lambda_k}\,x)/\sin(\sqrt{\lambda_k}\,l)$, $k = 1, 2, \ldots$, and that they are orthogonal with respect to the inner product of Exercise 6.6.1, where $n = 1$ and $x_1 = l$. Calculate the norm of these eigenfunctions.

6.6.6. Obtain the limit of the solutions (6.6.14) as $c_0 \to 0$ and discuss the resulting problem and its solution in that case.

6.6.7. Let $f_0(t) = A\cos(\omega t)$ in (6.6.15) and obtain the solution of the given problem in that case. Discuss the Doppler effect.

6.6.8. Use the matching conditions (6.6.9) to solve the heat equation $\rho u_t(x, t) - pu_{xx}(x, t) = 0$, $-\infty < x < \infty$, $t > 0$, $x \neq 0$, where ρ and p are positive constants, if $u(x, 0) = 0$ and $h(t) = 0$ in (6.6.9). Use the Laplace transform.

6.6.9. Solve the problem of Exercise 6.6.8 by means of the Fourier transform. *Hint*: Use the jump condition (6.6.9) to obtain an expression for the transform of the second derivative of $u(x, t)$ in terms of the transform of $u(x, t)$.

6.6.10. Derive the jump condition (6.6.9) for the hyperbolic case and show that the corresponding condition for the parabolic case is obtained by replacing the denominator in (6.6.9) by p.

6.6.11. Consider the heat equation $u_t(x,t) = c^2 u_{xx}(x,t)$, $-\infty < x < \infty$, $t > 0$, with $u(x,0) = 0$ and a source moving along the line $x = h(t) = c_0 t$, across which the jump conditions $[u] = 0$ and $[u_x] = -(\sqrt{1 + c_0^2/c^2})F_0(t)$ are satisfied. Use the transformation of Exercise 6.5.13 to change this problem into one with a stationary source point and solve the given problem.

6.6.12. Consider the following moving boundary (and source) problem $u_t(x,t) = c^2 u_{xx}(x,t)$, $0 < x < h(t)$, $t > 0$, $u(0,t) = u_0$, $u[h(t),t] = 0$, $u_x[h(t),t] = -ah'(t)$, where $h(0) = 0$, and the determination of $h(t)$ is part of the problem. (The constants $a > 0$ and $u_0 > 0$ are given.) This is an example of a one-phase *Stefan problem*, in which the region $x > 0$ is filled with ice, taken to be at zero temperature, and the boundary $x = 0$ is kept at a positive temperature u_0. As t increases from zero, the ice begins to melt and the curve $x = h(t)$ separates the water and the ice. In effect, we are solving a heat conduction problem in a composite region composed of water and ice, and the conditions along the separation curve $x = h(t)$ are jump conditions determined from the properties of heat flow across such a boundary. However, since the temperature $u(x,t)$ of the ice remains at zero while in its frozen state, we may treat the separation curve as a boundary curve for the liquid state, and treat the jump conditions as boundary conditions. Nevertheless, this problem is complicated by the fact that $x = h(t)$ is not known a priori.

To solve this problem, assume that $u(x,t)$ has the form of $u_2(x,t)$ given in Exercise 6.2.4, and substitute it into the three boundary conditions. Conclude that $h(t)$ must be a constant multiple of $\sqrt{t}$, and show how this constant can be determined numerically.

6.7 POINT SOURCES AND FUNDAMENTAL SOLUTIONS

In this section we assume that the inhomogeneous term ρF in the differential equations of Section 4.1 that represents a source or force term is effectively concentrated at a point P_0. (To simplify the presentation, the term *source* will be used in the following, even though it may represent a force in some contexts.) We consider two cases. To begin, we deal with higher-dimensional hyperbolic and parabolic equations and assume that ρF is concentrated at the point $P_0 = \mathbf{x}_0$ for all time $t \geq 0$. The one-dimensional version of this problem was considered in Section 6.6. This represents a *stationary point source*. Next, we assume that the point of concentration in the hyperbolic or parabolic case is $P_0 = (\mathbf{x}_0, t_0)$ and in the elliptic case, $P_0 = \mathbf{x}_0$. For time-dependent problems this constitutes an *instantaneous point source*. (We use notation for the spatial variable appropriate for the two- or three-dimensional cases. However, we will also consider problems in one dimension.)

In both cases, ρF can be expressed in terms of the Dirac delta function, which was discussed briefly in Chapter 1. Since the theory of the delta function and other generalized functions is not presented until Chapter 7, we discuss the foregoing problems, as far as possible, on the basis of the integral relations of Section 6.1. However, the theory of Green's functions developed in Chapter 7 is closely related to the point source problem.

Hyperbolic and Parabolic Equations: Stationary Point Sources

We start with the integral relations (6.1.6) and (6.1.9) for the hyperbolic and parabolic cases, respectively. We assume that ρF is concentrated at the point $P_0 = \mathbf{x}_0$ in the region $R_\mathbf{x}$ for all $t > 0$ and that it has the property $\iint_{R_\mathbf{x}} \rho F \, d\mathbf{x} = f(t)$, where $f(t)$ is a (smooth) function that characterizes the source strength for $t > 0$ and vanishes for $t < 0$. In the limit as $R_\mathbf{x}$ shrinks down to the point P_0, we obtain from (6.1.6) and (6.1.9)

$$\lim_{\partial R_\mathbf{x} \to P_0} \int_{\partial R_\mathbf{x}} p \frac{\partial u}{\partial n_\mathbf{x}} \, ds_\mathbf{x} = -f(t). \tag{6.7.1}$$

Although (6.7.1) should, in fact, be integrated over the interval $t_0 < t < t_1$, as is the case in (6.1.6) and (6.1.9), our result is valid since the interval of integration is arbitrary. The integrals in (6.1.6) and (6.1.9) that involve u or u_t vanish in the limit as $R_\mathbf{x} \to P_0$ since these terms are not more singular than $\mathbf{x}$-derivatives in u and the relevant integrals are higher dimensional than the integral in (6.7.1).

The result (6.7.1) is valid for an arbitrary region $R_\mathbf{x}$ and its boundary $\partial R_\mathbf{x}$. It takes on a simpler and more useful form if $R_\mathbf{x}$ is chosen to be a disk or a ball. In the two-dimensional case, let $R_\mathbf{x}$ be the disk of radius r with center at $P_0 = (x_0, y_0)$ and introduce polar coordinates with the pole at P_0. Then

$$\lim_{\partial R_\mathbf{x} \to (x_0, y_0)} \int_{\partial R_\mathbf{x}} p \frac{\partial u}{\partial n_\mathbf{x}} \, ds_\mathbf{x} = \lim_{r \to 0} \int_0^{2\pi} p \frac{\partial u}{\partial r} r \, d\theta = \lim_{r \to 0} \left(2\pi p r \frac{\partial u}{\partial r} \right) = -f(t), \tag{6.7.2}$$

with the θ-integral taken over the circle of radius r. In the three-dimensional case let $R_\mathbf{x}$ be a ball of radius r with center at $P_0 = (x_0, y_0, z_0)$ and introduce spherical coordinates with the pole at P_0. Then (6.7.1) easily yields

$$\lim_{\partial R_\mathbf{x} \to (x_0, y_0, z_0)} \int_{\partial R_\mathbf{x}} p \frac{\partial u}{\partial n_\mathbf{x}} \, ds_\mathbf{x} = \lim_{r \to 0} \left(4\pi p r^2 \frac{\partial u}{\partial r} \right) = -f(t), \tag{6.7.3}$$

on integrating over the sphere of radius r. In (6.7.2) and (6.7.3) we have used the fact that $\partial u / \partial n_\mathbf{x} = \partial u / \partial r$ on $\partial R_\mathbf{x}$ and assumed that the mean value theorem for integrals can be applied in both cases.

The *stationary point source problem* can now be formulated as follows. The function $u(\mathbf{x}, t)$ satisfies the homogeneous form of the relevant hyperbolic or parabolic equation for $\mathbf{x} \neq \mathbf{x}_0$ and $t > 0$, since ρF vanishes away from the source point P_0. At P_0, $u(\mathbf{x}, t)$ must satisfy the condition (6.7.2) or (6.7.3). Initially, at $t = 0$ it is required that $u(\mathbf{x}, t)$, as well as $u_t(\mathbf{x}, t)$ in the hyperbolic case, vanish. If the problem is given over a bounded spatial region, $u(\mathbf{x}, t)$ must satisfy homogeneous boundary conditions.

The foregoing problem requires a different approach from that used in solving the concentrated source problem of the preceding section or, for that matter, all the problems of the previous sections. In the earlier problems it was possible to divide the region R over which the problems were formulated into two (or more) subregions separated by the curves or surfaces on which the sources or singularities were

concentrated. Consequently, these curves or surfaces could be treated as additional boundaries for the problems, on which boundary conditions (i.e., matching or jump conditions) could be assigned. For the point source problem above and the one treated in the following, such a subdivision is not a priori possible because the source is concentrated either on a line or at a point and, as such, cannot represent a boundary for the region, given the space-time or space dimension of the problems we are considering. This makes it difficult to obtain exact results in the general case. However, if the problem has certain symmetries, such as axial symmetry, it may be possible to reduce the number of independent variables and treat the source points as boundary points, as was done in Example 5.9. This is the case in the examples presented below.

The nature of the singularity of the solution $u(\mathbf{x}, t)$ of this problem at the source point may be determined by treating (6.7.2)–(6.7.3) as approximate equations near P_0. Integrating (6.7.2) yields, with $r^2 = (x - x_0)^2 + (y - y_0)^2$,

$$u(x, y, t) \approx \frac{-f(t)}{2\pi p(x_0, y_0)} \log r, \quad (x, y) \approx (x_0, y_0), \tag{6.7.4}$$

in the two-dimensional case. In three dimensions (6.7.3) gives, with $r^2 = (x-x_0)^2 + (y - y_0)^2 + (z - z_0)^2$,

$$u(x, y, z, t) \approx \frac{-f(t)}{4\pi p(x_0, y_0, z_0) r}, \qquad (x, y, z) \approx (x_0, y_0, z_0). \tag{6.7.5}$$

In both cases only the most singular part of the solutions has been retained.

In the following two examples we solve the stationary point source problems for the wave and heat equations. Because the equations have constant coefficients, exact solutions can be found.

Example 6.11. The Stationary Point Source Problem for the Wave Equation. Let $u(\mathbf{x}, t)$ satisfy the two- or three-dimensional wave equation

$$u_{tt}(\mathbf{x}, t) - c^2 \nabla^2 u(\mathbf{x}, t) = 0, \qquad \mathbf{x} \neq \mathbf{x}_0, \quad t > 0, \tag{6.7.6}$$

everywhere except at the source point $P_0 = \mathbf{x}_0$. We assume that $u(\mathbf{x}, 0) = u_t(\mathbf{x}, 0) = 0$, and that at P_0 the condition (6.7.2) or (6.7.3) is satisfied.

To solve this problem we let $r = |\mathbf{x} - \mathbf{x}_0|$, the (spatial) distance from the source point $P_0 = \mathbf{x}_0$, and look for a solution of (6.7.6) in the form $u = u(r, t)$. Then (6.7.6) becomes

$$u_{tt}(r, t) - c^2 \left[u_{rr}(r, t) + \frac{n-1}{r} u_r(r, t) \right] = 0, \; r > 0, t > 0, \quad n = 2, 3, \tag{6.7.7}$$

in the case of two or three (space) dimensions.

We begin with the *three-dimensional case* since it is easier to handle. Let $n = 3$ in (6.7.7) and put $v(r, t) = ru(r, t)$, where $r^2 = (x-x_0)^2 + (y-y_0)^2 + (z-z_0)^2$. Then we easily verify that $v(r, t)$ satisfies the one-dimensional wave equation $v_{tt}(r, t) -$

$c^2 v_{rr}(r,t) = 0$, which has the general solution $v(r,t) = F[t - r/c] + G[t + r/c]$. Thus $u(r,t) = (1/r)(F[t - r/c] + G[t + r/c])$. As $u_t(r,t)$ must vanish at $t = 0$ and $u(r,t)$ satisfies the source point condition (6.7.3) where $f(t) = 0$ for $t < 0$, we find that $G[t + r/c] = 0$ and obtain

$$u(r,t) = \frac{1}{4\pi c^2 r} f\left(t - \frac{r}{c}\right). \tag{6.7.8}$$

Because $f(t)$ vanishes for negative t, $u(r,t)$ vanishes at $t = 0$ when $r > 0$. Near $r = 0$, (6.7.8) behaves like (6.7.5), where we must set $p = c^2$.

The solution (6.7.8) represents a *propagating spherical wave* that starts to spread out from the source point P_0 at the time $t = 0$ with speed c. The spherical wave fronts are given as $r = ct$ because $u(r,t) = 0$ when $r > ct$, since $f(t)$ vanishes when its argument is negative.

In the *two-dimensional case*, (6.7.7) cannot be simplified as nicely as was done for $n = 3$. However, the point source problem for $n = 2$ can be solved in terms of *Hankel transforms*, in which case the condition (6.7.2) at the source can effectively be treated as a boundary condition at $r = 0$ (see Examples 5.7 and 5.8).

A simpler method for solving this problem uses the three-dimensional result (6.7.8) and proceeds as follows. (This approach is analogous to *Hadamard's method of descent* of Section 5.4.) We think of the point source at $P_0 = (x_0, y_0)$ as representing a line source for the three-dimensional problem, located on the line $x = x_0$, $y = y_0$, $z = z$ in space. By summing (i.e., integrating) over all source points (x_0, y_0, z_0) along this line, we obtain the solution to the two-dimensional problem. Thus we integrate (6.7.8) [where $r^2 = (x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2$] with respect to z_0 from $-\infty$ to $+\infty$ and obtain

$$u(x,y,t) = \frac{1}{4\pi c^2}\int_{-\infty}^{\infty} \frac{1}{r} f\left(t - \frac{r}{c}\right)\, dz_0 = \frac{1}{2\pi c^2}\int_0^{\infty} \frac{1}{r} f\left(t - \frac{r}{c}\right)\, d\hat{z}$$

$$= \frac{1}{2\pi c}\int_0^{t-\hat{r}/c} [c^2(t-\tau)^2 - \hat{r}^2]^{-1/2} f(\tau)\, d\tau, \tag{6.7.9}$$

where $\hat{z} = z - z_0$, $\hat{r}^2 = (x - x_0)^2 + (y - y_0)^2$, $\tau = t - r/c$ and we have used the fact that $f(t) = 0$ for $t < 0$. Also, (6.7.9) vanishes at $t = 0$ since $f(t)$ vanishes for negative values of its argument. In fact, (6.7.9) vanishes when $\hat{r} > ct$. Thus, the solution (6.7.9) is a *propagating circular (or cylindrical) wave* and the circles (or cylinders) $\hat{r} = ct$ are wave fronts for the two-dimensional point source problem.

For the sake of completeness, we include the solution of the point source problem for the *one-dimensional wave equation*. This problem can be solved by the methods of the preceding section. If the source acts at the point x_0, we require that $u(x,t)$ and $u_t(x,t)$ vanish at $t = 0$ and that the foregoing jump conditions are satisfied, with $\rho = 1$, $p = c^2$, and $F_0(t) = f(t)$. To solve, we set $u = F[t + (x - x_0)/c]$ for $x < x_0$ and $u = G[t - (x - x_0)/c]$ for $x > x_0$. Using the initial and jump conditions, we

easily obtain the solution

$$u(x,t) = \frac{1}{2c}\int_0^{t-|x-x_0|/c} f(s)\,ds. \tag{6.7.10}$$

Since $f(t) = 0$ for $t < 0$, we have $u(x,t) = 0$ for $|x - x_0| > ct$ and (6.7.10) represents two *propagating plane waves* that travel away from the source.

The solutions of the point source problem demonstrate an important distinction in the nature of wave propagation in one, two, and three dimensions. To see this we assume that the source function $f(t)$ vanishes for $t > \tau > 0$ (i.e., it acts only for a finite time). Then (6.7.8) in the three-dimensional case vanishes when $r < c(t-\tau)$, in addition to the fact that it is zero ahead of the spherical wave front $r = ct$. That is, the (moving) disturbance is concentrated in the region $c(t-\tau) \le r \le ct$. The signal has a sharp leading and trailing edge. In the two- and one-dimensional cases, the solution or signal has a sharp leading edge but has a nonzero wake for all time. This distinction represents what is known as *Huygens' principle* (see Section 5.4).

Example 6.12. The Stationary Point Source Problem for the Heat Equation. Let $u(\mathbf{x},t)$ satisfy the two- or three-dimensional equation of heat conduction

$$u_t(\mathbf{x},t) - c^2\nabla^2 u(\mathbf{x},t) = 0, \qquad \mathbf{x} \neq \mathbf{x}_0, \quad t > 0, \tag{6.7.11}$$

everywhere except at the source point $P_0 = \mathbf{x}_0$. We assume that $u(\mathbf{x},0) = 0$, and that at P_0 the condition (6.7.2) or (6.7.3) is satisfied.

To solve this problem we look for a solution in the form $u = u(r,t)$, where $r = |\mathbf{x} - \mathbf{x}_0|$ is the distance from the source point P_0. Then (6.7.11) becomes

$$u_t(r,t) - c^2\left[u_{rr}(r,t) + \frac{n-1}{r}u(r,t)\right] = 0, \; r > 0, t > 0, \quad n = 2,3, \tag{6.7.12}$$

in two or three dimensions.

For the *three-dimensional case*, with $n = 3$ in (6.7.12), we put $v(r,t) = ru(r,t)$ and find that $v(r,t)$ satisfies the one-dimensional heat equation $v_t(r,t) = c^2 v_{rr}(r,t)$. At $t = 0$ we have $v(r,0) = 0$. We assume that $v(r,t)$ and $v_r(r,t)$ are bounded at $r = 0$, so that (6.7.3) yields the condition $v(0,t) = (1/4\pi c^2)f(t)$. The resulting initial and boundary value problem for $v(r,t)$ for $r > 0$ and $t > 0$ was solved in Example 5.5 [see (5.3.22)]. In terms of that result for $v(r,t)$ we obtain the solution

$$u(r,t) = \int_0^t [4\pi c^2(t-\tau)]^{-3/2} \exp\left[-\frac{r^2}{4c^2(t-\tau)}\right] f(\tau)\,d\tau. \tag{6.7.13}$$

In the *two-dimensional case*, we set $n = 2$ in (6.7.12). We have $u(r,0) = 0$ and at $r = 0$ the condition (6.7.2) must be satisfied. We use the zero-order Hankel transform to solve this problem and proceed as in Example 5.9. We easily determine

that the Hankel transform $U(\lambda, t)$ satisfies the equation $U_t(\lambda, t) + (c\lambda)^2 U(\lambda, t) = (1/2\pi) f(t)$. Also, $U(\lambda, 0) = 0$. Solving this equation, inverting the transform, interchanging the order of integration in the resulting double integral, and using (5.5.7) yields the solution

$$u(r,t) = \int_0^t [4\pi c^2 (t-\tau)]^{-1} \exp\left[-\frac{r^2}{4c^2(t-\tau)}\right] f(\tau)\, d\tau. \tag{6.7.14}$$

For the *one-dimensional case*, with the source located at $x = x_0$, we cite the result of Exercise 6.6.8, which yields the solution

$$u(x,t) = \int_0^t G(x - x_0, t - \tau) f(\tau)\, d\tau, \tag{6.7.15}$$

where $G(x,t)$ is the fundamental solution of the heat equation defined in equation (5.2.39).

In fact, in each of the three problems considered in this example, the solution of the point source problem is given as a product of the fundamental solution of the heat equation with the density function $f(\tau)$, integrated from 0 to t.

Point Sources and Instantaneous Point Sources

Next we consider *point source problems* for *elliptic equations* and *instantaneous point source problems* for *parabolic* and *hyperbolic equations*. In the elliptic case the source point is $P_0 = \mathbf{x}_0$, while in the hyperbolic and parabolic cases it is $P_0 = (\mathbf{x}_0, t_0)$. The source is assumed to have unit strength, so that the concentration of ρF at the point P_0 implies that

$$\iiint_R \rho F\, dv = 1 \tag{6.7.16}$$

for any region R in space or space-time that contains P_0, while ρF vanishes when $P \neq P_0$. Thus ρF is identical with the Dirac delta function with singular point at P_0.

Since $\rho F = 0$ for $P \neq P_0$, the solution $u(\mathbf{x})$ or $u(\mathbf{x}, t)$ of the point source problem satisfies the homogeneous form of the relevant differential equation or the equivalent integral relation away from P_0. The behavior of u at P_0 must be determined from the integral relation, since u is not expected to be smooth at P_0.

Proceeding formally, we consider the integral relation (6.1.4) for the *hyperbolic case* and assume that the region R contains the source point P_0. Then as R shrinks down to P_0, we obtain the formal limit

$$\lim_{\partial R \to P_0} \int_{\partial R} [p \nabla u, -\rho u_t] \cdot \mathbf{n}\, ds = -1 \tag{6.7.17}$$

on using (6.7.16) and assuming the other terms in the integral relation vanish in the limit. Similarly, the integral relation (6.1.8) for the *parabolic case* yields, as $R \to P_0$,

$$\lim_{\partial R \to P_0} \int_{\partial R} [p \nabla u, -\rho u] \cdot \mathbf{n}\, ds = -1. \tag{6.7.18}$$

In the *elliptic case* the integral relation (6.1.11) gives

$$\lim_{\partial R \to P_0} \int_{\partial R} p \frac{\partial u}{\partial n}\, ds = -1. \tag{6.7.19}$$

In obtaining (6.7.18) and (6.7.19) we have used (6.7.16) and assumed that all other terms in the integral relations (6.1.8) and (6.1.11) vanish in the limit as $R \to P_0$.

Now for the *hyperbolic problem* the singularity of $u(\mathbf{x}, t)$ does not remain isolated at the source point P_0 but is transmitted along the characteristics that contain that point. Thus $u(\mathbf{x}, t)$ may not possess the derivatives near P_0 required to make the integral relation (6.1.4) and its limit (6.7.17) valid. In that case we must replace the foregoing integral relation by that obtained in Section 6.4 for which the solutions $u(\mathbf{x}, t)$ were not required to be differentiable or even continuous. The relevant integral relation is (6.4.5), and in terms of that expression it is easy to see that (6.7.17) should be replaced by

$$\lim_{\partial R \to P_0} \iint_R u[\rho v_{tt} - \nabla \cdot (p \nabla v) + qv]\, dV = v(P_0). \tag{6.7.20}$$

This condition must be satisfied for every smooth function $v(\mathbf{x}, t)$ that vanishes near the boundary of R. More precisely, we require that u be a weak solution of the given hyperbolic equation, in the sense that it satisfies (6.4.5), and this takes the form of (6.7.20) with the limit process removed.

In the hyperbolic problem, the family of characteristic curves or surfaces containing a given point $P_0 = (\mathbf{x}_0, t_0)$ envelops a (singular) characteristic known as the *characteristic conoid*, which has a vertex at P_0. This conoid reduces to the two characteristic curves that pass through the point $P_0 = x_0$ in the case of one space dimension. For the wave equation in three dimensions, the conoid coincides with the characteristic cone (3.3.21). The solution of the point source problem is thus expected to be singular along the entire characteristic conoid and it appears that (6.7.20) rather than (6.7.17) is the relevant condition at P_0 for the hyperbolic case.

In fact, if we formally set $f(t) = \delta(t - t_0)$ [where $\delta(t)$ is the Dirac delta function and $t_0 > 0$] in the point source problem for the wave equation considered in Example 6.11, the problem reduces to that for an instantaneous point source at $(\mathbf{x}_0, t_0)$. It has the formal solution $u(r, t) = 1/(4\pi c^2 r)\delta[t - t_0 - r/c]$, as follows from (6.7.8). This shows the solution to be singular not just at P_0 but along the entire forward characteristic cone $r = c(t - t_0)$.

In the *parabolic problem*, the characteristic curves or surfaces that contain the point $P_0 = (\mathbf{x}_0, t_0)$ are the lines or planes $t = t_0$, as follows from the results of Chapter 3. However, as shown, the singularity in $u(\mathbf{x}, t)$ is confined to the point P_0, and does not spread throughout the characteristic $t = t_0$. Thus, it is unnecessary to replace (6.7.18) by a condition based on the weak formulation of the problem given in (6.4.9). Although this simplifies the parabolic case in comparison with the hyperbolic case, the condition (6.7.18) is not always easy to apply.

The source point condition (6.7.19) for the *elliptic case* poses the fewest problems. Since elliptic equations have no real characteristics, the singularity of $u(\mathbf{x})$ must be

confined to the source point $P_0 = \mathbf{x}_0$, and u and its derivatives are expected to be smooth on ∂R in (6.7.19). Effectively, the elliptic point source problem can be treated in the same way as the continuous point source problem for hyperbolic and parabolic equations.

In view of the foregoing, we discuss the elliptic point source problem on the basis of the condition (6.7.19), while the point source problems for the hyperbolic and parabolic cases are dealt with in a different way. It is generally easier to deal with the latter problems by expressing ρF as a δ-function and using the theory of generalized functions to solve the problem, as is done in Chapter 7. Nevertheless, many results for the point source and related problems were obtained before the theory of generalized functions was developed.

Fundamental Solutions

Solutions of the foregoing point source problems are called *fundamental solutions* of the given differential equation. As such, they are merely required to satisfy the homogeneous form of the equation away from the source point P_0 and the appropriate condition (6.7.17), (6.7.18), (6.7.19), or (6.7.20) at P_0. Fundamental solutions are, therefore, not determined uniquely since any smooth solution of the homogeneous differential equation can be added to them. Here we will be mostly concerned with fundamental solutions that are required to satisfy additional conditions that serve to specify them uniquely.

For the *hyperbolic* and *parabolic problems* we require that the fundamental solution $u(\mathbf{x}, t)$ satisfies a *causality condition*. This states that the solution $u(\mathbf{x}, t)$ must vanish identically for $t < t_0$ when $P_0 = (\mathbf{x}_0, t_0)$ is the source point. Thus we are considering the effect of the source point P_0 on the solution $u(\mathbf{x}, t)$ in the region $t > t_0$. This is referred to as the *causal fundamental solution*. The causality assumption translates into the condition that u and u_t vanish at $t = t_0$ for all $P \neq P_0$. That is,we effectively have an initial value problem for the given equations with zero data away from P_0 and with the data at P_0 determined from the source point conditions. Since $t = t_0$ is a characteristic for the parabolic problem, we can arbitrarily set $u(\mathbf{x}, t) = 0$ for $t < t_0$ since singularities in the solution are confined to the point P_0. For the hyperbolic problem (as we have already observed in a number of special cases), the domain of influence of a point $P_0 = (\mathbf{x}_0, t_0)$ is the forward characteristic conoid (i.e., that part of it for which $t > t_0$) with vertex at P_0. Then with $u(\mathbf{x}, t) = 0$ for $t < t_0$, the causal fundamental solution vanishes identically in the exterior of the forward characteristic conoid.

For the *elliptic problem* we require that the fundamental solution have a specified behavior at infinity. It is then known as the *free space Green's function*. Here we consider only fundamental solutions in unbounded regions. If the point source problem is given for a bounded region, the fundamental solution must satisfy homogeneous boundary conditions. Solutions of such problems are referred to as *Green's functions*. Green's functions for equations of all three types are considered in Chapter 7. The Green's functions determined in that chapter for hyperbolic and parabolic equations in unbounded regions differ from the causal fundamental solutions obtained here.

Fundamental Solutions of Elliptic Equations

Elliptic problems are easiest to deal with, so we begin our discussion of fundamental solutions with the elliptic case. In one dimension, where we are dealing effectively with an ODE for $u = u(x)$, the source point condition (6.7.19) is reduced to a jump condition for $u'(x)$ at the point x_0. In two and three dimensions the condition (6.7.19) is essentially of the form (6.7.1) if we set $f(t) = 1$. Consequently, the simplified forms of (6.7.1) given in (6.7.2) and (6.7.3) are valid for the present case as well if we set $f(t) = 1$ in these equations. With the number of dimensions given by n, we conclude that the source point condition (6.7.19) can be expressed as

$$\begin{cases} [du/dx]_{x=x_0} = -1/p(x_0), & n = 1, \\ \lim_{r\to 0} (2\pi p r \partial u/\partial r) = -1, & n = 2, \\ \lim_{r\to 0} \left(4\pi p r^2 \partial u/\partial r\right) = -1, & n = 3. \end{cases} \tag{6.7.21}$$

When $n = 1$, the bracket represents the jump in $u'(x)$ at $x = x_0$, which is given as $u'(x_0+) - u'(x_0-)$. For $n = 2$ we have $r^2 = (x - x_0)^2 + (y - y_0)^2$, and for $n = 3$, $r^2 = (x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2$.

Away from the source point P_0, the fundamental solution u satisfies the homogeneous equation

$$-\nabla \cdot (p\nabla u) + qu = 0, \qquad P \neq P_0, \tag{6.7.22}$$

where ∇ is replaced by d/dx in the one-dimensional case. For the one-dimensional problem, the source condition (6.7.21) can be treated as a matching condition for two solutions of (6.7.22) in the intervals $x < x_0$ and $x > x_0$. This is not so in the higher-dimensional problems. We also see that the conditions (6.7.21) are insufficient to determine solutions of (6.7.22) uniquely.

Fundamental solutions can be found by looking for singular solutions of (6.7.22) that have the behavior (6.7.21) at the source point. This approach is used when we obtain fundamental solutions of PDEs with constant coefficients in Example 6.13. Even if the solutions do not satisfy the auxiliary conditions placed on the problem, they are important because we can add smooth solutions of (6.7.22) to the results to account for the additional conditions. (This method is used in Chapter 7 to construct Green's functions.)

To examine the form of the singularity at the source point P_0, we consider (6.7.21) to represent approximate equations near P_0, just as was done in obtaining (6.7.4) and (6.7.5). Solving these (approximate) equations yields

$$u \approx \begin{cases} -|x - x_0|/2p(x_0), & n = 1, \\ -\log(r)/2\pi p(x_0, y_0), & n = 2, \\ 1/4\pi p(x_0, y_0, z_0) r, & n = 3, \end{cases} \tag{6.7.23}$$

which is valid for points P near P_0 and where r is defined as in (6.7.21) for $n = 2$ and $n = 3$. The one-dimensional form of $u(x)$ in (6.7.23) was chosen to make the solution symmetric with respect to x_0. In all cases only the most singular terms of

the solution were retained. We note that the singularity of u is localized at the source point P_0 in all cases, and that the strength of the singularity increases as the dimension n increases.

Example 6.13. Fundamental Solutions for Elliptic Equations with Constant Coefficients. We consider the elliptic equation

$$p\nabla^2 u - qu = 0, \tag{6.7.24}$$

where $p > 0$ and q are constants. Guided by (6.7.23), we look for a fundamental solution of (6.7.30) that depends on the distance r from the source point P_0. In one dimension (6.7.24) is an ODE and $r = |x - x_0|$. In two and three dimensions r is the radial variable in polar and spherical coordinates, respectively, with the pole at P_0.

Let $u = \tilde{u}(r)$ in (6.7.24) and we obtain the ordinary differential equation

$$p\left[\tilde{u}''(r) + \frac{n-1}{r}\,\tilde{u}'(r)\right] - q\,\tilde{u}(r) = 0, \qquad r > 0, \quad n = 1, 2, 3, \tag{6.7.25}$$

with n as the number of dimensions in the problem. For $n = 2$ and $n = 3$ we assume that the fundamental solution is independent of the angular variables.

To begin, we set $q = 0$ in (6.7.24), which reduces it to *Laplace's equation* when $n = 2$ and $n = 3$. For each n we require a singular solution of (6.7.25), and a set of such solutions is easily found to be given by

$$u = \tilde{u}(r) = \begin{cases} c|x - x_0|, & n = 1, \\ c\log\sqrt{(x-x_0)^2 + (y-y_0)^2}, & n = 2, \\ c[(x-x_0)^2 + (y-y_0)^2 + (z-z_0)^2]^{-1/2}, & n = 3, \end{cases} \tag{6.7.26}$$

with the constant c to be determined. On using the conditions (6.7.21) at P_0 or, more simply, by comparing (6.7.26) with (6.7.23), we find that

$$c = \begin{cases} -1/2p, & n = 1, \\ -1/2\pi p, & n = 2, \\ 1/4\pi p, & n = 3. \end{cases} \tag{6.7.27}$$

Other solutions of (6.7.25) exist that also satisfy the conditions (6.7.21) at P_0. However, the fundamental solutions (6.7.26)–(6.7.27) are in the standard form associated with (6.7.24) when $q = 0$, particularly when $n = 2$ or $n = 3$, in which case it reduces to Laplace's equation.

If $q \neq 0$ in (6.7.24)–(6.7.25) and $n = 1$, we obtain

$$u = \tilde{u}(r) = \begin{cases} c_1 \exp(\sqrt{q/p}\, r) + c_2 \exp(-\sqrt{q/p}\, r), & q > 0, \\ c_1 \exp(i\sqrt{-q/p}\, r) + c_2 \exp(-i\sqrt{-q/p}\, r), & q < 0, \end{cases} \tag{6.7.28}$$

with the constants c_1 and c_2 to be determined. This is done so that we obtain the *free space Green's functions* for these equations by specifying the behavior of the fundamental solutions at infinity, as well as at the source point.

For $q > 0$ we require that (6.7.28) tend to zero at infinity—that is, as $|x - x_0| \to \infty$—and that it satisfy (6.7.21) at x_0. Since $r = |x - x_0|$ in (6.7.28), we conclude that

$$u(x) = \frac{1}{2}(pq)^{-1/2} \exp\left(-\sqrt{q/p}\,|x - x_0|\right), \qquad q > 0, \tag{6.7.29}$$

is the (uniquely determined) *free space Green's function* for this case.

With $q < 0$ and $n = 1$, (6.7.24) has the form of the *Helmholtz* or *reduced wave equation* in one dimension. Given the wave equation $v_{tt}(x,t) = pv_{xx}(x,t)$, if we set $v(x,t) = u(x)\exp(-i\sqrt{-q}\,t)$ with $q < 0$, we find that $u(x)$ satisfies the one-dimensional form of (6.7.24). The *free space Green's function* for the Helmholtz equation is a fundamental solution that represents a wave traveling away from the source point. Combined with the source point condition (6.7.21), it implies that (6.7.28) has the form

$$u(x) = \frac{i}{2}(-pq)^{-1/2} \exp\left(i\sqrt{-q/p}\,|x - x_0|\right), \qquad q < 0. \tag{6.7.30}$$

When multiplied by $\exp(-i\sqrt{-q}\,t)$, (6.7.30) is a plane wave traveling away from the source point $x = x_0$ with speed $\sqrt{p}$.

In both of the foregoing problems, a uniquely determined free space Green's function was obtained, because the presence of exponentially growing solutions at infinity when $q > 0$ and the existence of waves traveling from infinity to the source point when $q < 0$ is physically unreasonable. (A physical interpretation of the case $q > 0$ is given below.)

With $n = 2$ and $q < 0$, (6.7.25) has the form of *Bessel's equation of zero order*. We require a singular solution of this equation and

$$u = \tilde{u}(r) = c_0 Y_0\left[\sqrt{-q/p}\sqrt{(x - x_0)^2 + (y - y_0)^2}\right], \qquad q < 0, \tag{6.7.31}$$

where c_0 is a constant and $Y_0(z)$ is the *Neumann function of zero order*, is singular at $r = 0$. With $n = 2$ and $q > 0$ in (6.7.25) we obtain the *modified Bessel equation of zero order*, and a singular solution is

$$u = \tilde{u}(r) = c_1 K_0\left[\sqrt{q/p}\sqrt{(x - x_0)^2 + (y - y_0)^2}\right], \qquad q > 0, \tag{6.7.32}$$

where c_1 is a constant and $K_0(z)$ is the *modified zero-order Bessel function of the second kind*. For $z \approx 0$ these functions have the following behavior: $Y_0(z) \approx 2/\pi\,\log(z)$, $K_0(z) \approx -\log(z)$, $z \approx 0$. Comparing (6.7.31)–(6.7.32) with (6.7.23) and using the above, we conclude that $c_0 = -1/4p$, $c_1 = 1/2\pi p$.

With $n = 3$ and $q \neq 0$ in (6.7.24), we easily obtain the solutions

$$u(x,y,z) = \begin{cases} (1/4\pi pr)\exp(-\sqrt{q/p}\,r), & q > 0, \\ (1/4\pi pr)\exp(i\sqrt{-q/p}\,r), & q < 0, \end{cases} \tag{6.7.33}$$

where $r^2 = (x-x_0)^2+(y-y_0)^2+(z-z_0)^2$. As $r \to 0$ these fundamental solutions have the required behavior (6.7.23).

The foregoing two- and three-dimensional fundamental solutions were specified solely on the basis of their behavior at the source point. To determine the *free space Green's functions* for these problems, we require that the fundamental solutions have a prescribed behavior at infinity. If $q > 0$, (6.7.24) represents a stationary form of the diffusion equation $v_t + qv = p\nabla^2 v$ that is obtained on setting $v(\mathbf{x}, t) \equiv u(\mathbf{x})$. We require that $v(\mathbf{x}, t)$ and, consequently, $u(\mathbf{x})$ tend to zero as the distance r from the source point P_0 tends to infinity. If $q < 0$, (6.7.24) is again the reduced wave equation or Helmholtz's equation. If $v(\mathbf{x}, t)$ is a solution of the wave equation $v_{tt}+qv = p\nabla^2 v$, we require that $v(\mathbf{x}, t) = u(\mathbf{x}) \exp(-i\sqrt{-q}\, t)$ represent a diverging wave traveling away from the source point P_0 as $r \to \infty$.

To apply the conditions at infinity we must know the behavior of the Bessel functions Y_0 and K_0 for large values of the argument. We have

$$Y_0(z) \approx [2/\pi z]^{1/2} \sin(z - \pi/4),\ K_0(z) \approx [\pi/2z]^{1/2}\, e^{-z}, \quad z \to \infty. \quad (6.7.34)$$

We conclude that for $q > 0$ the fundamental solutions (6.7.32) and (6.7.33) both decay exponentially as $r \to \infty$ and are, therefore, the *free space Green's functions* in two and three dimensions for this case.

When $n = 3$ and $q < 0$, the fundamental solution (6.7.33) represents a diverging spherical wave traveling away from the source point P_0 with speed $\sqrt{p}$. Thus it is the *free space Green's function* for this problem. However, since $\sin z = (1/2i)[e^{iz} - e^{-iz}]$, the fundamental solution (6.7.31) with $n = 2$ contains both diverging and converging cylindrical waves for large r, in view of (6.7.34). Thus (6.7.31) is not the free space Green's function for this problem.

The correct form for the *free space Green's function* in two dimensions is given by

$$u(x, y) = (i/4p)H_0^{(1)}\left[\sqrt{-q/p}\sqrt{(x-x_0)^2+(y-y_0)^2}\right], \qquad q < 0, \quad (6.7.35)$$

where $H_0^{(1)}(z)$ is the *zero-order Hankel function of the first kind*. $H_0^{(1)}(z)$ is a singular solution of the zero-order Bessel equation with the following behavior:

$$H_0^{(1)}(z) \approx (2i/\pi)\log(z),\ z \approx 0, \quad H_0^{(1)}(z) \approx [2/\pi z]^{1/2} \exp[i(z-\pi/4)],\ z \to \infty. \quad (6.7.36)$$

Consequently, (6.7.35) exhibits the correct behavior (6.7.23) at the source point and represents a diverging cylindrical wave as the distance r from the source point tends to infinity.

Fundamental Solutions of Hyperbolic Equations

We now turn to the consideration of fundamental solutions for hyperbolic equations. Because of the aforementioned complications with the interpretations of these solutions and the determination of their behavior at the source point, we concentrate mostly on one-dimensional problems for equations with constant coefficients, but some higher-dimensional results are also presented.

The solution of the *instantaneous point source problem* for *hyperbolic equations* can be determined from the solution of the stationary source problem discussed previously, by specializing the result to the case where source acts instantaneously at the time $t = t_0$. We show how this can be done for the wave equation, for which the stationary point source problem was solved in Example 6.11. Let $f(t) = \delta(t - t_0)$, the Dirac delta function, with $t_0 > 0$. Since the delta function vanishes for $t \neq t_0$, the stationary source problem of Example 6.11 is formally converted into an instantaneous source problem. The solution of the instantaneous source problem is obtained by replacing $f(t)$ by the Dirac delta function in each of the solution formulas obtained in Example 6.11.

We obtain the following results. With $n = 3$, (6.7.8) yields

$$u(r,t) = \frac{1}{4\pi c^2 r}\,\delta\left(t - t_0 - \frac{r}{c}\right). \tag{6.7.37}$$

In two dimensions, (6.7.9) yields

$$u(\hat{r},t) = \frac{1}{2\pi c}\,\frac{H[c(t-t_0)-\hat{r}]}{\sqrt{c^2(t-t_0)^2 - \hat{r}^2}}. \tag{6.7.38}$$

In the one-dimensional case, (6.7.10) yields

$$u(x,t) = \frac{1}{2c}\,H[c(t-t_0) - |x - x_0|]. \tag{6.7.39}$$

In (6.7.38) and (6.7.39), $H(z)$ is the Heaviside function. Its presence signifies that the solutions vanish outside the forward characteristic cone or sector with vertex at P_0 and t_0. The causal fundamental solution (6.7.37) vanishes everywhere except on the forward characteristic cone. Each of these solutions is singular not only at the vertex of the characteristic cone, but along the entire forward characteristic cone. This shows why they must be interpreted in the weak sense. (They are rederived and discussed in Section 7.4.)

We would like to determine the fundamental solutions directly, independent of having to solve more general problems, as has been done above. Although such direct determinations of fundamental solutions are given in Chapter 7, using the theory of generalized functions and transform methods, we now present an alternative method for obtaining fundamental solutions for hyperbolic equations with constant coefficients.

Example 6.14. Fundamental Solutions for Hyperbolic Equations with Constant Coefficients. We consider the hyperbolic equation

$$\rho u_{tt} - p\nabla^2 u + qu = 0, \tag{6.7.40}$$

with constant coefficients and determine fundamental solutions of (6.7.40) in a manner similar to that given for the elliptic equations of Example 6.13. Let $\hat{P}_0$ be the source point in one, two, or three (spatial) dimensions and t_0 the time when the source acts. In space-time, the instantaneous source point is $P_0 = (\hat{P}_0, t_0)$. With $r = |\mathbf{x} - \mathbf{x}_0|$ as the distance from $\hat{P}_0$ (in one dimension $r = |x - x_0|$), we define the *hyperbolic distance* γ as $\gamma = \sqrt{(t-t_0)^2 - r^2/c^2}$, where $c^2 = p/\rho$. In a hyperbolic (non-Euclidean) geometry appropriate for the hyperbolic equation (6.7.40), γ represents distance from the point $(\hat{P}_0, t_0)$. It is positive when $r^2 < c^2(t-t_0)^2$ and is imaginary when $r^2 > c^2(t-t_0)^2$. It vanishes when $r^2 = c^2(t-t_0)^2$, and this is the equation of the *characteristic cone* for (6.7.40) with vertex at $(\hat{P}_0, t_0)$. [For $n = 1$ this yields the two characteristic lines through the point (x_0, t_0).] As shown for the wave equation, fundamental solutions for (6.7.40) are expected to vanish outside the characteristic cone, so this distance function is appropriate for our problem.

Proceeding as in Example 6.13, we look for a singular solution of (6.7.40) in the form $u = u(\gamma)$. This gives the ordinary differential equation

$$\rho\left[u''(\gamma) + \frac{n}{\gamma}\,u'(\gamma)\right] + q\,u(\gamma) = 0, \qquad n = 1, 2, 3, \tag{6.7.41}$$

of the general form of (6.7.26). To avoid using the complex-valued variable γ when considering the fundamental solution outside the characteristic cone, we introduce the squared hyperbolic distance $\sigma = \gamma^2$. Then $\sigma > 0$ within the cone and $\sigma < 0$ outside the cone. With $u = \hat{u}(\sigma)$ in (6.7.41), we obtain

$$\rho\left[\sigma\,\hat{u}''(\sigma) + \frac{n+1}{2}\,\hat{u}(\sigma)\right] + \frac{q}{4}\,\hat{u}(\sigma) = 0, \qquad n = 1, 2, 3. \tag{6.7.42}$$

First we set $q = 0$ and $n = 1$ in (6.7.42) to obtain $\sigma\hat{u}''(\sigma) + \hat{u}'(\sigma) = 0$. [With $n = 1$ and $q = 0$, (6.7.41) and (6.7.42) have the same form.] The general solution of this equation is $\hat{u}(\sigma) = a + b\log(\sigma)$, with a and b as arbitrary constants. While the solution $\hat{u}(\sigma) = b\log(\sigma)$ is singular on the characteristic cone, it differs from the fundamental solution (6.7.39) obtained previously. The second (independent) solution $\hat{u}(\sigma) = a$ is not singular at all, but it is constant within the (forward) characteristic cone as is the case for (6.7.39). If we formally put $\hat{u}(\sigma) = 0$ outside the characteristic cone, the solution has a jump discontinuity across the cone and is no longer a classical solution of the equation for $\hat{u}(\sigma)$. It must be interpreted as a *generalized solution.*

To show that such a solution is possible, we require some elementary concepts from the *theory of generalized functions*, whose properties are discussed fully in Section 7.2. If $H(x)$ is the Heaviside function, then $H'(x) = 0$ for all $x \neq 0$, but the derivative is not defined at $x = 0$. However, if we assume that $H'(x)$ can be defined at $x = 0$ as well (this will be done in Section 7.2), we have $\int_{-\infty}^{\infty} H'(x)\,dx = 1$,

since $H(\infty) = 1$ and $H(-\infty) = 0$. Noting the results of Example 1.2, we put $H'(x) = \delta(x)$, the Dirac delta function. Additionally, since $f(x)\delta(x) = f(0)\delta(x)$ (see Example 1.2), we find that $x\delta(x) = 0$. Formally differentiating this equation gives $x\,\delta'(x) + \delta(x) = 0$. Consequently, $\hat{u}(\sigma) = aH(\sigma)$ is a generalized solution of $\sigma\hat{u}''(\sigma) + \hat{u}'(\sigma) = 0$. It vanishes outside the characteristic cone where $\sigma < 0$, and $\hat{u}(\sigma) = a$ within the cone where $\sigma > 0$.

For u to be a fundamental solution of (6.7.40) we must specify the value of a. If u is to be a causal fundamental solution, it must vanish for $t < t_0$ as well. We cannot obtain a causal fundamental solution strictly in terms of the variable σ, which can distinguish only between the interior or the exterior of the full characteristic cone, not between the forward and backward parts of the cone. Therefore, we set $u(\sigma) = aH(\sigma)$ for $t > t_0$ and $u = 0$ for $t < t_0$.

In effect, we define the causal fundamental solution of (6.7.40) (with $q = 0$) in the one-dimensional case to be $u = a$ for $|x - x_0| < c(t - t_0)$ and $u = 0$ for $|x - x_0| > c(t - t_0)$. It follows from the discussion in Example 6.4 that this is a *weak solution* of the differential equation for $(x, t) \neq (x_0, t_0)$. We now determine the constant a so that the source point condition (6.7.20) is satisfied.

Let the point (x_0, t_0) lie in the interior of the rectangular region R [of (6.7.20)] whose boundary is parallel to the coordinate axes. The function $v(x, t)$ in (6.7.20) is assumed to vanish on and near ∂R. Noting the definition of $u(x, t)$, the integral (6.7.20) reduces to an integral over the region A, whose boundary comprises a portion of the rectangle and the two characteristics $|x - x_0| = c(t - t_0)$. Thus,

$$\begin{aligned} \iint_R u[\rho v_{tt} - p v_{xx}]\, dV &= \iint_A u[\rho v_{tt} - p v_{xx}]\, dV \\ = a \iint_A [\rho v_{tt} - p v_{xx}]\, dV &= -a \int_{\partial A} \rho v_t\, dx + p v_x\, dt \\ = -a\sqrt{p\rho} \int_{P_0}^{P_1} dv + a\sqrt{p\rho} \int_{P_2}^{P_0} dv &= 2a\sqrt{p\rho} v(P_0) = v(P_0), \end{aligned} \tag{6.7.43}$$

where $P_0 = (x_0, t_0)$, P_1 and P_2 are the points where the characteristics intersect the rectangle and the techniques used in Example 6.4 have been applied. We find that $a = 1/2\sqrt{p\rho} = 1/2c\rho$. Consequently, the causal fundamental solution of the one-dimensional version of (6.7.40) where we put $q = 0$ is given as

$$u(x, t) = \begin{cases} 1/2c\rho, & |x - x_0| < c(t - t_0), \\ 0, & |x - x_0| > c(t - t_0). \end{cases} \tag{6.7.44}$$

This solution coincides with that obtained in (6.7.39) if we put $\rho = 1$ and $c^2 = p$.

If $n = 1$ and $q > 0$, (6.7.41) has the form of Bessel's equation of zero order. The nonsingular solution is $u = aJ_0[(q/\rho)^{1/2}\gamma]$, where a is an arbitrary constant and $J_0(z)$ is the zero-order Bessel function. To obtain a causal fundamental solution for (6.7.40) in this case, we put $a = 1/2c\rho$ and set $u = 0$ outside the forward

characteristic cone. This gives

$$u(x,t) = \begin{cases} (1/2c\rho)J_0\left[\sqrt{q/p}\sqrt{c^2(t-t_0)^2-(x-x_0)^2}\right], & |x-x_0| < c(t-t_0), \\ 0, & |x-x_0| > c(t-t_0). \end{cases} \tag{6.7.45}$$

If $n = 1$ and $q < 0$ in (6.7.41), we have a modified Bessel equation of zero order. With $I_0(z)$ as the modified Bessel function of zero order, the causal fundamental solution of (6.7.40) for this case is

$$u(x,t) = \begin{cases} (1/2c\rho)I_0\left[\sqrt{-q/p}\sqrt{c^2(t-t_0)^2-(x-x_0)^2}\right], & |x-x_0| < c(t-t_0), \\ 0, & |x-x_0| > c(t-t_0). \end{cases} \tag{6.7.46}$$

The verification that (6.7.45) and (6.7.46) are weak solutions of (6.7.40) is somewhat more complicated than that given above for the case when $q = 0$ and is not be given here. We note that when $q = 0$, both (6.7.45) and (6.7.46) reduce to (6.7.44). The derivation of causal fundamental solutions of (6.7.40) in higher dimensions by the foregoing method is not presented here. It requires a greater familiarity with the properties of generalized functions than was needed for the one-dimensional case. These solutions are considered in Section 7.4.

Fundamental Solutions of Parabolic Equations

Finally, we consider fundamental solutions of parabolic equations. Because of the aforementioned complications with the interpretations of these solutions and the determination of their behavior at the source point, we concentrate mostly on one-dimensional problems for equations with constant coefficients. Some results for higher-dimensional problems are also presented.

The solution of the instantaneous point source problem for parabolic equations can be determined from the solution of the stationary point source problem, by specializing the result to the case where the source acts instantaneously at the time $t = t_0$. We carry this out for the *heat equation*, for which the stationary point source problem was solved in Example 6.12. Let $f(t) = \delta(t - t_0)$, the Dirac delta function, with $t_0 > 0$. Since the delta function vanishes for $t \neq t_0$, the stationary point source problem of Example 6.12 is formally converted into an instantaneous point source problem. The solution of the instantaneous point source problem is obtained by replacing $f(t)$ by the Dirac delta function in each of the solution formulas obtained in Example 6.12.

For the heat equation, we find from Example 6.12 that the *causal fundamental solution* has the form

$$u(r,t) = G(r, t - t_0)H(t - t_0), \tag{6.7.47}$$

where r is the (spatial) distance from the source point P_0, $H(z)$ is the Heaviside function and $G(r, t - t_0)$ is the fundamental solution of the heat equation in one, two,

or three dimensions. The Heaviside function is required in (6.7.47) to make $u(r,t)$ vanish for $t < t_0$. However, as shown previously, since the fundamental solution G tends to zero at an exponential rate as $t \downarrow t_0$, $u(r,t)$ is a smooth function across $t = t_0$, for all points $P \neq P_0$. Therefore, the integral condition (6.7.18) can be used to determine the behavior of the solution at the singular point $(r,t) = (0,t_0)$. Its use is demonstrated below for the one-dimensional problem.

As was done for the point source problems treated above for elliptic and hyperbolic equations, we want to find fundamental solutions for parabolic equations directly, independent of having to solve more general problems. Although such direct determinations of fundamental solutions are given in Chapter 7 using the theory of generalized functions and transform methods, we now present an alternative method for obtaining fundamental solutions for parabolic equations with constant coefficients.

The causal fundamental solution for the heat equation in one, two or three dimensions was given in (6.7.47). By looking for a solution of the heat equation that depends on a *similarity variable* (as defined in the exercises), the heat equation is reduced to an ODE and the fundamental solution can then be constructed in terms of its solutions. This approach parallels that given in Examples 6.13 and 6.14 for elliptic and hyperbolic equations and is considered in the exercises.

In Example 6.15 a different approach is used, and the *causal fundamental solution* for the one-dimensional *heat equation* is obtained as a limit of the causal fundamental solution for the *telegrapher's equation*. This is of interest in connection with *Brownian motion*, as discussed in Sections 1.1 and 1.2.

Example 6.15. The Fundamental Solution of the Heat Equation in One Dimension. The (hyperbolic) *telegrapher's equation*

$$\epsilon\, u_{tt}(x,t) - c^2 u_{xx}(x,t) + u_t(x,t) = 0, \tag{6.7.48}$$

where $\epsilon > 0$, formally reduces to the *heat equation* if we set $\epsilon = 0$. We might expect that as $\epsilon \to 0$, the *causal fundamental solution* of (6.7.48) corresponding to a source point at $P_0 = (x_0,t_0)$ reduces to the causal fundamental solution of the heat equation. Indeed, the causality condition for both problems requires that $u(x,t) = 0$ for $t < t_0$. Furthermore, the causal fundamental solution of (6.7.48) vanishes outside the forward characteristic sector $|x - x_0| < \hat{c}(t - t_0)$ where $\hat{c} = c/\sqrt{\epsilon}$. As $\epsilon \to 0$, the boundary $\sqrt{\epsilon}|x - x_0| = c(t - t_0)$ of this sector tends to the line $t = t_0$, which is a characteristic line for the heat equation. In the limit, therefore, the fundamental solution of (6.7.48) is expected to be nonzero in the region $t > t_0$, a result appropriate for the heat equation.

Let

$$u(x,t) = \exp\left[-(t - t_0)/2\epsilon\right] v(x,t), \tag{6.7.49}$$

in (6.7.48) and we obtain

$$\epsilon\, v_{tt}(x,t) - c^2 v_{xx}(x,t) - \frac{1}{4\epsilon} v(x,t) = 0. \tag{6.7.50}$$

At the source point $P_0 = (x_0,t_0)$, (6.7.49) implies that $u(x_0,t_0) = v(x_0,t_0)$, so that the fundamental solution of (6.7.50) should yield the fundamental solution of (6.7.48)

by means of (6.7.49). With $\rho = \epsilon$, $p = c^2$, and $q = -1/4\epsilon$, we obtain the causal fundamental solution of (6.7.48) from that of (6.7.40) in the form

$$u(x,t) = \begin{cases} \dfrac{e^{-(t-t_0)/2\epsilon}}{\sqrt{4c^2\epsilon}} I_0\left[\dfrac{\sqrt{c^2(t-t_0)^2/\epsilon - (x-x_0)^2}}{\sqrt{4c^2\epsilon}}\right], & |x-x_0| < c(t-t_0)/\sqrt{\epsilon}, \\ 0, & |x-x_0| > c(t-t_0)/\sqrt{\epsilon}, \end{cases} \tag{6.7.51}$$

where (6.7.49) was used in connection with the result (6.7.46). As $z \to \infty$, the modified Bessel function $I_0(z)$ has the asymptotic behavior $I_0(z) \approx e^z/\sqrt{2\pi z}$. Also, for $t > t_0$ and $0 < \epsilon \ll 1$ we have

$$\sqrt{\frac{c^2(t-t_0)^2}{\epsilon} - (x-x_0)^2} = \frac{c(t-t_0)}{\sqrt{\epsilon}}\left[1 - \frac{\epsilon}{2c^2}\frac{(x-x_0)^2}{t-t_0} + O(\epsilon^2)\right], \tag{6.7.52}$$

which follows from the binomial expansion.

As $\epsilon \to 0$ the argument of I_0 in (6.7.51) can be given as $(t-t_0)/2\epsilon - (x-x_0)^2/(4c^2(t-t_0)) + O(\epsilon)$, and it tends to infinity. Then, we easily obtain in the limit as $\epsilon \to 0$,

$$u(x,t) = \begin{cases} \dfrac{1}{\sqrt{4\pi c^2(t-t_0)}} \exp\left[-\dfrac{(x-x_0)^2}{4c^2(t-t_0)}\right], & t > t_0, \\ 0, & t < t_0, \end{cases} \tag{6.7.53}$$

which is identical to the causal fundamental solution of the heat equation given in (6.7.47).

Exercises 6.7

6.7.1. Use (6.7.21) to determine that the constants c in (6.7.26) are given as in (6.7.27).

6.7.2. Derive the result (6.7.10).

6.7.3. Derive the result (6.7.14).

6.7.4. Show that the results (6.7.37)–(6.7.39) follow from the corresponding results for a continuous point source.

6.7.5. Derive the equations (6.7.42) and (6.7.43).

6.7.6. Show that if $u(x,t)$ satisfies the heat equation $u_t(x,t) = c^2 u_{xx}(x,t)$, so does $u(ax, a^2t)$, where a is a constant. Put $a = 1/\sqrt{t}$ and show that the heat equation has a *similarity solution* of the form $u(x,t) = F(x/\sqrt{t})$, where $F(z)$ satisfies the equation $F''(z) + (z/2c^2)F'(z) = 0$, and that $F'(z) = \hat{a}\exp(-z^2/4c^2)$. Noting that $u_x(x,t)$ is a solution of the heat equation if $u(x,t)$ is a solution, conclude that $u_x(x,t) = \partial/\partial x\left[F\left(x/\sqrt{t}\right)\right] = (1/\sqrt{t})F'\left(x/\sqrt{t}\right) = (\hat{a}/\sqrt{t})\exp(-x^2/4c^2t)$ (with $\hat{a} = 1/\sqrt{4\pi c^2}$ this is the fundamental solution of the heat equation).

6.7.7. Let $r^2 = x^2 + y^2$ and $u(r,t)$ be a solution of the heat equation in two dimensions. Show that $u(ar, a^2t)$ is also a solution and construct a *similarity solution* of the form $u(r,t) = F(r/\sqrt{t})$. Noting that $u_t(r,t)$ is also a solution of the heat equation if $u(r,t)$ is a solution, conclude that $u_t(r,t) = \partial/\partial t\left[F\left(r/\sqrt{t}\right)\right] = (\hat{a}/t)\exp(-r^2/4c^2t)$ satisfies the heat equation $u_t = c^2\nabla^2 u$ (with $\hat{a} = 1/4\pi c^2$ this is the fundamental solution of the heat equation).

6.7.8. Let $r^2 = x^2+y^2+z^2$ and show that the heat equation $u_t(\mathbf{x},t) = c^2\nabla^2 u(\mathbf{x},t)$ in three dimensions has a solution of the form $u(r,t) = F(r/\sqrt{t})/r$, where $ru(r,t) = v(r,t) = F(r/\sqrt{t})$ is a *similarity solution* of $v_t(r,t) = c^2 v_{rr}(r,t)$. Show that $u_t(r,t) = \partial/\partial t\left[(1/r)F\left(r/\sqrt{t}\right)\right] = (\hat{a}/t^{3/2})\exp(-r^2/4c^2t)$ satisfies the heat equation [with $\hat{a} = (1/\sqrt{4\pi c^2})^3$ this is the fundamental solution of the heat equation in three dimensions].

6.7.9. Obtain the solution of the (continuous) point source problem for the two-dimensional heat equation discussed in the text by constructing a superposition of appropriately modified fundamental solutions as given in Exercise 6.7.7.

6.7.10. Proceeding as in Exercise 6.7.9 and using Exercise 6.7.8, solve the (continuous) point source problem for the three-dimensional heat equation.

6.8 ENERGY INTEGRALS

For the most part, we have been concerned with the construction of solutions of partial differential equations in this and the preceding chapters, but not with the question of uniqueness and continuous dependence on the data. As was stated in Section 3.4, *existence, uniqueness*, and *continuous dependence on data* are required if an initial and/or boundary value problem is to be *well posed.* Using the concept of *energy integrals*, we now discuss briefly the uniqueness and continuous dependence on the data of the solutions of the second order differential equations we have considered in the last three chapters. (See Section 4.1 for a description of the boundary conditions.)

Energy Integrals for Hyperbolic Equations

We begin with the hyperbolic equation

$$\rho(\mathbf{x})u_{tt}(\mathbf{x},t) - \nabla\cdot(p(\mathbf{x})\nabla u(\mathbf{x},t)) + q(\mathbf{x})u(\mathbf{x},t) = \rho F(\mathbf{x},t),\ \mathbf{x}\in G,\ t>0, \quad (6.8.1)$$

where G is a region in $\mathbf{x}$-space with the boundary ∂G, $p(\mathbf{x})$ and $\rho(\mathbf{x})$ are positive, and $q(\mathbf{x})$ is nonnegative in G. Initial and boundary data are given as in (4.2.4) and (4.1.11). The coefficients and the data for (6.8.1) are assumed to be such that *classical solutions* can be constructed. We show that these solutions are unique and depend continuously on the data.

To construct an *energy integral* for (6.8.1) we multiply across by u_t and integrate over G. Since $u_t(\mathbf{x},t)\nabla\cdot(p(\mathbf{x})\nabla u(\mathbf{x},t)) = \nabla\cdot(p(\mathbf{x})u_t(\mathbf{x},t)\nabla u(\mathbf{x},t)) - p(\mathbf{x})\nabla u(\mathbf{x},t)\cdot\nabla u_t(\mathbf{x},t)$, we obtain

$$\iint_G u_t[\rho u_{tt} - \nabla\cdot(p\nabla u) + qu]\,dv = \frac{\partial}{\partial t}\iint_G \left[\frac{1}{2}\rho u_t^2 + \frac{1}{2}p(\nabla u)^2 + \frac{1}{2}qu^2\right] dv$$

$$- \int_{\partial G} p u_t \frac{\partial u}{\partial n}\,ds = \iint_G \rho F u_t\,dv, \tag{6.8.2}$$

where we have used the divergence theorem and the fact that the region G and the coefficients in (6.8.1) are time independent.

The *energy* of the system in the absence of external effects is

$$E(t) = \frac{1}{2}\iint_G \left[\rho u_t^2 + p(\nabla u)^2 + qu^2\right]\,dv + \frac{1}{2}\int_{S_3} p\,\frac{\alpha}{\beta}\,u^2\,ds, \tag{6.8.3}$$

where the surface integral term comes from substituting for $\partial u/\partial n$ in (6.8.2) on using the boundary condition. For the *wave equation* with homogeneous boundary data, the energy represents the sum of *kinetic* and *potential energies* in the case of a *vibrating string*. [We note that based on the assumptions on $p(\mathbf{x})$, $\rho(\mathbf{x})$, and $q(\mathbf{x})$, as well as $\alpha(\mathbf{x})$ and $\beta(\mathbf{x})$, all of which are nonnegative, each of the integrals in (6.8.3) is nonnegative.]

Thus on using (6.8.3) in (6.8.2), we have

$$E'(t) = \iint_G \rho F u_t\,dv + \int_{S_1}\frac{p}{\alpha}\,B_t\,\frac{\partial u}{\partial n}\,ds + \int_{S_2\cup S_3}\frac{p}{\beta}\,Bu_t\,ds. \tag{6.8.4}$$

The terms on the right in (6.8.4) represent changes in the energy of the system due to internal and boundary forces.

We use the energy integral to prove that the *initial and boundary value problem* for (6.8.1) has a unique solution. Suppose that there are two solutions $u_1(\mathbf{x},t)$ and $u_2(\mathbf{x},t)$ of the initial and boundary value problem for (6.8.1). Since the problem is linear, the difference $u(\mathbf{x},t) = u_1(\mathbf{x},t) - u_2(\mathbf{x},t)$ must be a solution of the homogeneous form of (6.8.1), that is, with $F = 0$ and with zero initial and boundary data. Since $F = B = 0$, we find from (6.8.4) that $E'(t) = 0$. Thus $E(t)$ = constant and energy is conserved. Furthermore,

$$E(t) = E(0) = \frac{1}{2}\iint_G \left[\rho(\mathbf{x})u_t^2(\mathbf{x},0) + p(\mathbf{x})(\nabla u(\mathbf{x},0))^2 + q(\mathbf{x})u^2(\mathbf{x},0)\right]\,dv$$

$$+\frac{1}{2}\int_{S_3} p(\mathbf{x})\,\frac{\alpha(\mathbf{x})}{\beta(\mathbf{x})}\,u^2(\mathbf{x},0)\,ds = 0, \tag{6.8.5}$$

since the initial data $u(\mathbf{x},0)$ and $u_t(\mathbf{x},0)$ vanish. [Note that $E(t)$ = constant implies that $E(t) = E(0)$.] Then for any $t > 0$ we have

$$E(t) = \frac{1}{2}\iint_G \left[\rho(\mathbf{x})u_t^2(\mathbf{x},t) + p(\mathbf{x})(\nabla u(\mathbf{x},t))^2 + q(\mathbf{x})u^2(\mathbf{x},t)\right]\, dv$$

$$+\frac{1}{2}\int_{S_3} p(\mathbf{x})\,\frac{\alpha(\mathbf{x})}{\beta(\mathbf{x})}\,u^2(\mathbf{x},t)\,ds = 0. \tag{6.8.6}$$

Since $\rho(\mathbf{x})$, $p(\mathbf{x})$ are positive and $q(\mathbf{x})$ is nonnegative in G, and $\alpha(\mathbf{x})$ and $\beta(\mathbf{x})$ are positive on S_3, both integrands in (6.8.6) are nonnegative and we conclude that $u_t(\mathbf{x},t) = 0$ and $\nabla u(\mathbf{x},t) = \mathbf{0}$, for $\mathbf{x} \in G$. The vanishing of all the first derivatives of $u(\mathbf{x},t)$ implies that $u(\mathbf{x},t)$ = constant. But $u(\mathbf{x},0) = 0$, so that $u(\mathbf{x},t) = 0$ and $u_1(\mathbf{x},t) = u_2(\mathbf{x},t)$. Thus we have shown that the (classical) solutions of the initial and boundary value problem are *unique*. In the above we have admitted the possibility that $q(\mathbf{x}) = 0$ over a portion of G. However, if $q(\mathbf{x}) > 0$ in G, we conclude directly from (6.8.6) that $u(\mathbf{x},t) = 0$.

If the coefficient $q(\mathbf{x})$ in (6.8.1) is permitted to be negative, the foregoing argument fails to prove uniqueness, as we can no longer conclude from the vanishing of $E(t)$ in (6.8.6) that the integrands and, consequently, $u_t(\mathbf{x},t)$ and $\nabla u(\mathbf{x},t)$ vanish. To obtain uniqueness for this case we set

$$u(\mathbf{x},t) = e^{rt}v(\mathbf{x},t), \tag{6.8.7}$$

where r is a positive constant to be specified. Inserting (6.8.7) into (6.8.1) yields

$$\rho v_{tt} + 2r\rho v_t - \nabla\cdot(p\nabla v) + (r^2\rho + q)v = e^{-rt}\rho F. \tag{6.8.8}$$

We assume that $\rho(\mathbf{x})$ and $q(\mathbf{x})$ are uniformly bounded in the region G, with $|q(\mathbf{x})| < M$, where M is some constant. Then we choose r such that $r^2\rho(\mathbf{x}) > M$ [recall that $\rho(\mathbf{x}) > 0$], and this implies that $r^2\rho(\mathbf{x}) + q(\mathbf{x})$, the coefficient of $v(\mathbf{x},t)$ in (6.8.8), is positive.

We multiply across by $v_t(\mathbf{x},t)$ in (6.8.8) and integrate over G. Then

$$\frac{1}{2}\frac{\partial}{\partial t}\left[\iint_G \left[\rho v_t^2 + p(\nabla v)^2 + (r^2\rho + q)v^2\right] dV + \int_{S_3}\frac{p\alpha}{\beta}v^2 ds\right] + \iint_G 2r\rho v_t^2\, dV$$

$$= \iint_G e^{-rt}\rho F v_t\, dV + \int_{S_1}\frac{p}{\alpha}\,(e^{-rt}B)_t\,\frac{\partial v}{\partial n}\,ds + \int_{S_2\cup S_3}\frac{p}{\beta}\,e^{-rt}Bv_t\,ds. \tag{6.8.9}$$

Let $\tilde{E}(t)$ equal half the term in braces in (6.8.9). Assuming there are two solutions $u_1(\mathbf{x},t)$ and $u_2(\mathbf{x},t)$ of the initial and boundary value problem for (6.8.1), we consider the difference $u(\mathbf{x},t) = u_1(\mathbf{x},t) - u_2(\mathbf{x},t)$. This leads to a homogeneous initial and boundary value problem for the solution of the homogeneous form of (6.8.8). Then since $F = B = 0$ in (6.8.9), we have

$$\tilde{E}'(t) = -2r\iint_G \rho(\mathbf{x})v_t^2(\mathbf{x},t)\, dV \le 0, \tag{6.8.10}$$

inasmuch as $\rho(\mathbf{x})$ and r are positive. From (6.8.10) we determine that $\tilde{E}(t)$ is nonincreasing in time. But $\tilde{E}(0) = 0$ since $v(\mathbf{x},t) = v_t(\mathbf{x},t) = 0$ at the time $t = 0$. Also,

$\tilde{E}(t) \geq 0$ for all t since $\rho(\mathbf{x})$, $p(\mathbf{x})$, and $r^2\rho(\mathbf{x})+q(\mathbf{x})$ are positive in G and $\alpha(\mathbf{x})/\beta(\mathbf{x})$ is positive in S_3. Since $\tilde{E}(t)$ cannot increase from its value $\tilde{E}(0) = 0$ as t increases, we conclude that $\tilde{E}(t) = 0$. Again the nonnegativity of the integrals implies that $v(\mathbf{x}, t) = 0$, so that $v_1(\mathbf{x}, t) = v_2(\mathbf{x}, t)$ and, consequently, $u_1(\mathbf{x}, t) = u_2(\mathbf{x}, t)$ and *uniqueness* is proven.

The energy integral method can also be used to prove *continuous dependence on the data* for the solution of the initial and boundary value problem for (6.8.1). For simplicity we assume that $F(\mathbf{x}, t) = B(\mathbf{x}, t) = 0$. The case with nonzero $F(\mathbf{x}, t)$ and $B(\mathbf{x}, t)$ requires additional estimates of the given integrals in $F(\mathbf{x}, t)$ and $B(\mathbf{x}, t)$. Let $u(\mathbf{x}, 0) = f(\mathbf{x})$ and $u_t(\mathbf{x}, 0) = g(\mathbf{x})$. We assume that there are two sets of initial data $f_1(\mathbf{x})$, $f_2(\mathbf{x})$ and $g_1(\mathbf{x})$, $g_2(\mathbf{x})$ with corresponding solutions $u_1(\mathbf{x}, t)$ and $u_2(\mathbf{x}, t)$. The data are assumed to be close to each other in the sense that

$$||f_1(\mathbf{x}) - f_2(\mathbf{x})|| < \epsilon_1, \; ||\nabla f_1(\mathbf{x}) - \nabla f_2(\mathbf{x})|| < \epsilon_2, \; ||g_1(\mathbf{x}) - g_2(\mathbf{x})|| < \epsilon_3, \quad (6.8.11)$$

where the norm $||\cdots||$ represents the absolute value of the maximum difference between the functions indicated over the region G, and the constants ϵ_1, ϵ_2, and ϵ_3 are small. We consider the case where $q(\mathbf{x}) > 0$ and $\beta(\mathbf{x}) = 0$ in the boundary condition. Thus the energy $E(t)$ is conserved.

Let $u(\mathbf{x}, t) = u_1(\mathbf{x}, t) - u_2(\mathbf{x}, t)$. The energy integral for $u(\mathbf{x}, t)$ yields

$$E(t) = E(0) = \frac{1}{2} \iint_G \left[\rho(g_1 - g_2)^2 + p(\nabla f_1 - \nabla f_2)^2 + q(f_1 - f_2)^2\right] \, dv, \quad (6.8.12)$$

since, for example, $u_t(\mathbf{x}, 0) = g_1(\mathbf{x}) - g_2(\mathbf{x})$. Using (6.8.11), we can bound $E(t)$ in terms of ϵ_1, ϵ_2, and ϵ_3. We write this as $E(t) \leq M(\epsilon_1, \epsilon_2, \epsilon_3)$, where $M(\epsilon_1, \epsilon_2, \epsilon_3) \to 0$ as ϵ_1, ϵ_2, and ϵ_3 tend to zero Thus we can make $E(t)$ arbitrarily small, say, less than ϵ, if the ϵ_i $(i = 1, 2, 3)$ are small. This yields

$$E(t) = \frac{1}{2} \iint_G \left[\rho \left(\frac{\partial u_1}{\partial t} - \frac{\partial u_2}{\partial t}\right)^2 + p(\nabla u_1 - \nabla u_2)^2 + q(u_1 - u_2)^2\right] \, dv < \epsilon. \quad (6.8.13)$$

The positivity of $\rho(\mathbf{x})$, $p(\mathbf{x})$, and $q(\mathbf{x})$ implies that each term in the integral is small, so that $u_1(\mathbf{x}, t)$ and its first derivatives are close to $u_2(\mathbf{x}, t)$ and its first derivatives. Thus, if the data $f(\mathbf{x})$, $\nabla f(\mathbf{x})$, and $g(\mathbf{x})$ undergo small perturbations, the solutions are perturbed only slightly. This signifies *continuous dependence on the data.* We note that if $F(\mathbf{x}, t)$ and $B(\mathbf{x}, t)$ are not zero, slight variations in their values lead to variations of the corresponding solutions that increase in time. However, in a fixed time interval $0 \leq t \leq T$ the difference between the original and perturbed solutions can be kept uniformly small.

If the domain G is of infinite extent and we are considering the *Cauchy problem* for (6.8.1) with initial data in G, we must assume that the data and the solution $u(\mathbf{x}, t)$ vanish for sufficiently large $|\mathbf{x}|$. Then uniqueness can be proven by the energy integral approach. The surface integral in (6.8.3) is absent and we can conclude that zero data implies that $u(\mathbf{x}, t) = 0$. In the same fashion, continuous dependence on the data can

be shown. In addition, if G is a semi-infinite region, say, a half-space, the preceding arguments also carry through.

We have seen that in the absence of interior or boundary force terms $F(\mathbf{x}, t)$ and $B(\mathbf{x}, t)$, the energy $E(t)$—that is, (6.8.3)—is conserved if $q \geq 0$. However, for (6.8.8), which is essentially a dissipative hyperbolic equation with $2r\rho v_t$ representing the dissipative effect, the energy $E(t)$ decreases in time as seen from (6.8.10). Assuming that the energy integral represents the kinetic and potential energy of the system, the energy is dissipated as t increases and is converted into another energy form, such as heat energy. This is consistent with our results in Chapter 1 relating the telegrapher's equation to the diffusion or heat equation.

We conclude our discussion of the hyperbolic case by noting that more precise results could have been obtained for (6.8.1) that would have given domains of dependence for both the Cauchy and the initial and boundary value problems for (6.8.1). For the sake of simplicity we did not present the more complicated analysis required to obtain these sharper results.

Energy Integrals for Parabolic Equations

Next we consider the *parabolic equation*

$$\rho(\mathbf{x})u_t(\mathbf{x}, t) - \nabla \cdot (p(\mathbf{x})\nabla u(\mathbf{x}, t)) + q(\mathbf{x})u(\mathbf{x}, t) = \rho(\mathbf{x})F(\mathbf{x}, t), \quad \mathbf{x} \in G,\ t > 0, \tag{6.8.14}$$

with the initial condition $u(\mathbf{x}, 0) = f(\mathbf{x})$ and the boundary data (4.1.11). We assume that $p(\mathbf{x})$ and $\rho(\mathbf{x})$ are positive, and $q(\mathbf{x})$ is nonnegative in G, and that $u(\mathbf{x}, t)$ is a classical solution of the initial and boundary value problem.

To obtain an energy integral for (6.8.14) we multiply across by $u(\mathbf{x}, t)$ and integrate over G. Since

$$u(\mathbf{x}, t)\nabla \cdot (p(\mathbf{x})\nabla u(\mathbf{x}, t)) = \nabla \cdot (p(\mathbf{x})u(\mathbf{x}, t)\nabla u(\mathbf{x}, t)) - p(\mathbf{x})\nabla u(\mathbf{x}, t) \cdot \nabla u(\mathbf{x}, t), \tag{6.8.15}$$

we obtain, on using the divergence theorem,

$$\frac{\partial}{\partial t}\left[\frac{1}{2}\iint_G \rho u^2\, dv\right] + \iint_G [p(\nabla u)^2 + qu^2]\, dv = \int_{\partial G} pu\frac{\partial u}{\partial n}\, ds + \iint_G \rho F u\, dv. \tag{6.8.16}$$

We have

$$\int_{\partial G} pu\frac{\partial u}{\partial n}\, ds = -\int_{S_3} \frac{\alpha}{\beta}\, pu^2\, ds + \int_{S_1} \frac{p}{\alpha}\, B\, \frac{\partial u}{\partial n}\, ds + \int_{S_2 \cup S_3} \frac{p}{\beta}\, Bu\, ds. \tag{6.8.17}$$

We define the *energy term* for (6.8.14) as

$$E(t) = \frac{1}{2}\iint_G \rho(\mathbf{x})u^2(\mathbf{x}, t)\, dv. \tag{6.8.18}$$

Assuming that there are two solutions $u_1(\mathbf{x}, t)$ and $u_2(\mathbf{x}, t)$ for the initial and boundary value problem for (6.8.14), the difference $u(\mathbf{x}, t) = u_1(\mathbf{x}, t) - u_2(\mathbf{x}, t)$

satisfies the homogeneous form of (6.8.14) with homogeneous initial and boundary data. Thus we obtain from (6.8.16)–(6.8.17),

$$E'(t) = -\iint_G [p(\nabla u)^2 + qu^2]\, dv - \int_{S_3} \frac{\alpha}{\beta} pu^2\, ds \le 0, \tag{6.8.19}$$

because $p(\mathbf{x})$, $q(\mathbf{x})$, $\rho(\mathbf{x})$, $\alpha(\mathbf{x})$, and $\beta(\mathbf{x})$ are nonnegative. Since $E(t) \ge 0$, $E(0) = \frac{1}{2}\iint_G \rho u(\mathbf{x}, 0)^2\, dv = 0$, and $E(t)$ is nonincreasing in view of (6.8.19), we conclude that $E(t) = 0$ for all t. Thus $u(\mathbf{x}, t) = 0$ and $u_1(\mathbf{x}, t) = u_2(\mathbf{x}, t)$, so that *uniqueness* is established. *Continuous dependence on the data* follows by an argument similar to that in the hyperbolic case. If the region G is of infinite extent and $u(\mathbf{x}, t)$ vanishes as $|\mathbf{x}| \to \infty$, we can also prove uniqueness and continuous dependence on the data.

Again, if $q(\mathbf{x}) < 0$, this method fails, but if $q(\mathbf{x})$ is uniformly bounded in G, the transformation (6.8.7) reduces (6.8.14) to a form for which this approach proves uniqueness.

The equation (6.8.19) shows that the energy $E(t)$ is not conserved and this is the case even if $u(\mathbf{x}, t) = 0$ on the boundary and $q(\mathbf{x}) = 0$. There is, however, another conserved quantity if $q(\mathbf{x}) = 0$ and $\partial u(\mathbf{x}, t)/\partial n$ vanishes on the boundary [i.e., there is no flux of $u(\mathbf{x}, t)$ through the boundary]. With $F(\mathbf{x}, t) = q(\mathbf{x}) = 0$ in (6.8.14), we obtain on integrating the equation over G,

$$\iint_G [\rho u_t - \nabla \cdot (p \nabla u)]\, dv = \frac{\partial}{\partial t} \iint_G \rho u\, dv - \int_{\partial G} p \frac{\partial u}{\partial n}\, ds = \frac{\partial}{\partial t} \iint_G \rho u\, dv = 0. \tag{6.8.20}$$

The quantity $\iint_G \rho u\, dv$ is conserved and if, initially, $u(\mathbf{x}, 0) = f(\mathbf{x})$, we have

$$\iint_G \rho(\mathbf{x}) u(\mathbf{x}, t)\, dv = \iint_G \rho(\mathbf{x}) f(\mathbf{x})\, dv. \tag{6.8.21}$$

The integrals in (6.8.21) represent the total amount of heat, say, in the region G. Even if G is of infinite extent and $u(\mathbf{x}, t)$ vanishes as $|x| \to \infty$, (6.8.21) remains valid.

Energy Integrals for Elliptic Equations

For the *elliptic equation*

$$-\nabla \cdot (p(\mathbf{x}) \nabla u(\mathbf{x})) + q(\mathbf{x}) u(\mathbf{x}) = \rho(\mathbf{x}) F(\mathbf{x}), \qquad \mathbf{x} \in G, \tag{6.8.22}$$

with the boundary condition (4.1.11), we multiply across by $u(\mathbf{x})$ and integrate over G. Using the divergence theorem and (6.8.15), this gives the *energy integral*,

$$\iint_G [p(\nabla u)^2 + qu^2]\, dv = \int_{\partial G} pu \frac{\partial u}{\partial n}\, ds + \iint_G \rho F u\, dv. \tag{6.8.23}$$

The *energy term* is

$$E = \iint_G [p(\mathbf{x})(\nabla u(\mathbf{x}))^2 + q(\mathbf{x}) u^2(\mathbf{x})]\, dv + \int_{S_3} \frac{\alpha(\mathbf{x})}{\beta(\mathbf{x})} p(\mathbf{x}) u^2(\mathbf{x})\, ds. \tag{6.8.24}$$

Proceeding as before, we easily prove *uniqueness* of classical solutions if $\alpha(\mathbf{x}) \neq 0$ in the boundary condition $\alpha(\mathbf{x})u(\mathbf{x}) + \beta(\mathbf{x})\,\partial u(\mathbf{x})/\partial n\big|_{\partial G} = B(\mathbf{x})$. However, if $\alpha(\mathbf{x})$ and $q(\mathbf{x})$ are zero, the solution is unique only up to an additive constant since $u(\mathbf{x})$ = constant is a solution of $\nabla \cdot (p(\mathbf{x})\nabla u(\mathbf{x})) = 0$ with the boundary condition $\beta(\mathbf{x})\,\partial u(\mathbf{x})/\partial n\big|_{\partial G} = 0$. Furthermore, uniqueness can be demonstrated only if $q(\mathbf{x}) > 0$ in (6.8.22). If $q(\mathbf{x}) < 0$, the energy integral method (in the present form) cannot be used to prove uniqueness. In fact, as shown in Chapter 8, the solution need not be unique in that case.

The uniqueness theorem can also be extended to exterior boundary value problems for (6.8.22), with G is a semi-infinite region or a region of infinite extent outside a closed and bounded region with the conditions (4.1.11) assigned on the finite boundaries. The solution is required to satisfy appropriate conditions at infinity. (See the exercises for a discussion of Laplace's and Poisson's equations.) Also, the *continuous dependence on the data* for solutions of the boundary value problem for (6.8.22) can be proven by this method.

The energy integral method can be applied to systems of equations. It can also be used to prove existence of solutions, not only uniqueness and continuous dependence on data, but we do not consider this here.

To conclude our discussion, we note that for *parabolic* and *elliptic equations* there exist *maximum principles* whereby uniqueness and continuous dependence on data can be proven for equations of a more general form than those considered in this section. They are considered in Chapter 8.

Exercises 6.8

6.8.1. Use the energy integral method to prove uniqueness for the solution of the Cauchy problem for the hyperbolic equation (6.8.1).

6.8.2. Construct an energy integral for the one-dimensional form $\rho(x)u_{tt}(x,t) - (p(x)u_x(x,t))_x + q(x)u(x,t) = p(x)F(x,t)$ of the hyperbolic equation (6.8.1) and prove uniqueness and continuous dependence on the data for the initial and boundary value problem for this equation. [In proving continuous dependence, set $F(x,t)$ and the boundary terms equal to zero.]

6.8.3. Prove that the solution of the initial and boundary value problem for the parabolic equation (6.8.14) is continuously dependent on the data, assuming that $F(\mathbf{x},t)$ and the boundary terms are zero.

6.8.4. Determine conditions at infinity that would guarantee uniqueness for the exterior Dirichlet problem for Laplace's and Poisson's equations on the basis of the energy integral method. *Hint*: Apply the energy integral method to the region bounded by the curve or surface where the data are assigned and by a circle or sphere that completely contains the boundary on which those data are given. Allow the radius of the circle or sphere to tend to infinity and obtain conditions on the solution that make the integral over the circle or sphere tend to zero as the radius tends to infinity.

6.8.5. Show that $u(x,y) = \log(\sqrt{x^2+y^2}/a)$ and $u(x,y) = 0$ are both solutions of $u_{xx}(x,y) + u_{yy}(x,y) = 0$ for $x^2 + y^2 > a^2$ with $u(x,y) = 0$ on $x^2 + y^2 = a^2$. Use the conditions at infinity determined in Exercise 6.8.4 to eliminate one of these solutions and to obtain uniqueness for the foregoing exterior Dirichlet problem.

6.8.6. Show that $u(x,y,z) = a/\sqrt{x^2+y^2+z^2}$ and $u(x,y,z) = 1$ are both solutions of the exterior Dirichlet problem for $u_{xx}(x,y,z)+u_{yy}(x,y,z)+u_{zz}(x,y,z) = 0$ in the region $x^2 + y^2 + z^2 > a^2$ with the boundary condition $u(x,y,z) = 1$ on $x^2 + y^2 + z^2 = a^2$. Use the conditions at infinity obtained in Exercise 6.8.4 to eliminate one of the solutions above so that this problem has a unique solution.

6.8.7. Develop an energy integral approach for the first order linear equation $u_t(x,t)+ a(x,t)u_x(x,t) + b(x,t)u(x,t) = 0$, $-\infty < x < \infty$, $t > 0$, by multiplying across by $u(x,t)$, using the identity $auu_x = (\frac{1}{2}au^2)_x - \frac{1}{2}a_xu^2$ and integrating with respect to x from $-\infty$ to ∞. Assuming that all coefficients have the necessary behavior at $\pm\infty$ and using the change of variable (6.8.7), show how the energy integral can be used to prove uniqueness for the Cauchy problem for the given equation.

6.8.8. Consider the first order hyperbolic system with constant coefficients $\mathbf{u}_t(x,t) + A\mathbf{u}_x(x,t) + B\mathbf{u}(x,t) = \mathbf{0}$, $-\infty < x < \infty$, $t > 0$, where A is a real symmetric matrix. Multiply across by the vector $\mathbf{u}^T(x,t)$ and integrate from $-\infty$ to ∞ with respect to x. Assuming that $\mathbf{u}(x,t)$ vanishes as $|x| \to \infty$, show that the energy integral $E(t) = \int_{-\infty}^{\infty} |\mathbf{u}(x,t)|^2\, dx \leq 0$ if B is a nonnegative matrix with the property that $\mathbf{u}^T(x,t)B\mathbf{u}(x,t) \geq 0$ for all vectors $\mathbf{u}(x,t)$. If B is not nonnegative, the transformation $\mathbf{u}(x,t) = e^{rt}\mathbf{v}(x,t)$ with $r > 0$ yields a system for $\mathbf{v}(x,t)$ in which the coefficient $\hat{B}$ of $\mathbf{v}(x,t)$ can be made nonnegative for an appropriate choice of r. Use the energy integral $E(t)$ to prove uniqueness for the Cauchy problem for the given system.

6.8.9. Use the method of Exercise 6.8.8 to show that the Cauchy problem for the system (2.1.2), which is equivalent to the wave equation, has a unique solution.

6.8.10. Apply the method of Exercise 6.8.8 to prove uniqueness for the solution of the Cauchy problem for the system (1.2.23)–(1.2.24), which is equivalent to the telegrapher's equation.

6.8.11. Use an energy integral to show that the initial and boundary value problem for the equation of a vibrating rod (see Section 8.5) $u_{tt}(x,t) + u_{xxxx}(x,t) = 0$, $0 < x < l$, $t > 0$, with $u(x,0)$ and $u_t(x,0)$ specified and $u(x,t)$ and $u_x(x,t)$ given at $x = 0$ and $x = l$, has a unique solution. *Hint*: Use the identity $u_tu_{xxxx} = (u_tu_{xxx})_x - (u_{tx}u_{xx})_x + \frac{1}{2}(u_{xx}^2)_t$.

6.8.12. Show that the boundary value problem for the biharmonic equation (see Section 8.5) $\nabla^2\nabla^2u(x,y) = 0$, in a bounded region R, with $u(x,y)$ and $\partial u(x,y)/\partial n$ specified on ∂R has a unique solution by using an appropriate energy integral. *Hint*: $u\nabla^2\nabla^2u = u\nabla\cdot(u\nabla\nabla^2u) = \nabla\cdot(u\nabla\nabla^2u) - \nabla\cdot(\nabla u\nabla^2u) + (\nabla^2u)^2$. Deduce that if $u(x,y) = \partial u(x,y)/\partial n = 0$ on ∂R, we must have $\nabla^2u(x,y) = 0$, and conclude from the uniqueness theorem for Laplace's equation that $u(x,y) = 0$.

6.8.13. Combine the methods of Exercises 6.8.11 and 6.8.12 to develop an energy integral that yields uniqueness for the initial and boundary value problem for the equation of a vibrating plate (see Section 8.5) $u_{tt}(x,y,t) + c^2\nabla^2\nabla^2 u(x,y,t) = 0$, in a bounded region R where $u(x,y,0)$ and $u_t(x,y,0)$ are prescribed and $u(x,y,t)$, together with $\partial u(x,y,t)/\partial n$, is specified on ∂R.

6.9 MAPLE METHODS

There are essentially no built-in procedures in Maple that deal directly with the general concepts introduced and discussed in this chapter. We have seen, however, that in determining fundamental solutions it becomes necessary to solve ordinary differential equations and to examine the properties of their solutions, which may involve special functions. Some of these matters are considered here.

Integral Wave Equation

Section 6.5, which deals with the integral wave equation, presents a number of results that represent solution formulas or techniques for the solution of the wave equation. They lend themselves to the creation of Maple procedures that automate the construction of solutions via these processes.

We present only one procedure $IntWaveEqNum$, which finds numerical solutions of the initial and boundary value problem for the (one-dimensional) homogeneous wave equation for $u(x,t)$ in a finite interval $0 < x < l$ with Dirichlet conditions assigned at the endpoints. The problem is given as

$$u_{tt}(x,t) = c^2 u_{xx}(x,t),\ u(x,0) = f(x),\ u_t(x,0) = g(x),$$

$$u(0,t) = r(t),\ u(l,t) = s(t). \tag{6.9.1}$$

To find the solution of (6.9.1) at the point $(x,t) = (\xi,\tau)$, the arguments are given as *IntWaveEqNum*$(c, u(x,t), [f(x), g(x)], [r(t), s(t)], x = 0..l, [\xi,\tau])$. The procedure finds the solution at the point $x = \xi$ at the time $t = \tau$ in terms of the initial and boundary data in a stepwise fashion.

Example 6.16. The Wave Equation in a Finite Interval. We put $c = 1$ and $l = 1$ in the procedure $IntWaveEqNum$ and determine the solution at points in the four regions R_1, R_2, R_3, R_4 displayed in Figure 6.11. The initial and boundary data are kept arbitrary. Let $\hat{u}(\xi,\tau)$ represent the output of the procedure at the point (ξ,τ). Then we have

$$\hat{u}(0.5, 0.35) = 0.5\,(f\,(0.85) + f\,(0.15)) + 0.5\int_{0.15}^{0.85} g\,(s)\,ds. \tag{6.9.2}$$

The point $(0.5, 0.35)$ lies in the region R_1 and (6.9.2) is just *d'Alembert's solution* evaluated at that point. Next, we consider a point in R_2 and obtain

$$\hat{u}(0.9, 0.35) = 0.50\, f(0.75) - 0.50\, f(0.55) + 0.50 \int_{0.55}^{0.75} g(s)\, ds + s(0.25). \tag{6.9.3}$$

Only the effect of the boundary value at $x = 1$ is felt at this time. Then we consider a point in R_3 and obtain

$$\hat{u}(0.1, 0.65) = -0.5\,(f(0.55) - f(0.75)) + 0.5 \int_{0.55}^{0.75} g(s)\, ds + r(0.55). \tag{6.9.4}$$

Now, the boundary value at $x = 0$ but not at $x = 1$ appears in the solution. Finally, we consider a point in R_4 and find that both boundary conditions occur in the result, as expected.

$$\hat{u}(0.5, 1.25) = r(0.75) + s(0.75) - 0.5\,(f(0.75) + f(0.25)) - 0.5 \int_{0.25}^{0.75} g(s)\, ds. \tag{6.9.5}$$

As a concrete example we consider the function $u(x,t) = (x+t)^2$, which is a solution of the wave equation (6.9.1) with $c = 1$. We set $l = 1$, evaluate the initial and boundary values for this solution and determine its value at the point $(x,t) = (0.4, 46.73)$. On using the initial and boundary conditions in the procedure $IntWaveEqNum(1, u(x,t), [x,t], [x^2, 2x], [t^2, (1+t)^2], x = 0 \ldots 1, [0.4, 46.73])$ the output $\hat{u}(0.4, 46.73) = 2221.2369$, equals the value of the solution at the given point, which is $u(0.4, 46.73) = 2221.2369$.

Fundamental Solutions

A basic technique for finding fundamental solutions of partial differential equations with constant coefficients involved a transformation of the variables that reduced the PDEs to ODEs. The solutions of the ODEs were found and on the basis of their behavior at the singular point and at infinity, in a number of cases, fundamental solutions were constructed. Maple can be used to convert the PDEs to ODEs, and then procedures such as *dsolve*, *series*, and *asympt* can be invoked to solve the ODEs and analyze their behavior. We reexamine one of the elliptic PDEs considered in Example 6.13 from this point of view.

Example 6.17. The Free Space Green's Function for an Elliptic Equation. We consider the elliptic equation (6.7.24) with constant coefficients in two dimensions,

$$p\left(\partial^2 u(x,y)/\partial x^2 + \partial^2 u(x,y)/\partial y^2\right) - q u(x,y) = 0, \tag{6.9.6}$$

where $p > 0$ and $q < 0$. Then $u(x,y) = \tilde{u}\left(\sqrt{(x-x_0)^2 + (y-y_0)^2}\right)$ yields

$$p(\tilde{u}''(r) + \frac{1}{r}\tilde{u}'(r)) - q\tilde{u}(r) = 0, \tag{6.9.7}$$

where $r = \sqrt{(x-x_0)^2 + (y-y_0)^2}$. The Maple procedure *dsolve* gives

$$\tilde{u}(r) = c_1 J_0\left(\sqrt{-q/p}\, r\right) + c_2 Y_0\left(\sqrt{-q/p}\, r\right), \tag{6.9.8}$$

in terms of the Bessel functions J_0 and Y_0 with arbitrary constants c_1 and c_2.

The Bessel function Y_0 is singular at $r = 0$. *series* determines the behavior of $\tilde{u}(r)$ near $r = 0$ to be $\tilde{u}(r) \approx c_2[2\ln(r)/\pi]$, $r \to 0$. On proceeding as in Example 6.13, we find that we must put $c_2 = -1/4p$. To determine the constant c_1, we must examine the behavior of $\tilde{u}(r)$ at infinity. The Maple procedure *asympt* finds the required asymptotic behavior as

$$\tilde{u}(r) \approx \sqrt{\frac{1}{2\pi r}}\,(-p/q)^{1/4}\,(c_1 - ic_2)\exp\left[i\left(\sqrt{-q/p}\, r - \frac{\pi}{4}\right)\right]$$

$$+ \sqrt{\frac{1}{2\pi r}}\,(-p/q)^{1/4}\,(c_1 + ic_2)\exp\left[-i\left(\sqrt{-q/p}\, r - \frac{\pi}{4}\right)\right], \quad r \to \infty. \tag{6.9.9}$$

To satisfy the *Sommerfeld radiation condition*, we must put $c_1 = -ic_2 = i/4p$.

As a result, (6.9.8) becomes

$$\tilde{u}(r) = \frac{i}{4p}\left\{J_0\left(\sqrt{-q/p}\, r\right) + iY_0\left(\sqrt{-q/p}\, r\right)\right\} = \frac{i}{4p} H_0^{(1)}\left(\sqrt{-q/p}\, r\right). \tag{6.9.10}$$

Exercises 6.9

6.9.1. Use the procedure $IntWaveEqNum$ to obtain the results given in Example 6.16.

6.9.2. Use the Maple procedures referred to in Example 6.17 to reproduce the results given there.

CHAPTER 7

GREEN'S FUNCTIONS

The method of *Green's functions* is an important technique for solving boundary value, initial and boundary value, and Cauchy problems for partial differential equations. It is most commonly identified with the solution of boundary value problems for Laplace's equation and a Green's function has already been introduced in that context in Chapter 1. It was also seen in Section 1.3 and in our study of point source problems in Section 6.7 that the Green's function is often worthwhile determining in its own right rather than as a tool to be used only for solving another problem.

In this chapter we begin by constructing generalizations of *Green's second theorem* that are appropriate for the second order differential equations introduced in Chapter 4. These integral theorems [which are special cases of the general result (3.6.7) given in Section 3.6] are then used to show how boundary value, initial and boundary value, and Cauchy problems can be solved in terms of appropriately defined Green's functions for each of these problems. Even though the construction of Green's functions requires that a problem similar to the original (given) problem must be solved, it is often easier to solve the Green's function problem in a number of important cases, as we shall see. In this regard the *fundamental solutions* considered in Section 6.7, of which Green's functions are a special case, play an important role. Since the determination and use of Green's functions require the use of *generalized functions* such as the Dirac

Partial Differential Equations of Appplied Mathematics, Third Edition. By Erich Zauderer

delta function, a brief discussion of the theory of generalized functions is given in this chapter. Most of the chapter, however, is devoted to the construction and use of Green's functions for problems involving equations of elliptic, hyperbolic, and parabolic types.

7.1 INTEGRAL THEOREMS AND GREEN'S FUNCTIONS

In this section we construct *integral theorems* appropriate for the *elliptic, hyperbolic,* and *parabolic equations* introduced in Section 4.1. Each of these theorems follows from an application of the divergence theorem and represents a generalization of Green's second theorem. These theorems form the basis for the construction of the Green's functions we consider in this chapter. Technically, the theorems are valid only if the functions occurring in the integrals are sufficiently smooth, and as we have seen in Section 6.7, this is generally not the case for Green's functions. Nevertheless, we shall assume that these theorems are formally valid in all cases and rely on the theory of generalized functions presented in Section 7.2 to form a basis for their validity, even though this is not demonstrated. We begin our discussion with problems in two or three space dimensions and present the one-dimensional results at the end of this section. Even though these integral theorems are special cases of the general result (3.6.7), we include some details of their derivation.

Integral Theorems and Green's Functions for Elliptic Equations

We start with the *elliptic equation*

$$Lu(\mathbf{x}) = -\nabla \cdot (p(\mathbf{x})\nabla u(\mathbf{x})) + q(\mathbf{x})u(\mathbf{x}) = \rho(\mathbf{x})F(\mathbf{x}) \tag{7.1.1}$$

in two or three dimensions given over a bounded region G with the boundary conditions

$$\alpha(\mathbf{x})\, u(\mathbf{x}) + \beta(\mathbf{x}) \left.\frac{\partial u(\mathbf{x})}{\partial n}\right|_{\partial G} = B(\mathbf{x}). \tag{7.1.2}$$

The conditions on the coefficients in (7.1.1) and (7.1.2) given in Section 4.1 are assumed to remain in effect. Introducing a function $w(\mathbf{x})$ whose properties are to be specified and proceeding as in Example 4.2, we obtain

$$\iint_G [wLu - uLw]\, dv = -\int_{\partial G} p[w\nabla u - u\nabla w] \cdot \mathbf{n}\, ds = \int_{\partial G} p\left[u\frac{\partial w}{\partial n} - w\frac{\partial u}{\partial n}\right] ds, \tag{7.1.3}$$

on applying the divergence theorem with $\mathbf{n}$ as the exterior unit normal on ∂G. Equation (7.1.3) is the basic *integral theorem* from which the *Green's function method* proceeds in the *elliptic case.*

The function $w(\mathbf{x})$ is now determined such that (7.1.3) expresses $u(\mathbf{x})$ at an arbitrary point ξ in the region G in terms of $w(\mathbf{x})$ and known functions in (7.1.1) and (7.1.2).

Let $w(\mathbf{x})$ be a solution of $Lw(\mathbf{x}) = \delta(\mathbf{x} - \boldsymbol{\xi})$, where $\delta(\mathbf{x} - \boldsymbol{\xi})$ is a two- or three-dimensional Dirac delta function. The substitution property of the delta function then yields

$$\iint_G u(\mathbf{x})Lw(\mathbf{x})\, dv = \iint_G u(\mathbf{x})\, \delta(\mathbf{x} - \boldsymbol{\xi})\, dv = u(\boldsymbol{\xi}). \tag{7.1.4}$$

In view of (7.1.1) we also have

$$\iint_G w(\mathbf{x})Lu(\mathbf{x})\, dv = \iint_G \rho(\mathbf{x})w(\mathbf{x})F(\mathbf{x})\, dv. \tag{7.1.5}$$

It now remains to choose boundary conditions for $w(\mathbf{x})$ on ∂G so that the boundary integral in (7.1.3) involves only $w(\mathbf{x})$ and known functions. This can be accomplished by requiring $w(\mathbf{x})$ to satisfy the homogeneous version of the boundary condition (7.1.2); that is, $\alpha(\mathbf{x})\, w(\mathbf{x}) + \beta(\mathbf{x})\, \partial w(\mathbf{x})/\partial n\big|_{\partial G} = 0$. If $\mathbf{x} \in S_1$ on ∂G, we have

$$u\frac{\partial w}{\partial n} - w\frac{\partial u}{\partial n} = \frac{1}{\alpha}B\frac{\partial w}{\partial n}, \tag{7.1.6}$$

in view of (7.1.2). If $\mathbf{x} \in S_2 \cup S_3$ on ∂G, we have

$$u\frac{\partial w}{\partial n} - w\frac{\partial u}{\partial n} = -\frac{1}{\beta}Bw. \tag{7.1.7}$$

The function $w(\mathbf{x})$ is called the *Green's function* for the boundary value problem (7.1.1)–(7.1.2). To indicate its dependence on the point $\boldsymbol{\xi}$, we denote the Green's function by $w(\mathbf{x}) = K(\mathbf{x}; \boldsymbol{\xi})$, as in Section 1.3. In terms of the Green's function $K(\mathbf{x}; \boldsymbol{\xi})$, the foregoing implies that (7.1.3) takes the form

$$u(\boldsymbol{\xi}) = \iint_G \rho K(\mathbf{x}; \boldsymbol{\xi})F\, dv - \int_{S_1} \frac{pB}{\alpha}\frac{\partial K(\mathbf{x}; \boldsymbol{\xi})}{\partial n}\, ds + \int_{S_2 \cup S_3} \frac{pB}{\beta}K(\mathbf{x}; \boldsymbol{\xi})\, ds. \tag{7.1.8}$$

The Green's function $K(\mathbf{x}; \boldsymbol{\xi})$ thus satisfies the equation

$$-\nabla \cdot (p(\mathbf{x})\nabla K(\mathbf{x}; \boldsymbol{\xi})) + q(\mathbf{x})K(\mathbf{x}; \boldsymbol{\xi}) = \delta(\mathbf{x} - \boldsymbol{\xi}), \qquad \mathbf{x},\ \boldsymbol{\xi} \in G, \tag{7.1.9}$$

and the boundary condition

$$\alpha(\mathbf{x})K(\mathbf{x}; \boldsymbol{\xi}) + \beta(\mathbf{x})\left.\frac{\partial K(\mathbf{x}; \boldsymbol{\xi})}{\partial n}\right|_{\partial G} = 0, \tag{7.1.10}$$

with the derivatives taken in the $\mathbf{x}$-variables. It follows from (7.1.9) and Section 6.7 that the Green's function is a *fundamental solution* of (7.1.1). This fact will be exploited in the construction of certain Green's functions.

Not all Green's function problems (7.1.9)–(7.1.10) have solutions. In certain cases considered later, a generalized or modified Green's function must be constructed that satisfies an equation that differs from (7.1.9) or boundary conditions that differ from (7.1.10). However, once the Green's function has been determined, the formulas

(7.1.8) or slightly modified ones in the generalized case yield the solution $u(\mathbf{x})$ of the boundary value problem (7.1.1)–(7.1.2) at any point in G. By introducing appropriate assumptions on the behavior of the solutions at infinity, the Green's function technique can also be applied to problems over unbounded regions. We construct Green's functions for specific elliptic equations of the form (7.1.1) over bounded and unbounded regions.

Integral Theorems and Green's Functions for Hyperbolic Equations

We consider the initial and boundary value problem for the *hyperbolic equation*

$$\rho(\mathbf{x})u_{tt}(\mathbf{x},t) + Lu(\mathbf{x},t) = \rho(\mathbf{x})F(\mathbf{x},t), \qquad \mathbf{x} \in G,\ t > 0, \tag{7.1.11}$$

where the operator L is defined as in (7.1.1) and G is a bounded region in two or three-dimensional space. The initial conditions for $u(\mathbf{x},t)$ are

$$u(\mathbf{x},0) = f(\mathbf{x}), \qquad u_t(\mathbf{x},0) = g(\mathbf{x}), \qquad \mathbf{x} \in G. \tag{7.1.12}$$

The boundary conditions on ∂G are given as in Section 4.1 in the form

$$\alpha(\mathbf{x})u(\mathbf{x},t) + \beta(\mathbf{x}) \left.\frac{\partial u(\mathbf{x},t)}{\partial n}\right|_{\partial G} = B(\mathbf{x},t), \qquad t > 0. \tag{7.1.13}$$

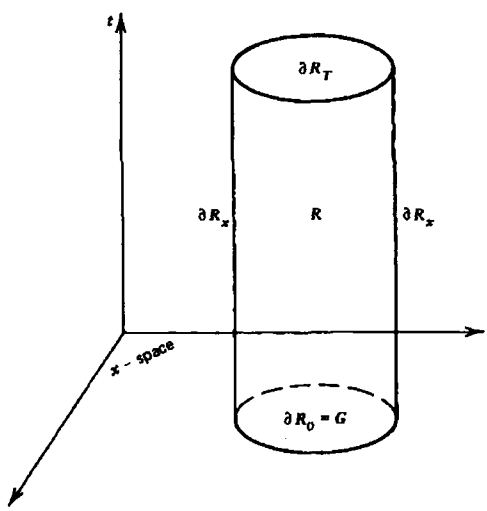

Figure 7.1 The region R.

The integral theorem appropriate for the problem is given over the bounded cylindrical region $R = G \times [0,\ T]$ in $(\mathbf{x},t)$-space , as shown in Figure 7.1 ($T > 0$ is an arbitrary number). The lateral boundary of R is denoted by $\partial R_{\mathbf{x}}$ and the two caps of

the cylinder, which are portions of the planes $t = 0$ and $t = T$, are denoted by ∂R_0 and ∂R_T, respectively. ∂R_0 is identical to the region G and the initial conditions for $u(\mathbf{x}, t)$ are assigned on it. The boundary conditions for $u(\mathbf{x}, t)$ are assigned on $\partial R_{\mathbf{x}}$. The exterior unit normal $\mathbf{n}$ to ∂R has the form $\mathbf{n} = [\mathbf{n}_{\mathbf{x}}, 0]$ on $\partial R_{\mathbf{x}}$, where $\mathbf{n}_{\mathbf{x}}$ is the exterior unit normal to ∂G. On ∂R_0, $\mathbf{n}$ has the form $\mathbf{n} = [\mathbf{0}, -1]$, and on ∂R_T it has the form $\mathbf{n} = [\mathbf{0}, 1]$.

It follows from Section 6.4 that for an arbitrary function $w(\mathbf{x}, t)$ we have

$$\iint_R [w(\rho u_{tt} + Lu) - u(\rho w_{tt} + Lw)]\, dv = \int_{\partial R} [-pw\nabla u + pu\nabla w, \rho w u_t - \rho u w_t] \cdot \mathbf{n}\, ds$$
$$= \int_{\partial R_{\mathbf{x}}} (-pw\nabla u + pu\nabla w) \cdot \mathbf{n}_{\mathbf{x}}\, ds + \int_{\partial R_T} (\rho w u_t - \rho u w_t)\, d\mathbf{x} - \int_{\partial R_0} (\rho w u_t - \rho u w_t)\, d\mathbf{x}, \tag{7.1.14}$$

where $\tilde{\nabla} = [\nabla, \partial/\partial t]$, the gradient operator in space-time, the divergence theorem, as well as the foregoing results concerning the exterior unit normal to the boundary ∂R have been used. The *integral relation* (7.1.14) forms the basis for the *Green's function method* for solving the initial and boundary value problem (7.1.11)–(7.1.13).

We now show how $w(\mathbf{x}, t)$ is specified so that the solution $u(\mathbf{x}, t)$ of (7.1.11)–(7.1.13) can be determined at an arbitrary point $(\boldsymbol{\xi}, \tau)$ in the region R from (7.1.14). First we require that $w(\mathbf{x}, t)$ be a solution of

$$\rho(\mathbf{x}) w_{tt}(\mathbf{x}, t) + Lw(\mathbf{x}, t) = \delta(\mathbf{x} - \boldsymbol{\xi})\delta(t - \tau), \qquad \boldsymbol{\xi} \in G,\ 0 < \tau < T. \tag{7.1.15}$$

The product of the Dirac delta functions in (7.1.15) has the effect

$$\iint_R u(\rho w_{tt} + Lw)\, dv = \iint_R u\, \delta(\mathbf{x} - \boldsymbol{\xi})\delta(t - \tau)\, dv = u(\boldsymbol{\xi}, \tau). \tag{7.1.16}$$

In addition, we obtain from (7.1.11),

$$\iint_R w(\rho u_{tt} + Lu)\, dv = \iint_R \rho w F\, dv, \tag{7.1.17}$$

so that this term is known once $w(\mathbf{x}, t)$ is specified.

Since

$$\int_{\partial R_{\mathbf{x}}} p(-w\nabla u + u\nabla w) \cdot \mathbf{n}_{\mathbf{x}}\, ds = \int_{\partial R_{\mathbf{x}}} p\left(-w\frac{\partial u}{\partial n} + u\frac{\partial w}{\partial n}\right) ds, \tag{7.1.18}$$

we see that if we require, as in the elliptic case, that

$$\alpha w + \beta \left.\frac{\partial w}{\partial n}\right|_{\partial R_{\mathbf{x}}} = 0, \tag{7.1.19}$$

we obtain

$$\int_{\partial R_{\mathbf{x}}} p\left(-w\frac{\partial u}{\partial n} + u\frac{\partial w}{\partial n}\right) ds = \int_{\hat{S}_1} \frac{p}{\alpha} B \frac{\partial w}{\partial n}\, ds - \int_{\hat{S}_2 \cup \hat{S}_3} \frac{p}{\beta} Bw\, ds, \tag{7.1.20}$$

where $\hat{S}_1$, $\hat{S}_2$, and $\hat{S}_3$ are the portions of $\partial R_{\mathbf{x}}$ that correspond to S_1, S_2, and S_3 on ∂G, respectively.

To complete the determination of $w(\mathbf{x}, t)$, we expect that initial conditions must be assigned to it for some value of t. If $w(\mathbf{x}, t)$ and $w_t(\mathbf{x}, t)$ are specified at $t = 0$, the integral over ∂R_0 in (7.1.14) is determined completely since $u(\mathbf{x}, t)$ and $u_t(\mathbf{x}, t)$ are given at $t = 0$. However, $u(\mathbf{x}, t)$ and $u_t(\mathbf{x}, t)$ at $t = T$ (i.e., on ∂R_T) are not known. If we specify $w(\mathbf{x}, t)$ and $w_t(\mathbf{x}, t)$ at $t = T$, it must be done in such a way that the unknown values of $u(\mathbf{x}, t)$ and $u_t(\mathbf{x}, t)$ play no role in the integral over ∂R_T. The only possible choice is to set

$$w(\mathbf{x}, T) = 0, \qquad w_t(\mathbf{x}, T) = 0, \tag{7.1.21}$$

so that the entire integral over ∂R_T vanishes.

The equation (7.1.15) together with the boundary condition (7.1.19) and the conditions (7.1.21) at $t = T$ constitutes a *backward initial and boundary value problem* for the function $w(\mathbf{x}, t)$. It differs from the types of problems considered previously for hyperbolic equations (see, however, Section 1.2), where initial conditions were assigned at $t = 0$ and the problem was solved for $t > 0$. Here we assign end conditions at $t = T$ and solve the problem for $t < T$. The problem for $w(\mathbf{x}, t)$ is *well posed* because if t is replaced by $-t$ in $w_{tt}(\mathbf{x}, t)$ its sign is unchanged. We will refer to problems for which either initial conditions or end conditions are assigned as initial value problems. The function $w(\mathbf{x}, t)$ determined from (7.1.15), (7.1.19), and (7.1.21) is called the *Green's function* for the initial and boundary value problem (7.1.11)–(7.1.13) for $u(\mathbf{x}, t)$. It is denoted as $w(\mathbf{x}, t) = K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)$.

Once the initial and boundary value problem for $K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)$ is solved, the values of $K(\mathbf{x}, 0; \boldsymbol{\xi}, \tau)$ and $K_t(\mathbf{x}, 0; \boldsymbol{\xi}, \tau)$ are known. Then the foregoing results yield the solution $u(\mathbf{x}, t)$ at an (arbitrary) point $(\boldsymbol{\xi}, \tau)$ as

$$u(\boldsymbol{\xi}, \tau) = \iint_R \rho K F \, dv + \int_{\partial R_0} (\rho K g - \rho K_t f) \, d\mathbf{x} - \int_{\hat{S}_1} \frac{pB}{\alpha} \frac{\partial K}{\partial n} \, ds + \int_{\hat{S}_2 \cup \hat{S}_3} \frac{pBK}{\beta} \, ds. \tag{7.1.22}$$

For completeness, we state the problem that the Green's function $K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)$ must satisfy. It is a solution of the equation

$$\rho(\mathbf{x}) K_{tt}(\mathbf{x}, t; \boldsymbol{\xi}, \tau) + LK(\mathbf{x}, t; \boldsymbol{\xi}, \tau) = \delta(\mathbf{x} - \boldsymbol{\xi}) \delta(t - \tau), \ \mathbf{x}, \boldsymbol{\xi} \in G, \ t, \tau < T, \ \tau > 0, \tag{7.1.23}$$

with the end conditions

$$K(\mathbf{x}, T; \boldsymbol{\xi}, \tau) = 0, \qquad K_t(\mathbf{x}, T; \boldsymbol{\xi}, \tau) = 0, \tag{7.1.24}$$

and the boundary condition

$$\alpha(\mathbf{x}) K(\mathbf{x}, t; \boldsymbol{\xi}, \tau) + \beta(\mathbf{x}) \left. \frac{\partial K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)}{\partial n} \right|_{\partial R_x} = 0, \qquad t < T. \tag{7.1.25}$$

It is shown in the exercises that $K(\mathbf{x}, t; \boldsymbol{\xi}, \tau) = K(\boldsymbol{\xi}, -\tau; x, -t)$. Therefore, as a function of $\boldsymbol{\xi}$ and τ, $K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)$ satisfies the same differential equation but with

time running forward instead of backward. In these variables it represents the *causal fundamental solution* for the given hyperbolic operator and the boundary condition (7.1.25).

Integral Theorems and Green's Functions for Parabolic Equations

The *parabolic equation*

$$\rho(\mathbf{x})u_t(\mathbf{x},t) + Lu(\mathbf{x},t) = \rho(\mathbf{x})F(\mathbf{x},t), \qquad \mathbf{x} \in G,\ t > 0, \tag{7.1.26}$$

with the initial and boundary conditions

$$u(\mathbf{x},0) = f(\mathbf{x}),\ \mathbf{x} \in G,\ \alpha(\mathbf{x})u(\mathbf{x},t) + \beta(\mathbf{x})\left.\frac{\partial u(\mathbf{x},t)}{\partial n}\right|_{\partial G} = B(\mathbf{x},t),\ t > 0, \tag{7.1.27}$$

can be treated in the same way as the hyperbolic problem (7.1.11)–(7.1.13). The operator L, the regions R and G, and their boundaries are defined as in the foregoing hyperbolic problem.

We introduce the function $w(\mathbf{x},t)$ and consider the integral relation

$$\iint_R [w(\rho u_t + Lu) - u(-\rho w_t + Lw)]\,dv = \iint_R \tilde{\nabla}\cdot[-pw\nabla u + pu\nabla w,\ \rho wu]\,dv$$

$$= \int_{\partial R_x}\left(-pw\frac{\partial u}{\partial n} + pu\frac{\partial w}{\partial n}\right)ds + \int_{\partial R_T}\rho wu\,d\mathbf{x} - \int_{\partial R_0}\rho wu\,d\mathbf{x}. \tag{7.1.28}$$

Again, $\tilde{\nabla} = [\nabla, \partial/\partial t]$ is the gradient operator in space-time and the region R and its boundaries are as shown in Figure 7.1. The result (7.1.28) is a consequence of the divergence theorem, but it differs from the preceding integral theorems for the elliptic and hyperbolic problems in the following respect. The operator $\rho(\partial/\partial t) + L$ in the parabolic equation (7.1.26) is not self-adjoint. Its *adjoint operator* is given as $-\rho(\partial/\partial t) + L$. With this choice for the adjoint operator we find that $w(\rho u_t + Lu) - u(-\rho w_t + Lw)$ is a divergence expression, as is shown in (7.1.28) (see Example 3.9).

We require $w(\mathbf{x},t)$ to be a solution of

$$-\rho(\mathbf{x})w_t(\mathbf{x},t) + Lw(\mathbf{x},t) = \delta(\mathbf{x}-\boldsymbol{\xi})\delta(t-\tau), \qquad \boldsymbol{\xi} \in G,\ 0 < \tau < T, \tag{7.1.29}$$

with the end and boundary conditions

$$w(\mathbf{x},T) = 0, \quad \alpha(\mathbf{x})w(\mathbf{x},t) + \beta(\mathbf{x})\left.\frac{\partial w(\mathbf{x},t)}{\partial n}\right|_{\partial R_x} = 0. \tag{7.1.30}$$

Then, $w(\mathbf{x},t) = K(\mathbf{x},t;\boldsymbol{\xi},\tau)$ is the Green's function for the initial and boundary value problem (7.1.26)–(7.1.27). It follows from (7.1.28) that

$$u(\boldsymbol{\xi},\tau) = \iint_R \rho KF\,dv + \int_{\partial R_0}\rho Kf\,d\mathbf{x} - \int_{\hat{S}_1}\frac{p}{\alpha}B\frac{\partial K}{\partial n}\,ds + \int_{\hat{S}_2\cup\hat{S}_3}\frac{p}{\beta}BK\,ds. \tag{7.1.31}$$

For completeness, we state the problem that the Green's function $K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)$ must satisfy. It is a solution of the equation

$$-\rho(\mathbf{x})K_t(\mathbf{x}, t; \boldsymbol{\xi}, \tau) + LK(\mathbf{x}, t; \boldsymbol{\xi}, \tau) = \delta(\mathbf{x}-\boldsymbol{\xi})\delta(t-\tau),\ \mathbf{x}, \boldsymbol{\xi} \in G,\ t, \tau < T,\ \tau > 0, \tag{7.1.32}$$

with the end and boundary conditions

$$K(\mathbf{x}, T; \boldsymbol{\xi}, \tau) = 0,\ \alpha(\mathbf{x})K(\mathbf{x}, t; \boldsymbol{\xi}, \tau) + \beta(\mathbf{x}) \left.\frac{\partial K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)}{\partial n}\right|_{\partial R_x} = 0,\ t < T. \tag{7.1.33}$$

The equation (7.1.32) satisfied by the *Green's function* $K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)$ is a *backward parabolic equation* that results on reversing the direction of time in the (forward) parabolic equation (7.1.26). Since the problem for the Green's function is to be solved backward in time, the initial and boundary value problem (7.1.32)–(7.1.33) for K is *well posed.* [That is, we must determine $K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)$ for $t < T$ with an end condition given at $t = T$.] Once $K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)$ has been determined, all the terms on the right side of (7.1.31) are known and the solution $u(\mathbf{x}, t)$ of the initial and boundary value problem (7.1.26)–(7.1.27) is specified completely.

It is shown in the exercises that $K(\mathbf{x}, t; \boldsymbol{\xi}, \tau) = K(\boldsymbol{\xi}, -\tau; x, -t)$. Therefore, as a function of $\boldsymbol{\xi}$ and τ, $K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)$ satisfies a forward parabolic differential equation, but with time now running forwards instead of backwards. In these variables it represents the *causal fundamental solution* for the given parabolic operator and the boundary condition (7.1.33).

Causal Fundamental Solutions and Green's Functions for Cauchy Problems

The Green's functions $K(\mathbf{x}; \boldsymbol{\xi})$ and $K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)$ defined above are *fundamental solutions* of the PDEs (7.1.9), (7.1.23) or (7.1.32) in the elliptic, hyperbolic, and parabolic cases, respectively. Each of these equations is the adjoint of the given equation for $u(\mathbf{x})$ or $u(\mathbf{x}, t)$. Since the elliptic and hyperbolic equations are self-adjoint, the Green's function is also a fundamental solution of the given equation. In the parabolic case, since the given equation is not self-adjoint, the Green's function $K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)$ is only a fundamental solution of the adjoint equation. We have shown, however, that as a function of $\boldsymbol{\xi}$ and τ, $K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)$ is a fundamental solution of the given parabolic equation.

Let us consider an *instantaneous point source problem* for the hyperbolic and parabolic cases, with the source acting at the time $t = \tau$ and located at the point $\mathbf{x} = \boldsymbol{\xi}$. We require that the solution $u(\mathbf{x}, t)$ satisfy the homogeneous boundary condition $\alpha(\mathbf{x})u(\mathbf{x}, t) + \beta(\mathbf{x})\partial u(\mathbf{x}, t)/\partial n = 0$ on ∂G for $t > \tau$ and that $u(\mathbf{x}, t) = 0$ for $t < \tau$. Then the solution is called a *causal fundamental solution* for the initial and boundary value problem. (In Section 6.7 we found causal fundamental solutions over unbounded regions.) We have already indicated how to obtain these solutions in terms of the Green's functions $K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)$.

The Green's function method can also be used to solve *Cauchy problems* for hyperbolic and parabolic equations. In the hyperbolic case we assume that $u(\mathbf{x}, t)$ is a solution of (7.1.11) with initial data (7.1.12) and in the parabolic case $u(\mathbf{x}, t)$ is a solution of (7.1.26) that satisfies the initial condition (7.1.27). Both problems are given over the entire two- or three-dimensional space. The integral theorems (7.1.14) and (7.1.28) can be used for these problems if we assume that the (spatial) boundary $\partial R_{\mathbf{x}}$ tends to infinity and the solution $u(\mathbf{x}, t)$ and the Green's function $K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)$ are such that the contributions from these integrals vanish in the limit.

The Green's function $K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)$ for the *hyperbolic case* is taken to be the solution of the backward Cauchy problem (7.1.23) and (7.1.24). Then the solution of the Cauchy problem (7.1.11)–(7.1.12) is given as

$$u(\boldsymbol{\xi}, \tau) = \int_0^T \int_{\mathbf{x}-space} \rho K F \, d\mathbf{x} \, dt + \int_{\mathbf{x}-space} [\rho K g - \rho K_t f]\big|_{t=0} \, d\mathbf{x}, \quad (7.1.34)$$

as is easily seen from the (modified) integral relation (7.1.14).

In the *parabolic case* the Green's function $K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)$ is chosen to satisfy the backward Cauchy problem (7.1.32)–(7.1.33), and it then follows from the (modified) integral theorem (7.1.28) that the solution of the Cauchy problem (7.1.26)–(7.1.27) takes the form

$$u(\boldsymbol{\xi}, \tau) = \int_0^T \int_{\mathbf{x}-space} \rho K F \, d\mathbf{x} \, dt + \int_{\mathbf{x}-space} [\rho K f]\big|_{t=0} \, d\mathbf{x}. \quad (7.1.35)$$

The foregoing results are easily modified to yield Green's functions and solution formulas for initial and boundary value problems for hyperbolic and parabolic equations given over semi-infinite spatial regions.

Green's Functions for Hyperbolic and Parabolic Equations: An Alternative Construction

There is an alternative approach to the construction of Green's functions that applies in the hyperbolic and parabolic cases. Instead of having $K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)$ satisfy the inhomogeneous equations (7.1.23) and (7.1.32), we require that they be solutions of the homogeneous equations

$$\rho(\mathbf{x}) K_{tt}(\mathbf{x}, t; \boldsymbol{\xi}, \tau) + L K(\mathbf{x}, t; \boldsymbol{\xi}, \tau) = 0, \qquad \mathbf{x}, \boldsymbol{\xi} \in G, \ t < \tau, \quad (7.1.36)$$

$$-\rho(\mathbf{x}) K_t(\mathbf{x}, t; \boldsymbol{\xi}, \tau) + L K(\mathbf{x}, t; \boldsymbol{\xi}, \tau) = 0, \qquad \mathbf{x}, \boldsymbol{\xi} \in G, \ t < \tau, \quad (7.1.37)$$

in the hyperbolic and parabolic cases, respectively. The homogeneous initial conditions (7.1.24) and (7.1.33) are replaced by

$$K(\mathbf{x}, \tau; \boldsymbol{\xi}, \tau) = 0, \qquad K_t(\mathbf{x}, \tau; \boldsymbol{\xi}, \tau) = -\frac{\delta(\mathbf{x} - \boldsymbol{\xi})}{\rho(\mathbf{x})}, \qquad \boldsymbol{\xi} \in G, \quad (7.1.38)$$

$$K(\mathbf{x}, \tau; \boldsymbol{\xi}, \tau) = \frac{\delta(\mathbf{x} - \boldsymbol{\xi})}{\rho(\mathbf{x})}, \qquad \boldsymbol{\xi} \in G, \quad (7.1.39)$$

respectively. The boundary conditions (7.1.25) and (7.1.33) for $K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)$ are retained.

In this formulation we obtain

$$\int_{\partial R_\tau} [\rho K u_t - \rho u K_t]\, d\mathbf{x} = \int_{\partial R_\tau} u\, \delta(\mathbf{x} - \boldsymbol{\xi})\, d\mathbf{x} = u(\boldsymbol{\xi}, \tau), \tag{7.1.40}$$

$$\int_{\partial R_\tau} \rho u K\, d\mathbf{x} = \int_{\partial R_\tau} u\, \delta(\mathbf{x} - \boldsymbol{\xi})\, d\mathbf{x} = u(\boldsymbol{\xi}, \tau), \tag{7.1.41}$$

for the hyperbolic and parabolic cases, respectively, when (7.1.38) and (7.1.39) are used. The solutions $u(\boldsymbol{\xi}, \tau)$ then have the form (7.1.22) and (7.1.31) in the hyperbolic and parabolic cases as is easily seen. The only difference is that the domain of integration in the original formulation of the Green's function problem extends from 0 to T whereas in the present formulation it extends from 0 to $\tau < T$. However, since the equation and the data for $K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)$ are all homogeneous for $\tau < t < T$, the Green's function $K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)$ vanishes identically in that interval. Consequently, the domains of integration are, in effect, identical for both formulations. The relation between these two approaches is connected with *Duhamel's principle* (see Section 4.5), which relates inhomogeneous equations with homogeneous initial conditions to homogeneous equations with inhomogeneous initial conditions.

Integral Theorems and Green's Functions in One Dimension

The preceding results are valid in two or three space dimensions. The case of one dimension for the *elliptic equation* (7.1.1) with the boundary condition (7.1.2), leads to the consideration of a boundary value problem for an ordinary differential equation for $u(x)$. It is given as, for $0 < x < l$,

$$-\frac{d}{dx}\left(p(x)\frac{du(x)}{dx}\right) + q(x)u(x) = \rho(x)F(x), \quad \begin{cases} \alpha_1 u(0) - \beta_1 u'(0) = B_1, \\ \alpha_2 u(l) + \beta_2 u'(l) = B_2, \end{cases} \tag{7.1.42}$$

where α_1, β_1, α_2, β_2 satisfy the conditions given in Chapter 4. The related Green's function $K(x; \xi)$ satisfies the equation

$$LK(x;\xi) = -\frac{\partial}{\partial x}\left(p(x)\,\frac{\partial K(x;\xi)}{\partial x}\right) + q(x)K(x;\xi) = \delta(x-\xi), \qquad 0 < x,\, \xi < l, \tag{7.1.43}$$

with the boundary conditions

$$\alpha_1 K(0;\xi) - \beta_1 \partial K(0;\xi)/\partial x = 0, \quad \alpha_2 K(l;\xi) + \beta_2 \partial K(l;\xi)/\partial x = 0. \tag{7.1.44}$$

The solution of the boundary value problem (7.1.43)–(7.1.44) is expressed at a point ξ with $0 < \xi < l$ in terms of the solution formula (7.1.8) specialized to the one-dimensional case. The region G is the interval $0 < x < l$ and the S_i $(i = 1, 2, 3)$ correspond to boundary conditions of the first, second, or third kinds at $x = 0$ and

$x = l$. Thus the integral over G in (7.1.8) becomes an integral over $0 < x < l$. The integrals over the S_i reduce to the integrands evaluated at $x = 0$ or $x = l$. The normal derivative $\partial/\partial n$ equals $-\partial/\partial x$ at $x = 0$ and $\partial/\partial x$ at $x = l$. For example, in the case of mixed boundary conditions in (7.1.42) with $\alpha_1 \neq 0$, $\beta_1 = 0$, $\alpha_2 \neq 0$, $\beta_2 \neq 0$, we obtain the solution formula

$$u(\xi) = \int_0^l \rho(x)K(x;\xi)F(x)\,dx + \frac{p(0)}{\alpha_1} B_1 \frac{\partial K(0;\xi)}{\partial x} + \frac{p(l)}{\beta_2} B_2 K(l;\xi). \quad (7.1.45)$$

The one-dimensional versions of the *hyperbolic* and *parabolic equations* (7.1.11) and (7.1.26) lead to the consideration of the region R given as $[0,l] \times [0,T]$. The boundary ∂R is made up of the portion ∂R_x, which comprises the lines $x = 0$ and $x = l$ with $0 \leq t \leq T$, and ∂R_0 and ∂R_T, which represent the lines $t = 0$ and $t = T$, respectively, with $0 < x < l$. The region R is depicted in Figure 7.2.

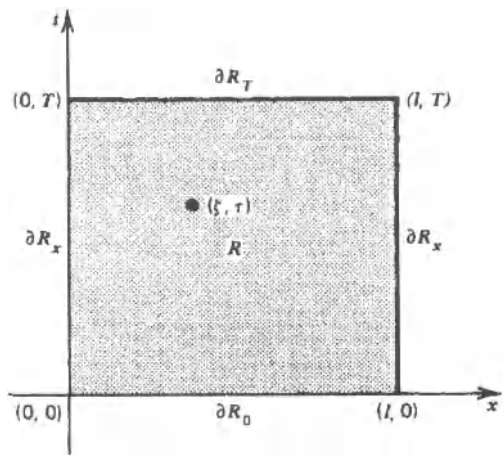

Figure 7.2 The region R.

For the *hyperbolic case* we have, with $0 < x < l$, $t > 0$,

$$\rho(x)\frac{\partial^2 u(x,t)}{\partial t^2} - \frac{\partial}{\partial x}\left(p(x)\frac{\partial u(x,t)}{\partial x}\right) + q(x)u(x,t) = \rho(x)F(x,t), \quad (7.1.46)$$

with the initial conditions

$$u(x,0) = f(x), \qquad u_t(x,0) = g(x), \qquad 0 < x < l, \quad (7.1.47)$$

and the boundary conditions

$$\alpha_1 u(0,t) - \beta_1 u_x(0,t) = g_1(t),\ \alpha_2 u(l,t) + \beta_2 u_x(l,t) = g_2(t). \quad (7.1.48)$$

With the operator L defined as in (4.1.7), we have

$$\int_0^T \int_0^l [K(\rho u_{tt} + Lu) - u(\rho K_{tt} + LK)]\, dx\, dt$$

$$= -\int_0^l [\rho K u_t - \rho u K_t]\big|_{t=0}\, dx + \int_0^T [-pKu_x + puK_x]\big|_{x=l}\, dt$$

$$+ \int_0^l [\rho K u_t - \rho u K_t]\big|_{t=T}\, dx - \int_0^T [-pKu_x + puK_x]\big|_{x=0}\, dt, \quad (7.1.49)$$

where the space-time gradient operator $\tilde{\nabla} = [\partial/\partial x, \partial/\partial t]$ and Green's theorem in the plane were used.

For the one-dimensional *parabolic case* we consider the equation

$$\rho(x)\frac{\partial u(x,t)}{\partial t} - \frac{\partial}{\partial x}\left(p(x)\,\frac{\partial u(x,t)}{\partial x}\right) + q(x)u(x,t) = \rho(x)F(x,t), \quad (7.1.50)$$

with $0 < x < l,\ t > 0$, the initial condition

$$u(x,0) = f(x), \quad (7.1.51)$$

and the boundary conditions (7.1.48). Proceeding as in the higher-dimensional case, with $\tilde{\nabla} = [\partial/\partial x, \partial/\partial t]$ as the space-time gradient operator, we obtain

$$\int_0^T \int_0^l [K(\rho u_t + Lu) - u(-\rho K_t + LK)]\, dx\, dt$$

$$= -\int_0^l [\rho K u]\big|_{t=0}\, dx + \int_0^T [-pKu_x + puK_x]\big|_{x=l}\, dt$$

$$+ \int_0^l [\rho K u]\big|_{t=T}\, dx - \int_0^T [-pKu_x + puK_x]\big|_{x=0}\, dt. \quad (7.1.52)$$

With $(\xi,\ \tau)$ as an interior point in the region R, we choose $K(x,t;\xi,\tau)$ to be a solution of

$$\rho(x)K_{tt}(x,t;\xi,\tau) + LK(x,t;\xi,\tau) = \delta(x-\xi)\delta(t-\tau), \qquad t < T, \quad (7.1.53)$$

$$-\rho(x)K_t(x,t;\xi,\tau) + LK(x,t;\xi,\tau) = \delta(x-\xi)\delta(t-\tau), \qquad t < T, \quad (7.1.54)$$

with $0 < x,\ \xi < l$, and $0 < t,\ \tau < T$ in the hyperbolic and parabolic problems, respectively. In addition, we have the end conditions

$$K(x,T;\xi,\tau) = 0, \qquad K_t(x,T;\xi,\tau) = 0, \quad (7.1.55)$$

in the former case and

$$K(x,T;\xi,\tau) = 0 \quad (7.1.56)$$

in the latter case. In both cases $K(x,t;\xi,\tau)$ is required to satisfy the homogeneous version of (7.1.48). If we have $\beta_1 = \beta_2 = 0$ in the boundary conditions, we obtain

$$\int_0^T [-pKu_x + puK_x]\big|_{x=l}\, dt = \int_0^T \frac{p(l)}{\alpha_2}\, g_2(t) K_x(l,t;\xi,\tau)\, dt, \tag{7.1.57}$$

$$\int_0^T [-pKu_x + puK_x]\big|_{x=0}\, dt = \int_0^T \frac{p(0)}{\alpha_1}\, g_1(t) K_x(0,t;\xi,\tau)\, dt. \tag{7.1.58}$$

Using these results, we easily obtain the following expression for the solution u at the arbitrary point $(\xi,\ \tau)$ in the hyperbolic case

$$\begin{aligned} u(\xi,\tau) = &\int_0^T \int_0^l \rho(x) F(x,t) K(x,t;\xi,\tau)\, dx\, dt \\ &+ \int_0^l \rho(x)[g(x)K(x,0;\xi,\tau) - f(x)K_x(x,0;\xi,\tau)]\, dx \\ &+ \int_0^T \left\{ \frac{p(0)g_1(t)}{\alpha_1} K_x(0,t;\xi,\tau) - \frac{p(l)g_2(t)}{\alpha_2} K_x(l,t;\xi,\tau) \right\} dt \end{aligned} \tag{7.1.59}$$

when $\beta_1 = \beta_2 = 0$ in the boundary conditions. For the parabolic case we have

$$\begin{aligned} u(\xi,\tau) = &\int_0^T \int_0^l \rho(x) F(x,t) K(x,t;\xi,\tau)\, dx\, dt + \int_0^l \rho(x) f(x) K(x,0;\xi,\tau)\, dx \\ &+ \int_0^T \left\{ \frac{p(0)g_1(t)}{\alpha_1} K_x(0,t;\xi,\tau) - \frac{p(l)g_2(t)}{\alpha_2} K_x(l,t;\xi,\tau) \right\} dt, \end{aligned} \tag{7.1.60}$$

with $\beta_1 = \beta_2 = 0$ in the boundary conditions.

If β_1 and β_2 are not zero or if there are mixed boundary conditions, a somewhat different expression for the solutions is readily obtained. Also, if the problem is given over a semi-infinite interval or we are dealing with the Cauchy problem over the infinite interval, appropriate expressions for the solution are easily found in a manner similar to that used previously in the higher-dimensional problems. Further, $K(x,t;\xi,\tau)$ can be characterized in an alternative manner as was done in (7.1.36)–(7.1.40) for higher dimensions.

Green's Functions for Nonself-Adjoint Elliptic Equations

The initial (or end) data for the Green's functions $K(\mathbf{x},\tau;\boldsymbol{\xi},\tau)$ in the hyperbolic and parabolic cases are assigned either at $t = T$ or at $t = \tau$. For the given problem for $u(\mathbf{x},t)$, however, the data are prescribed at $t = 0$. This results from the fact that the Green's function is the solution of an adjoint problem in which time runs backward rather than forward. The general form of the boundary conditions is the same for $K(\mathbf{x};\boldsymbol{\xi})$ and $u(\mathbf{x})$ in the elliptic problems, and for $K(\mathbf{x},\tau;\boldsymbol{\xi},\tau)$ and $u(\mathbf{x},t)$ in the hyperbolic and parabolic problems considered. This occurs because the elliptic

operator L that occurs in all three equations is *self-adjoint*, and it determines the choice of the boundary condition for the (adjoint) Green's function problem in each case. To see what happens if the elliptic (spatial) operator is *not self-adjoint*, we now consider a boundary value problem for a nonself-adjoint elliptic equation and determine the corresponding Green's function problem.

Consider the *elliptic equation*

$$\tilde{L}u(\mathbf{x}) = Lu(\mathbf{x}) + \mathbf{b}(\mathbf{x}) \cdot \nabla u(\mathbf{x}) = \rho(\mathbf{x})F(\mathbf{x}) \tag{7.1.61}$$

in a bounded region G, where L is defined in (7.1.1). With $\mathbf{b}(\mathbf{x}) \neq \mathbf{0}$, the operator $\tilde{L}$ is not self-adjoint. The boundary conditions for $u(\mathbf{x})$ are

$$\alpha(\mathbf{x})\, u(\mathbf{x}) + \beta(\mathbf{x}) \left.\frac{\partial u(\mathbf{x})}{\partial n}\right|_{\partial G} = B(\mathbf{x}), \tag{7.1.62}$$

as in the problem (7.1.1)–(7.1.2). The operator $\tilde{L}^*$, defined as $\tilde{L}^* u(\mathbf{x}) = Lu(\mathbf{x}) - \nabla \cdot (u(\mathbf{x})\mathbf{b}(\mathbf{x}))$, is the adjoint of L, as is easily checked, and we have

$$\begin{aligned} \iint_G [w\tilde{L}u - u\tilde{L}^* w]\, dv &= \iint_G \nabla \cdot [pu\nabla w - pw\nabla u + wu\mathbf{b}]\, dv \\ &= \int_{\partial G} \left[pu\frac{\partial w}{\partial n} - pw\frac{\partial u}{\partial n} + wu\mathbf{b} \cdot \mathbf{n} \right] ds \end{aligned} \tag{7.1.63}$$

on using the divergence theorem.

Proceeding as in the problem (7.1.1)–(7.1.2), we put $w(\mathbf{x}) = K(\mathbf{x}; \boldsymbol{\xi})$, the *Green's function*, and require that $K(\mathbf{x}; \boldsymbol{\xi})$ be a solution of

$$\tilde{L}^* K(\mathbf{x}; \boldsymbol{\xi}) = \delta(\mathbf{x} - \boldsymbol{\xi}), \qquad \mathbf{x},\ \boldsymbol{\xi} \in G. \tag{7.1.64}$$

Then $K(\mathbf{x}; \boldsymbol{\xi})$ must be specified on ∂G so that $u(\boldsymbol{\xi})$ is determined completely in terms of the boundary values for $u(\mathbf{x})$ and $K(\mathbf{x}; \boldsymbol{\xi})$. If $\mathbf{x} \in S_1$ on ∂G (see Section 4.1) we have

$$pu\frac{\partial K}{\partial n} - pK\frac{\partial u}{\partial n} + Ku\mathbf{b}\cdot\mathbf{n} = \frac{p}{\alpha}B\frac{\partial K}{\partial n} + K\left(u\mathbf{b}\cdot\mathbf{n} - p\frac{\partial u}{\partial n}\right) = \frac{p}{\alpha}B\frac{\partial K}{\partial n}, \tag{7.1.65}$$

if we set $K(\mathbf{x}; \boldsymbol{\xi}) = 0$. If $\mathbf{x} \in S_2 \cup S_3$ on ∂G, we have

$$pu\frac{\partial K}{\partial n} - pK\frac{\partial u}{\partial n} + Ku\mathbf{b}\cdot\mathbf{n} = -\frac{p}{\beta}BK + u\left(p\frac{\partial K}{\partial n} + \frac{p\alpha}{\beta}K + K\mathbf{b}\cdot\mathbf{n}\right) = -\frac{p}{\beta}BK \tag{7.1.66}$$

if we set $p(\mathbf{x})\partial K(\mathbf{x}; \boldsymbol{\xi})/\partial n + (p(\mathbf{x})\alpha(\mathbf{x})/\beta(\mathbf{x}))K(\mathbf{x}; \boldsymbol{\xi}) + K(\mathbf{x}; \boldsymbol{\xi})\mathbf{b}(\mathbf{x})\cdot\mathbf{n} = 0$. Thus, the (adjoint) boundary conditions for $K(\mathbf{x}; \boldsymbol{\xi})$ are

$$K(\mathbf{x};\boldsymbol{\xi}) = 0, \quad \mathbf{x} \in S_1, \tag{7.1.67}$$

$$\frac{p(\mathbf{x})}{\beta(\mathbf{x})}\left(\alpha(\mathbf{x})K(\mathbf{x};\boldsymbol{\xi}) + \beta(\mathbf{x})\frac{\partial K(\mathbf{x};\boldsymbol{\xi})}{\partial n}\right) + K(\mathbf{x};\boldsymbol{\xi})\mathbf{b}(\mathbf{x})\cdot\mathbf{n}, \quad \mathbf{x} \in S_2 \cup S_3.$$

In terms of the Green's function $K(\mathbf{x};\boldsymbol{\xi})$ that satisfies the adjoint equation (7.1.64) and the adjoint boundary conditions (7.1.67), the solution $u(\mathbf{x})$ of the problem (7.1.61)–(7.1.62) is given by the formula (7.1.8). We observe that only in the case of Dirichlet boundary conditions for $u(\mathbf{x})$ does the Green's function $K(\mathbf{x};\boldsymbol{\xi})$ satisfy the (homogeneous) Dirichlet condition $K(\mathbf{x};\boldsymbol{\xi}) = 0$ on the boundary. For boundary conditions of the second and third kind, $K(\mathbf{x};\boldsymbol{\xi})$ satisfies modified boundary conditions as given in (7.1.67) [unless $\mathbf{b}(\mathbf{x})\cdot\mathbf{n} = 0$ on the boundary]. The adjoint boundary conditions in the one-dimensional case are considered in the exercises.

If we replace the operator L in (7.1.11) and (7.1.26) (the *hyperbolic* and *parabolic cases*) by the operator $\tilde{L}$, but leave the initial and boundary conditions for these equations unchanged, we find that the Green's functions $K(\mathbf{x},t;\boldsymbol{\xi},\tau)$ for these problems are defined in terms of the adjoints of the modified equations. They have the same end conditions as before, but on the (lateral) boundary ∂R_x they satisfy (7.1.67).

Because of the greater difficulty involved in determining Green's functions for the foregoing nonself-adjoint problems, we restrict our discussion to problems with self-adjoint elliptic operators. However, before proceeding to construct Green's functions for various problems, we present the theory of generalized functions in the following section since these functions play an important role in Green's function theory.

Exercises 7.1

7.1.1. Show that for the elliptic case the Green's function $K(\mathbf{x};\boldsymbol{\xi})$ determined from (7.1.9)–(7.1.10) is symmetric [i.e., $K(\mathbf{x};\boldsymbol{\xi}) = K(\boldsymbol{\xi},\mathbf{x})$]. *Hint*: Let $u(\mathbf{x}) = K(\mathbf{x};\hat{\boldsymbol{\xi}})$ and $w(\mathbf{x}) = K(\mathbf{x};\boldsymbol{\xi})$ in (7.1.3).

7.1.2. Show that the Green's function for the hyperbolic problem (7.1.23)–(7.1.25) satisfies the equation $K(\mathbf{x},t;\boldsymbol{\xi},\tau) = K(\boldsymbol{\xi},-\tau;\mathbf{x},-t)$. *Hint*: Let $u(\mathbf{x},t) = K(\mathbf{x},-t;\hat{\boldsymbol{\xi}},-\hat{\tau})$ and $w(\mathbf{x},t) = K(\mathbf{x},t;\boldsymbol{\xi},\tau)$ in (7.1.14).

7.1.3. Show that the Green's function for the parabolic problem (7.1.32)–(7.1.33) satisfies $K(\mathbf{x},t;\boldsymbol{\xi},\tau) = K(\boldsymbol{\xi},-\tau;\mathbf{x},-t)$. *Hint*: Let $u(\mathbf{x},t) = K(\mathbf{x},-t;\hat{\boldsymbol{\xi}},-\hat{\tau})$ and $w(\mathbf{x},t) = K(\mathbf{x},t;\boldsymbol{\xi},\tau)$ in (7.1.28).

7.1.4. Let $F = 0$ in (7.1.8) and let B have delta function behavior with the singular point at $\mathbf{x} = \hat{\mathbf{x}}$ on the boundary so that (7.1.8) reduces to $u(\boldsymbol{\xi}) = -p/\alpha\,\partial K(\hat{\mathbf{x}};\boldsymbol{\xi})/\partial n$, $\hat{\mathbf{x}} \in S_1$, $u(\boldsymbol{\xi}) = p/\beta\,K(\hat{\mathbf{x}};\boldsymbol{\xi})$, $\hat{\mathbf{x}} \in S_2 \cup S_3$. Use these results to show that not only can the solution of (7.1.1)–(7.1.2) with $B = 0$ and $F \neq 0$ be expressed as the superposition of the solutions of point source or singularity problems, but the same can be done for (7.1.1)–(7.1.2) if $F = 0$ and $B \neq 0$. In the latter case the point sources lie on the boundary.

7.1.5. Use the expression (7.1.22) to characterize the solution of the initial and boundary value problem for the hyperbolic equation (7.1.11) as a superposition of solutions of point source problems, as in Exercise 7.1.4.

7.1.6. Use the expression (7.1.31) to characterize the solution of the initial and boundary value problem for the parabolic equation (7.1.26) as a superposition of solutions of point source problems.

7.1.7. Obtain an expression for the solution $u(\boldsymbol{\xi}, \tau)$ of the hyperbolic equation (7.1.11) based on the Green's function obtained from (7.1.36) and (7.1.39).

7.1.8. Using the Green's function determined from (7.1.37) and (7.1.40), obtain an expression for the solution $u(\boldsymbol{\xi}, \tau)$ of the parabolic equation (7.1.26).

7.1.9. Verify that $u(\xi, \tau)$ is given as in (7.1.59) if $\beta_1 = \beta_2 = 0$, and find the appropriate form for $u(\xi, \tau)$ if $\alpha_1 = \alpha_2 = 0$.

7.1.10. Verify the expression for $u(\xi, \tau)$ in (7.1.60) if $\beta_1 = \beta_2 = 0$, and obtain the correct form for $u(\xi, \tau)$ if $\alpha_1 = \alpha_2 = 0$.

7.1.11. Consider the initial and boundary value problem for the (damped) hyperbolic equation $\rho(\mathbf{x})u_{tt}(\mathbf{x}, t) + 2\lambda\rho(\mathbf{x})u_t(\mathbf{x}, t) + Lu(\mathbf{x}, t) = \rho(\mathbf{x})F(\mathbf{x}, t)$, where λ is a positive constant and L is given as in (7.1.1), with the initial and boundary conditions (7.1.12)–(7.1.13). Determine the appropriate Green's function problem and obtain a solution formula for $u(\mathbf{x}, t)$ that corresponds to (7.1.22).

7.1.12. Let $Mu(x, t) = u_{tt}(x, t) + c^2 u_{xxxx}(x, t)$. Express $wMu - uMw$ in divergence form and use this result as done in the text for the hyperbolic equation (7.1.11) to determine how a Green's function $K(x, t; \xi, \tau)$ should be constructed for the initial and boundary value problem for $Mu(x, t) = F(x, t)$. Here $u(x, t)$ and $u_t(x, t)$ are specified at $t = 0$ and, say, $u(x, t)$ and $u_x(x, t)$ are given on the boundary of the interval $0 < x < l$.

7.1.13. Generalize the result of Exercise 7.1.12 to $u_{tt}(\mathbf{x}, t) + c^2\nabla^2\nabla^2 u(\mathbf{x}, t) = F(\mathbf{x}, t)$. *Hint*: Use (8.5.38).

7.1.14. Consider the hyperbolic system $M\mathbf{u}(x, t) = \mathbf{u}_t(x, t) + A\mathbf{u}_x(x, t) + B\mathbf{u}(x, t) = \mathbf{c}(x, t)$, A, B constant, where A is a (real) symmetric matrix and B and $\mathbf{c}(x, t)$ are real valued. With M^* as the adjoint of M and B^T equal to the transpose of B, show that $\mathbf{w}^T\mathbf{u} - \mathbf{u}^T M^*\mathbf{w} = \mathbf{w}^T[\mathbf{u}_t + A\mathbf{u}_x + B\mathbf{u}] - \mathbf{u}^T[-\mathbf{w}_t - A\mathbf{w}_x + B^T\mathbf{w}]$ has the form of a divergence expression. Use the formula to construct a Green's matrix (or tensor) $\mathbf{K}(x, t; \xi, \tau)$ that is a solution of $M^*\mathbf{K}(x, t; \xi, \tau) = \delta(x-\xi)\delta(t-\tau)I$, $t < T$, $-\infty < x < \infty$, $\mathbf{K}(x, T; \xi, \tau) = \mathbf{0}$, where I is the identity matrix. Use $\mathbf{K}(x, t; \xi, \tau)$ to obtain a formula for the solution of the Cauchy problem for the given equation for $\mathbf{u}(x, t)$ with $\mathbf{u}(x, 0) = \mathbf{f}(x)$.

7.1.15. Determine the equation for the Green's function and the adjoint boundary conditions for the one-dimensional form of (7.1.61)–(7.1.62).

7.1.16. Let $\hat{K}(\mathbf{x}; \boldsymbol{\xi})$ be a solution of $\tilde{L}\hat{K}(\mathbf{x}; \boldsymbol{\xi}) = \delta(\mathbf{x} - \boldsymbol{\xi})$ [see (7.1.62)] with a homogeneous form of the boundary condition (7.1.63) given on ∂G. Show that $K(\mathbf{x}; \boldsymbol{\xi}) = \hat{K}(\boldsymbol{\xi}; \mathbf{x})$, where $K(\mathbf{x}; \boldsymbol{\xi})$ is the Green's function determined from (7.1.64) and (7.1.67).

7.2 GENERALIZED FUNCTIONS

Throughout the text and especially in connection with fundamental solutions and Green's functions, we have used the Dirac delta function in various calculations. Although the delta function is neither integrable nor differentiable in the conventional sense, we have integrated and differentiated this function. In this section, the Dirac delta and other functions are characterized as *generalized functions* and it is shown how the formal operations carried out on these functions are to be interpreted. Only the elementary and basic ideas of the theory of generalized functions are presented. Our discussion is restricted primarily to the one-dimensional case, as we shall see that the results for higher-dimensional problems can often be characterized in terms of the one-dimensional case. Many of the definitions and properties of generalized functions presented below carry over to higher dimensions. The approach we use was developed by *Laurent Schwartz* and defines generalized functions in terms of what are known as *linear functionals.*

Test Functions and Linear Functionals

We begin by considering the collection of *test functions* $\phi(x)$ that are assumed to have derivatives of all orders for all x (they are C^∞ functions) and vanish identically outside a bounded region. (Such functions are said to have *compact support*, and in the exercises it is shown that such functions exist.)

The generalized function $f(x)$ is defined with respect to the set of test functions $\phi(x)$ as the *linear functional* $(f, \phi(x))$. For each $\phi(x)$, the functional has a uniquely defined numerical value. Therefore, $f(x)$ is technically a function of a function that assumes a (numerical) value for each $\phi(x)$. So we call $(f, \phi(x))$ a *functional.* Strictly speaking, the generalized function $f(x)$ is not a function of x (i.e., it does not assume values for each x) but rather, of each test function $\phi(x)$. The $\phi(x)$ are called *test functions* since $f(x)$ is determined by specifying its effect on the test function $\phi(x)$. However, for the cases we consider, generalized functions can be characterized as ordinary functions of x for almost all values of x.

The functional $(f, \phi(x))$ is assumed to be *linear*; that is, if $\phi(x)$ and $\psi(x)$ are two test functions and a and b are constants,

$$(f, a\phi(x) + b\psi(x)) = a(f, \phi(x)) + b(f, \psi(x)). \tag{7.2.1}$$

In addition, $(f, \phi(x))$ is required to be *continuous* in the following sense. If $\{\phi_k(x)\}$ is a sequence of test functions, all of which vanish outside a common region and $\{\phi_k(x)\}$ converges (say, uniformly) to a function $\phi(x)$ as $k \to \infty$, we have

$$\lim_{k\to\infty} (f, \phi_k(x)) = (f, \phi(x)). \tag{7.2.2}$$

The collection of generalized functions $f(x)$ is a *linear (vector) space* in view of (7.2.1). Given a sequence of generalized functions $\{f_k(x)\}$ for which

$$\lim_{k\to\infty} (f_k, \phi(x)) = (f, \phi(x)), \tag{7.2.3}$$

we say that $f_k(x)$ converges to the generalized function $f(x)$ as $k \to \infty$. This type of convergence is called *weak convergence* and has been encountered earlier in the book in a different context.

A concrete general representation of the linear functional $(f, \phi(x))$ is given in the case where $f(x)$ is an (ordinary) *integrable function* over any finite region. We then associate the functional $(f, \phi(x))$ with an integral given as

$$(f, \phi(x)) = \int f(x)\phi(x)\, dx, \tag{7.2.4}$$

where the integration is carried out over the region where $\phi(x)$ is nonzero. The linearity of the functional follows from the linearity of the integral, and the continuity of the functional is not hard to show. In view of the representation (7.2.4), the reason for the use of the inner product notation $(f, \phi(x))$ for the functional becomes clear.

If $f(x)$ and $g(x)$ are two integrable functions such that $(f, \phi(x)) = (g, \phi(x))$ for all $\phi(x)$, we must have $f(x) = g(x)$ for almost all x. This follows since $(f, \phi(x)) - (g, \phi(x)) = \int (f(x) - g(x))\phi(x)\, dx = 0$. The vanishing of the integral for all test functions $\phi(x)$ implies (say, by using the fundamental lemma of the calculus of variations) that $f(x) = g(x)$ for almost all x. This shows that the functional representation (7.2.4) of $f(x)$ essentially specifies $f(x)$ uniquely if $f(x)$ is integrable. Thus with each integrable function a unique generalized function (f, ϕ) can be identified. Such generalized functions are often called *regular generalized functions*.

Generalized functions that are not ordinary integrable functions are often called *singular generalized functions*. The basic example is given by the *Dirac delta function*, $\delta(x)$ which is defined as

$$(\delta, \phi(x)) = \int \delta(x)\phi(x)\, dx = \phi(0). \tag{7.2.5}$$

[In higher dimensions with the test functions given as $\phi(\mathbf{x})$ we have $(\delta, \phi(\mathbf{x})) = \phi(\mathbf{0})$.] Although we have formally carried over the integral representation of $(f, \phi(x))$ given in (7.2.4) to this case, the delta function is not an integrable function. We recall that $\delta(x)$ was previously defined to be zero for all $x \neq 0$. Thus the improper integral in (7.2.5)—it is improper since $\delta(x)$ is assumed to be singular at zero—must vanish for all x. Consequently, since $\phi(0)$ need not vanish for all $\phi(x)$, we find that $\delta(x)$ is a singular generalized function. It may be noted that the integral representation (7.2.4) for singular generalized functions is not strictly valid, but it is generally used.

The generalized function $f(x)$ can be identified with (the values of) an ordinary function $g(x)$ in a region G if we have $(f, \phi(x)) = (g(x), \phi(x))$ for all test functions $\phi(x)$ that vanish outside G. We then say that $f(x) = g(x)$ in the region G. In particular, if $g(x) = 0$ in G, we say that $f(x)$ vanishes in the region G. As an example, we now show that $\delta(x)$ vanishes for all $x \neq 0$. We consider all test functions $\phi(x)$ that vanish in a neighborhood of $x = 0$, so that, in particular, $\phi(0) = 0$. Then, using the definition (7.2.5) of $\delta(x)$, we find that $(\delta, \phi(x)) = \phi(0) = 0 = (0, \phi(x))$. Thus $\delta(x) = 0$ for all $x \neq 0$.

In Example 1.1 we characterized the delta function as a limit of the sequence of integrable functions $f_k(x) = \begin{cases} k/2, & |x| < 1/k, \\ 0, & |x| > 1/k. \end{cases}$ We have

$$(f_k, \phi(x)) = \frac{k}{2}\int_{-1/k}^{1/k} \phi(x)\, dx = \frac{k\phi(\hat{x})}{2}\int_{-1/k}^{1/k} dx = \phi(\hat{x}) \tag{7.2.6}$$

on using the mean value theorem for integrals with $-1/k \leq \hat{x} \leq 1/k$. Since $\phi(x)$ is continuous at $x = 0$, we obtain $\lim_{k\to\infty}(f_k, \phi(x)) = \lim_{k\to\infty}\phi(\hat{x}) = \phi(0) = (\delta, \phi(x))$. Recalling our definition of the convergence of sequences of generalized functions, we conclude that the functions $f_k(x)$ converge to the delta function $\delta(x)$ as $k \to \infty$.

Properties of Generalized Functions

The properties of generalized functions are usually derived from results valid for the functional $(f, \phi(x))$ when $f(x)$ is a regular generalized function. They are then defined to be properties of generalized functions. Thus we have

$$(f(cx), \phi(x)) = \int_{-\infty}^{\infty} f(cx)\phi(x)\, dx = \begin{cases} (1/c)\int_{-\infty}^{\infty} f(x)\phi(x/c)\, dx, & c > 0, \\ -(1/c)\int_{-\infty}^{\infty} f_k(x)\phi(x/c)\, dx, & c < 0, \end{cases} \tag{7.2.7}$$

$$(f(x+\alpha), \phi(x)) = \int_{-\infty}^{\infty} f(x+\alpha)\phi(x)\, dx = \int_{-\infty}^{\infty} f(x)\phi(x-\alpha)\, dx, \tag{7.2.8}$$

with constant c and α. Therefore, we define for the generalized function $f(x)$

$$(f(cx), \phi(x)) = \frac{1}{|c|}\left(f(x), \phi\left(\frac{x}{c}\right)\right), \quad (f(x+\alpha), \phi(x)) = (f(x), \phi(x-\alpha)). \tag{7.2.9}$$

As a consequence of these definitions, we have for the delta function $\delta(x)$,

$$(\delta(cx), \phi(x)) = \frac{1}{|c|}\left(\delta(x), \phi\left(\frac{x}{c}\right)\right) = \frac{1}{|c|}\phi(0) = \left(\frac{1}{|c|}\delta(x), \phi(x)\right), \tag{7.2.10}$$

so that $\delta(cx) = (1/|c|)\delta(x)$ formally. In particular, if $c = -1$, we obtain $\delta(-x) = \delta(x)$, so that $\delta(x)$ may be said to be an *even function*. Similarly,

$$(\delta(x-y), \phi(x)) = (\delta(x), \phi(x+y)) = \phi(y), \tag{7.2.11}$$

which is known as the *substitution property* of $\delta(x)$ and is generally written as

$$\int \delta(x-y)\phi(x)) = \phi(y). \tag{7.2.12}$$

If $f(x)$ and $g(x)$ are regular functions, we have $\int (g(x)f(x))\phi(x)\,dx = \int f(x)(g(x)\phi(x))\,dx$, so that $(g(x)f(x), \phi(x)) = (f(x), g(x)\phi(x))$. If $g(x)$ is infinitely differentiable, $g(x)\phi(x)$ is a test function. We define the product of an infinitely differentiable function $g(x)$ and the generalized function $f(x)$ as

$$(g(x)f, \phi(x)) = (f, g(x)\phi(x)). \tag{7.2.13}$$

An example of this result is given by

$$g(x)\delta(x-y) = g(y)\delta(x-y), \tag{7.2.14}$$

which follows from $(g(x)\delta(x-y), \phi(x)) = (\delta(x-y), g(x)\phi(x)) = g(y)\phi(y) = (g(y)\delta(x-y), \phi(x))$, in view of (7.2.11). In particular, $x\delta(x) = 0 \cdot \delta(x) = 0$.

If $f(x)$ and $f'(x)$ are regular generalized functions, we obtain

$$(f'(x), \phi(x)) = \int_{-\infty}^{\infty} \phi(x)\,df = -\int_{-\infty}^{\infty} f(x)\phi'(x)\,dx = -(f(x), \phi'(x)), \tag{7.2.15}$$

since $\phi(x)$ vanishes as $|x| \to \infty$. If $f(x)$ has n derivatives, repeated integration by parts yields

$$\left(\frac{d^n f(x)}{dx^n}, \phi(x)\right) = (-1)^n \left(f(x), \frac{d^n \phi(x)}{dx^n}\right), \qquad n \geq 1. \tag{7.2.16}$$

We take (7.2.16) to be the definition of the nth derivative of the generalized function $f(x)$. Test functions $\phi(x)$ are assumed to be infinitely differentiable, so we conclude that any generalized function $f(x)$ has derivatives of all orders. In higher dimensions a similar result is valid for partial derivatives of $f(\mathbf{x})$.

As an example, we consider the function $f(x) = \begin{cases} 0, & x < 0, \\ x, & x \geq 0, \end{cases}$ which is continuous at $x = 0$ but is not differentiable there. To obtain the *generalized derivative* we have

$$(f'(x), \phi(x)) = -(f(x), \phi'(x)) = -\int_{-\infty}^{\infty} f(x)\phi'(x)\,dx = -\int_{0}^{\infty} x\phi'(x)\,dx$$

$$= -\int_{0}^{\infty} x\,d\phi = -x\phi\Big|_{0}^{\infty} + \int_{0}^{\infty} \phi(x)\,dx = \int_{0}^{\infty} \phi(x)\,dx. \tag{7.2.17}$$

Now the functional that represents the *Heaviside function* $H(x)$ defined as

$$H(x) = \begin{cases} 0, & x < 0, \\ 1, & x \geq 0, \end{cases} \tag{7.2.18}$$

is given by

$$(H(x), \phi(x)) = \int_{-\infty}^{\infty} H(x)\phi(x)\,dx = \int_{0}^{\infty} \phi(x)\,dx. \tag{7.2.19}$$

Thus (7.2.17) shows that $f'(x) = H(x)$, with $f(x)$ defined as above.

As $(H'(x), \phi(x)) = -(H(x), \phi'(x)) = -\int_0^\infty \phi'(x)\,dx = \phi(0) = (\delta(x), \phi(x))$, we conclude for the Heaviside function

$$H'(x) = \delta(x). \tag{7.2.20}$$

Further, we have

$$(\delta'(x), \phi(x)) = -(\delta(x), \phi'(x)) = -\phi'(0), \tag{7.2.21}$$

$$(\delta''(x), \phi(x)) = (\delta(x), \phi''(x)) = \phi''(0). \tag{7.2.22}$$

An additional result that has already been used in Section 6.7 is the following: $(x\delta'(x), \phi(x)) = (\delta'(x), x\phi(x)) = -(\delta(x), (x\phi(x))') = -(\delta(x), x\phi'(x)+\phi(x)) = -(x\phi'(x) + \phi(x))\big|_{x=0} = -\phi(0) = -(\delta(x), \phi(x))$. This may be formally expressed as the equation

$$x\,\delta'(x) + \delta(x) = 0. \tag{7.2.23}$$

In the following two examples we use the foregoing properties to construct generalized power functions and show how functions with jump discontinuities can be expressed as generalized functions.

Example 7.1. Generalized Power Functions. The functions x^m, where m is a negative integer, $x^a H(x)$, where a is any real number and $H(x)$ is the *Heaviside function*, are either ordinary power functions or the zero function, when $x \neq 0$. However, either the functions themselves or their derivatives (from some order on) are singular at $x = 0$. We introduce generalized functions associated with these functions that retain the basic properties of (nonsingular) power functions.

First we define the generalized functions

$$f_n(x) = x^n H(x), \qquad n = 0, 1, 2, \ldots. \tag{7.2.24}$$

If $n = 0$, this is just the Heaviside function, and for $x > 0$, it is identical with x^n. If $n > 0$, we obtain from (7.2.15)

$$(f_n'(x), \phi(x)) = (-1)\int_{-\infty}^{\infty} x^n H(x)\phi'(x)\,dx = (-1)\int_0^\infty x^n\phi'(x)\,dx$$

$$= -x^n\phi(x)\Big|_0^\infty + n\int_0^\infty x^{n-1}\phi(x)\,dx = (nf_{n-1}(x), \phi(x)), \tag{7.2.25}$$

so that $f_n'(x) = nf_{n-1}(x)$. Also, $xf_n(x) = f_{n+1}(x)$. We see that $f_n(x)$ retains the basic properties of the power function x^n (at $x \neq 0$), but it has generalized derivatives of all orders. The kth derivative of $f_n(x)$ is given by

$$\left(\frac{d^k f_n(x)}{dx^k}, \phi(x)\right) = (-1)^k\left(f_n(x), \frac{d^k\phi(x)}{dx^k}\right) = (-1)^k\int_0^\infty x^n\phi^{(k)}(x)\,dx \tag{7.2.26}$$

in view of (7.2.16).

Next we define the generalized functions

$$f_a(x) = x^a H(x), \qquad a > -1, \quad a \neq \text{integer}. \tag{7.2.27}$$

Since $a > -1$, the integral representation of $(f_a(x), \phi(x))$ converges. Thus, $f_a(x)$ has generalized derivatives of all orders. If $a < -1$, the integral representation of $f_a(x)$ diverges. In that case we define the generalized functions

$$f_a(x; k) = \frac{1}{(a+1)\cdots(a+k)} x^{a+k} H(x), \qquad a + k > -1, \tag{7.2.28}$$

where k can be chosen as the smallest integer such that $a + k > -1$, and a is not an integer. In terms of the $f_a(x; k)$, we define $f_a(x)$ as

$$f_a(x) = \frac{d^k}{dx^k}[f_a(x; k)], \tag{7.2.29}$$

with the kth generalized derivative of $f_a(x; k)$ understood to mean

$$(f_a(x), \phi(x)) = (f_a^{(k)}(x; k), \phi(x)) = \frac{(-1)^k}{(a+1)\cdots(a+k)} \int_0^\infty x^{a+k} \phi^{(k)}(x)\, dx \tag{7.2.30}$$

on using (7.2.16). (The integral converges since $a + k > -1$.) We note that for $a > -1$, we would set $k = 0$ in (7.2.28) and $f_a(x; 0) = f_a(x)$, so that the definitions (7.2.27) and (7.2.29) agree in that case. If $x > 0$, both (7.2.27) and (7.2.29) yield $f_a(x) = x^a$. It is shown in the exercises that $f_a'(x) = a f_{a-1}(x)$ and that $x f_a(x) = f_{a+1}(x)$.

The product of the function $1/x^n$, where n is a positive integer, with a test function $\phi(x)$ is not integrable at $x = 0$, in general. To deal with this problem, we first let $n = 1$ and define the generalized function $g_1(x) = 1/x$ as

$$(g_1(x), \phi(x)) = \left(\frac{1}{x}, \phi(x)\right) = \text{P.V.} \int_{-\infty}^{\infty} \frac{1}{x} \phi(x)\, dx, \tag{7.2.31}$$

where P.V. signifies that the Cauchy principal value of the integral is to be taken. In terms of $g_1(x)$, we define the generalized functions $g_n(x)$ as

$$g_n(x) = \frac{(-1)^{n-1}}{(n-1)!} \frac{d^{n-1}}{dx^{n-1}}[g_1(x)], \qquad n = 2, 3, \ldots. \tag{7.2.32}$$

This means that

$$(g_n(x), \phi(x)) = \frac{1}{(n-1)!} \text{P.V.} \int_{-\infty}^{\infty} \frac{1}{x} \phi^{(n-1)}(x)\, dx \tag{7.2.33}$$

on using (7.2.16). The relation between $g_n(x)$ and $1/x^n$ and some additional properties of these functions are considered in the exercises.

Finally we consider the function $(1/x^n)H(x)$, where n is a positive integer. The singularity at $x = 0$ is not integrable. It is not possible to define a unique generalized function associated with $(1/x^n)H(x)$ that retains the basic properties of the power function $1/x^n$, as shown in the exercises. We define the generalized functions $h_n(x)$ as

$$h_n(x) = \frac{(-1)^{n-1}}{(n-1)!}\frac{d^n}{dx^n}[(\log(x) + c)H(x)], \qquad n = 1, 2, 3, \ldots, \tag{7.2.34}$$

where the arbitrary constant c can be different for each n. [If $x > 0$, $h_n(x) = 1/x^n$.] We have

$$(h_n(x), \phi(x)) = \frac{-1}{(n-1)!}\int_0^\infty (\log(x) + c)\phi^{(n)}(x)\, dx. \tag{7.2.35}$$

Some additional properties of $h_n(x)$ are considered in the exercises.

The generalized power functions defined above can be used to assign a meaning to divergent improper integrals with algebraic singularities. In his treatment of fundamental solutions for hyperbolic equations Hadamard introduced the *method of finite parts* to evaluate divergent improper integrals with algebraic singularities. His analysis preceded the development of generalized functions as presented above. A connection between the foregoing discussion and Hadamard's method is considered in the exercises.

Example 7.2. The Derivative of a Discontinuous Function. Let the function $f(x)$ be continuously differentiable everywhere except at $x = \alpha$, where it has a jump discontinuity. Consequently, $f(x)$ does not have an ordinary derivative at $x = \alpha$, but the generalized function associated with $f(x)$ is differentiable everywhere and we now obtain its first derivative.

Using (7.2.15), we have

$$(f'(x), \phi(x)) = -(f(x), \phi'(x)) = -f(x)\phi(x)\Big|_{-\infty}^{\alpha} + \int_{-\infty}^{\alpha} f'(x)\phi(x)\, dx$$

$$-f(x)\phi(x)\Big|_{\alpha}^{\infty} + \int_{\alpha}^{\infty} f'(x)\phi(x)\, dx = [f(x)]_{x=\alpha}\phi(\alpha) + \int_{-\infty}^{\infty} f'(x)\phi(x)\, dx, \tag{7.2.36}$$

where $[f(x)]_{x=\alpha}$ is the jump in $f(x)$ at $x = \alpha$ and the derivative in the last integral is defined at all x except $x = \alpha$. Since $f'(x)$ is assumed to have a finite limit as x approaches α from the left and from the right, the value of $f'(x)$ at $x = \alpha$ plays no role in the integral. We can express $[f(x)]_{x=\alpha}\phi(\alpha)$ in the terms of the delta function in the form $([f(x)]_{x=\alpha}\delta(x-\alpha), \phi(x))$ and the last integral in (7.2.36) can be written as $(f'(x), \phi(x))$. Thus the *generalized derivative* f' of $f(x)$ can be expressed as

$$f' = [f(x)]_{x=\alpha}\delta(x-\alpha) + f'(x)|_{x\neq\alpha}. \tag{7.2.37}$$

It may be noted that (7.2.20) is a special case of this result.

An application of (7.2.37) may be made to the problem of determining the Green's function $K(x;\xi)$ for the ODE (7.1.42). We have

$$\frac{\partial}{\partial x}\left(p(x)\,\frac{\partial K(x;\xi)}{\partial x}\right) = -\delta(x-\xi) + q(x)K(x;\xi). \tag{7.2.38}$$

Using (7.2.37), we conclude that $q(x)K(x;\xi)$ is the ordinary derivative of $p(x)K(x;\xi)$ for $x \neq \xi$, and -1 is the jump of $p(x)K(x;\xi)$ at $x = \xi$. Since $p(x)$ is continuous at $x = \xi$, the jump in $\partial K(x;\xi)/\partial x$ at $x = \xi$ is $[\partial K(x;\xi)/\partial x]_{x=\xi} = -1/p(\xi)$, a result that agrees with (6.7.21).

The expression (7.2.37) for the generalized derivative of $f(x)$ suggests that we can express $f(x)$ as a sum of a continuous function and a piecewise constant function with a jump at $x = \alpha$. In fact, we can write

$$f(x) = \tilde{f}(x) + [f(x)]_{x=\alpha}H(x-\alpha). \tag{7.2.39}$$

Alternatively, we can set

$$f(x) = \hat{f}(x) - [f(x)]_{x=\alpha}H(\alpha-x). \tag{7.2.40}$$

The functions $\tilde{f}(x)$ and $\hat{f}(x)$ are given as the difference of $f(x)$ and the jump terms on the right of (7.2.39) and (7.2.40), respectively. Each function is continuous and has a jump in the first derivative at $x = \alpha$. Thus, the derivative functions $\tilde{f}'(x)$ and $\hat{f}'(x)$ are regular generalized functions, and the need for a generalized derivative of $f(x)$ is restricted to the terms that involve the Heaviside functions.

If $f(x)$ has derivatives of all orders for $x < \alpha$ and for $x > \alpha$ and $f(x)$ and its derivatives have (at most) jumps at $x = \alpha$, we can express $f(x)$ as the sum of a C^∞ function and a series of terms that accounts for the jumps at $x = \alpha$. To do so, we introduce the generalized functions

$$H_n(x) = \frac{1}{n!}f_n(x) = \frac{1}{n!}x^n H_n(x), \qquad n = 0, 1, \ldots, \tag{7.2.41}$$

given in terms of (7.2.24). Then we express $f(x)$ as

$$f(x) = \tilde{f}(x) + \sum_{n=0}^{\infty}[f^{(n)}(x)]_{x=\alpha}H_n(x-\alpha), \tag{7.2.42}$$

where we have used the form (7.2.39). The function $\tilde{f}(x)$ equals the difference between $f(x)$ and the sum on the right side of (7.2.42). [The jumps in the derivatives of $f(x)$ are expected to be known.] Noting that $H_n'(x) = H_{n-1}(x)$, we easily verify that $\tilde{f}(x)$ has derivatives of all orders.

The representation (7.2.42) is valid even if $f(x)$ is continuous and has continuous derivatives of order k at $x = \alpha$, but the derivatives of order $k+1$ and higher have jump discontinuities there. Additionally, if $f(x)$ has a finite (or even an infinite) number of jumps, the foregoing representation can be carried out at each singular point.

As an example we consider the function $f(x) = x$, $x < \alpha$, $f(x) = x^2$, $x > \alpha$. We easily obtain $f(x) = \tilde{f}(x) + \sum_{n=0}^{2}[f^{(n)}(x)]_{x=\alpha} H_n(x-\alpha) = x + (\alpha^2 - \alpha)H(x-\alpha) + (2\alpha - 1)H_1(x-\alpha) + 2H_2(x-\alpha)$. If $\alpha = 0$, $f(x)$ is continuous at $x = 0$ but has a jump in the first two derivatives there. Then the above takes the form $f(x) = x - H_1(x) + 2H_2(x)$. The Heaviside function $H(x)$ that determines the jump at $x = 0$ is absent from the expression.

The representation (7.2.42) and generalizations thereof are used in Section 10.2 in the analysis of the propagation of discontinuities and singularities of hyperbolic equations.

Fourier Transforms of Generalized Functions

In determining Green's functions for differential equations, it is often necessary to take transforms of the delta function. Therefore, we now discuss how Fourier transforms of generalized functions are to be defined. Since we are mostly concerned with transforms of the delta function and possibly its derivatives, we begin by considering generalized functions that vanish outside a bounded region.

For the one-dimensional problem, let the generalized function $f(x)$ vanish outside the interval $[-R, R]$ (i.e., it has compact support). It is possible to construct test functions $\hat{\phi}(x)$ with the property that $\hat{\phi}(x) = 1$ for $x \in [-R, R]$ and $\hat{\phi}(x) = 0$ outside an interval that contains $[-R, R]$. Assuming for the moment that $f(x)$ is a regular function, the Fourier transform $F(\lambda)$ of $f(x)$ is given as, since $f(x) = 0$ outside $[-R, R]$,

$$F(\lambda) = \frac{1}{\sqrt{2\pi}} \int_{-R}^{R} e^{i\lambda x} f(x)\, dx = \frac{1}{\sqrt{2\pi}} \int_{-R}^{R} e^{i\lambda x} f(x)\hat{\phi}(x)\, dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{i\lambda x} f(x)\hat{\phi}(x)\, dx = \left(f(x), \frac{1}{\sqrt{2\pi}} e^{i\lambda x}\hat{\phi}(x)\right). \tag{7.2.43}$$

Since $1/\sqrt{2\pi}\, e^{i\lambda x}\hat{\phi}(x)$ is again a test function, we define the *Fourier transform of the generalized function* $f(x)$ to be

$$F(\lambda) = \left(f(x), \frac{1}{\sqrt{2\pi}} e^{i\lambda x}\hat{\phi}(x)\right). \tag{7.2.44}$$

For example, the Fourier transform of the delta function $\delta(x)$ is given as

$$F(\lambda) = \left(\delta(x), \frac{1}{\sqrt{2\pi}} e^{i\lambda x}\hat{\phi}(x)\right) = \frac{1}{\sqrt{2\pi}}, \tag{7.2.45}$$

since $\hat{\phi}(0) = 1$ by assumption. This result is formally equivalent to

$$F(\lambda) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{i\lambda x}\delta(x)\, dx = \frac{1}{\sqrt{2\pi}}. \tag{7.2.46}$$

All the properties of Fourier transforms given in Section 5.2 are valid for this case. Higher dimensional Fourier transforms are defined in a similar way.

To define Fourier transforms for generalized functions without compact support we proceed as follows. Let $f(x)$ and $g(x)$ be regular real-valued functions whose Fourier transforms are $F(\lambda)$ and $G(\lambda)$, respectively. Then the Fourier transform of $g(-x)$ is given by $\overline{G(\lambda)}$, the complex conjugate of $G(\lambda)$. On setting $x = 0$ in (5.2.26), we conclude that

$$\int_{-\infty}^{\infty} F(\lambda)\overline{G(\lambda)}\, d\lambda = \int_{-\infty}^{\infty} f(x)g(x)\, dx. \tag{7.2.47}$$

Noting the formulas (5.2.6)–(5.2.7) relating the transform $F(\lambda)$ and its inverse transform $f(x)$, we see that $g(x)$ can be considered to be the transform of $\overline{G(\lambda)}$. Then if $\overline{G(\lambda)}$ is assumed to be a test function $\phi(\lambda)$ and we denote its transform by $\Phi(\lambda)$, (7.2.47) may be written as

$$\int_{-\infty}^{\infty} F(\lambda)\phi(\lambda)\, d\lambda = \int_{-\infty}^{\infty} f(x)\Phi(x)\, dx. \tag{7.2.48}$$

Clearly, both sides of the equation are linear functionals of the form (7.2.4) and we can write (7.2.48) as

$$(F, \phi) = (f, \Phi). \tag{7.2.49}$$

This is taken as the definition of the *Fourier transform* $F(\lambda)$ of the generalized function $f(x)$.

Unfortunately, even though every test function $\phi(x)$ has a Fourier transform, the transform is not itself a test function unless $\phi(x)$ vanishes identically. Thus the right side of (7.2.49) does not define a generalized function in general since $\Phi(x)$ is not necessarily a test function, so that (7.2.49) does not result in a meaningful definition of the Fourier transform for all generalized functions. To obtain a useful definition of the Fourier transform on the basis of (7.2.49), we must introduce a new class of test functions whose properties are preserved under Fourier transformation and, correspondingly, a new class of generalized functions.

Accordingly, we define the class of *test functions of rapid decay* that are required to be C^∞ and to vanish, together with all their derivatives, more rapidly than any negative power of $|x|$ as $|x| \to \infty$. The function $\exp(-x^2)$ belongs to this class of test functions but not to the previously defined collection of test functions since it does not have compact support. However, every test function with compact support belongs to this class of test functions. It is not hard to show that the Fourier transform of the present class of test functions is again a test function of the same class.

If $\phi(x)$ is a test function of rapid decay, the linear functional $(f, \phi(x))$ determines a *generalized function of slow growth*. The basic definitions and properties given for the previously defined generalized functions carry over to this class of generalized functions. Thus they have derivatives of all orders. (We do not discuss convergence of test functions and weak convergence of generalized functions for this case.) Any regular function $f(x)$ that is integrable over any finite interval and does not grow

more rapidly than any power of $|x|$ as $|x| \to \infty$, determines a (regular) generalized function of slow growth by means of the formula (7.2.4). Since our main interest lies in the definition of the Fourier transform for the class of generalized functions of slow growth, we do not discuss all of their properties. We conclude with the observation that if $f(x)$ is a generalized function of slow growth and $\phi(x)$ is a test function of rapid decay, then the linear functional on the right side of (7.2.49) determines a generalized function of slow growth that we define to be the *Fourier transform* $F(\lambda)$ of $f(x)$.

As an example, we consider the Fourier transform of the (generalized) function of slow growth $f(x) = 1$. The conventional Fourier transform is not defined for $f(x)$ and $f(x)$ does not have compact support so that the definition (7.2.44) cannot be used. Using (7.2.49) where $F(\lambda)$ is the Fourier transform of $f(x) = 1$, we have

$$(F, \phi) = (1, \Phi) = \int_{-\infty}^{\infty} \Phi(x)\, dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i\lambda x} \phi(\lambda)\, d\lambda\, dx$$

$$= \sqrt{2\pi} \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-i\lambda(t-x)} \phi(\lambda)\ dx\, d\lambda \right] \Bigg|_{t=0} = \sqrt{2\pi}\, \phi(0), \tag{7.2.50}$$

where we have used the Fourier integral formula (5.2.5) after interchanging the order of integration. Thus $(F, \phi) = \sqrt{2\pi}\, \phi(0) = \sqrt{2\pi}\ (\delta(x), \phi(x))$, so that the Fourier transform of $f(x) = 1$ equals $\sqrt{2\pi}\, \delta(x)$. This result can be obtained formally from the inversion formula for the Fourier transform (5.2.7), where the use of (7.2.45) gives

$$\delta(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\lambda x} \frac{1}{\sqrt{2\pi}}\, d\lambda. \tag{7.2.51}$$

Multiplying across by $\sqrt{2\pi}$ and applying complex conjugation on both sides yields the required relationship.

Weak Convergence of Series

In connection with the application of Fourier series and eigenfunction expansion methods for the solution of partial differential equations, we consider a further property of generalized functions that assigns a significance to such series even if they do not converge everywhere. Let $\sum_{n=1}^{\infty} g_n(x)$ be an infinite series of functions that converges uniformly in any bounded region, and let the sequence of partial sums of the series be defined as $f_k(x) = \sum_{n=1}^{k} g_n(x)$, $k = 1, 2, \ldots$. Then if $f(x)$ is the limit of the sequence $\{f_k(x)\}$, we have $\lim_{k\to\infty} (f_k(x), \phi(x)) = \lim_{k\to\infty} \int f_k(x)\phi(x)\, dx = \int \lim_{k\to\infty} f_k(x)\phi(x)\, dx = \int f(x)\phi(x)\, dx = (f(x), \phi(x))$, since the uniform convergence of the sequence $\{f_k(x)\}$ permits the interchange of the limit process and integration. Note that the vanishing of the test function $\phi(x)$ outside a bounded region implies that the integration is carried out only over that region so that the convergence is uniform.

The foregoing result shows that the sequence $\{f_k(x)\}$ and, consequently, the given series converges to the function $f(x)$ in the sense of weak convergence. Thus

$f(x)$ and the $f_k(x)$ can be interpreted as generalized functions. If each term $g_n(x)$ in the series has m derivatives, so does each term $f_k(x)$ in the sequence. Using (7.2.16), we have $(d^m f_k(x)/dx^m, \phi(x)) = (-1)^m (f_k(x), d^m\phi(x)/dx^m)$. Since the right side of this equation converges as $k \to \infty$ (as $d^m\phi(x)/dx^m$ has compact support), we find that the sequence $\{d^m f_k(x)/dx^m\}$ converges weakly, as $k \to \infty$, to the mth derivative of $f(x)$. In fact, $\lim_{k\to\infty}(d^m f_k(x)/dx^m, \phi(x)) = (-1)^m \lim_{k\to\infty}(f_k(x), d^m\phi(x)/dx^m) = (-1)^m(f(x), d^m\phi(x)/dx^m)$ and this equals $(d^m f(x)/dx^m, \phi(x))$.

As an example of the use of these results, we consider the Fourier sine series $F(x) = \sum_{n=1}^{\infty} a_n \sin(nx)$, where we assume that $|a_n| < M < \infty$ for all n. On differentiating the series $G(x) = -\sum_{n=1}^{\infty} a_n \sin(nx)/n^2$, twice term by term, we obtain the sine series above. Since the terms in the series $G(x)$ are majorized by M/n^2, the series is uniformly convergent. Thus even though the series for $F(x)$ may not converge pointwise everywhere, we conclude on the basis of the foregoing that it converges weakly to the generalized function $F(x)$, which represents the second generalized derivative of $G(x)$.

Example 7.3. Generalized Solutions of the Wave Equation. It was shown in Example 6.6 that the solution of the wave equation

$$u_{tt}(x,t) - c^2 u_{xx}(x,t) = 0, \qquad x > 0,\ t > 0 \tag{7.2.52}$$

in the semi-infinite interval $x > 0$, with the initial and boundary data

$$u(x,0) = 0, \quad u_t(x,0) = 0,\ x > 0, \quad u(0,t) = 1,\ t > 0, \tag{7.2.53}$$

is given in terms of the Heaviside function $H(x)$ as

$$u(x,t) = H\Big(t - \frac{x}{c}\Big). \tag{7.2.54}$$

Since $H(x)$ has derivatives of all orders in the generalized sense, we have $u_{tt} = H''(t - x/c) = \delta'(t - x/c)$, $u_{xx} = H''(t - x/c)/c^2 = \delta'(t - x/c)/c^2$. Thus (7.2.54) is formally a solution of the wave equation if we admit generalized functions as solutions. This interpretation of a generalized solution of the wave equation is related to the concept of weak solutions of Section 6.4.

We next consider the wave equation (7.2.52) in the finite interval $0 < x < l$ with the boundary conditions $u(0,t) = u(l,t) = 0$, $t > 0$, and the initial conditions $u(x,0) = 1$, $u_t(x,0) = 0$, $0 < x < l$. Using the results of Example 4.9, we obtain the (formal) Fourier sine series solution

$$u(x,t) = \sqrt{\frac{2}{l}} \sum_{k=1}^{\infty} a_k \cos\left(\frac{\pi k c}{l}\, t\right) \sin\left(\frac{\pi k}{l}\, x\right). \tag{7.2.55}$$

The Fourier coefficients a_k are given as

$$a_k = \sqrt{\frac{2}{l}} \int_0^l \sin\left(\frac{\pi k}{l}\, x\right) dx = \frac{\sqrt{2l}}{\pi k}(1 - (-1)^k) = \begin{cases} 2\sqrt{2l}/\pi k, & k = 1,3,\ldots, \\ 0, & k = 2,4,\ldots. \end{cases} \tag{7.2.56}$$

Thus (7.2.55) can be written as

$$u(x,t) = \frac{4}{\pi} \sum_{k=1}^{\infty}{}' \frac{1}{k} \cos\left(\frac{\pi k c}{l}\, t\right) \sin\left(\frac{\pi k}{l}\, x\right), \tag{7.2.57}$$

where the prime indicates that the summation is taken only over the odd numbers $k = 1, 3, 5, \ldots$.

In view of (4.4.22) we can write (7.2.57) as

$$u(x,t) = \frac{2}{\pi} \sum_{k=1}^{\infty}{}' \frac{1}{k} \sin\left[\frac{\pi k}{l}\,(x+ct)\right] + \frac{2}{\pi} \sum_{k=1}^{\infty}{}' \frac{1}{k} \sin\left[\frac{\pi k}{l}\,(x-ct)\right], \tag{7.2.58}$$

where each term in the series represents a propagating or progressive wave. As noted in Section 4.3, the Fourier sine series (7.2.55) evaluated at $t = 0$ is the odd periodic extension of the function $f(x) = 1$ defined in $0 < x < l$. The graph of the extended function is shown in Figure 7.3.

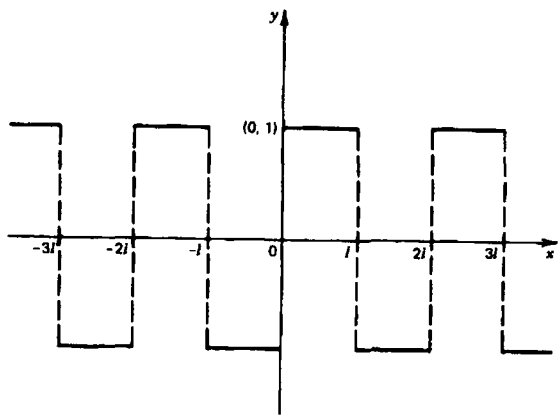

Figure 7.3 The odd periodic extension of $f(x) = 1$.

Let $F(x)$ denote the extended function of $f(x)$. It is seen that $F(x)$ has jumps of magnitude 2 at the points $x = nl$, $n = 0, \pm 1, \pm 2, \ldots$. Then (7.2.58) shows that $u(x,t)$ has the form

$$u(x,t) = \frac{1}{2}\, F(x+ct) + \frac{1}{2}\, F(x-ct). \tag{7.2.59}$$

The Fourier series does not converge pointwise to $F(x)$ at $x = nl$, $n = 0, \pm 1, \pm 2, \ldots$, and $F(x)$ is certainly not differentiable there. Consequently, (7.2.59) cannot be interpreted as a classical solution of the initial and boundary value problem for the wave equation.

However, each of the Fourier sine series in (7.2.58) converges in the generalized sense to the functions $\frac{1}{2}F(x \pm ct)$ in view of our discussion. In addition, they can

be differentiated twice to show that (7.2.59) is a generalized solution of the wave equation. For example, we have for $F(x+ct)$,

$$F(x+ct) = \frac{4}{\pi} \sum_{k=1}^{\infty}{}' \frac{1}{k} \sin\left[\frac{\pi k}{l}(x+ct)\right], \tag{7.2.60}$$

and on using Example 7.2 we obtain

$$\frac{\partial F(x+ct)}{\partial x} = \sum_{n=-\infty}^{\infty} 2(-1)^n \delta(x+ct-nl) = \frac{4}{l} \sum_{k=1}^{\infty}{}' \cos\left[\frac{\pi k}{l}(x+ct)\right], \tag{7.2.61}$$

since $F(x+ct) = \pm 1$ at $x+ct \neq nl$ and has jumps of magnitude 2 at $x+ct = nl$. A second derivative yields

$$\frac{\partial^2 F(x+ct)}{\partial x^2} = \sum_{n=-\infty}^{\infty} 2(-1)^n \delta'(x+ct-nl) = -\frac{4\pi}{l^2} \sum_{k=1}^{\infty}{}' k \sin\left[\frac{\pi k}{l}(x+ct)\right]. \tag{7.2.62}$$

Since $\delta'(x) = 0$ for $x \neq 0$, the series of derivatives of delta functions, as well as the sine series, in (7.2.62) vanishes for $x+ct \neq nl$. However, at $x+ct = nl$, each term in the sine series is zero. Thus it would appear that the sine series is identically zero and this is not so. [We have shown in (7.2.23) that $x\delta'(x) \neq 0$ but equals $-\delta(x)$.] This observation demonstrates that infinite series that are valid in a weak or generalized sense cannot always be dealt with in the same way as ordinary series. Again, this interpretation of the solution is closely related to that given in Example 6.5, where such Fourier series were characterized as weak solutions.

Properties of the Dirac Delta Function

The product of an ordinary function and a generalized function was defined in (7.2.13). Yet there does not seem to exist a useful definition of the ordinary product of two generalized functions. Thus we cannot define the product $\delta(x)\delta(x)$ or $H(x)H(x)$ in a consistent manner (see Exercise 7.2.19). Yet $\delta(x)\delta(x-a)$ with $a \neq 0$ is well defined since the singular points of both delta functions do not coincide. Thus at $x = 0$, $\delta(x-a)|_{x=0} = \delta(-a) = 0$ and at $x = a$, $\delta(x)|_{x=a} = \delta(a) = 0$. Consequently, we are effectively considering the product of an ordinary function and the delta function, and we have $\delta(x)\delta(x-a) = 0$.

In a similar fashion it is possible to consider the product of two generalized functions if they involve different variables. (This is sometimes called the *direct product*.) For example, the two-dimensional delta function $\delta(x,y)$ can be written as

$$\delta(x,y) = \delta(x)\delta(y). \tag{7.2.63}$$

Using the definition of the delta function (7.2.5) which is valid in any number of dimensions, we have

$$\iint_{(x,y)-plane} \delta(x,y)\phi(x,y)\,dx\,dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \delta(x)\delta(y)\phi(x,y)\,dx\,dy$$

$$= \int_{-\infty}^{\infty} \delta(y)\phi(0,y)\,dy = \phi(0,0) = (\delta(x,y), \phi(x,y)), \tag{7.2.64}$$

so that the two expressions in (7.2.63) are seen to be equivalent. Similarly, we obtain in three dimensions

$$\delta(x,y,z) = \delta(x)\delta(y)\delta(z). \tag{7.2.65}$$

By changing variables in the integral in (7.2.64) it is possible to obtain delta function representations in polar or other coordinate systems. This can also be done in the three-dimensional case.

In addition, it is possible to define *generalized composite functions* such as $\delta[g(x)]$, where $g(x)$ is assumed to be an ordinary function. The interpretation of these functions is based on the integral representation (7.2.4) and a change of variables, in the manner used in (7.2.7)–(7.2.8).

For example, if $g(x)$ is a smooth monotonic function that vanishes at $x = x_0$ and is such that $g'(x) > 0$ for all x, we have, formally,

$$(\delta[g(x)], \phi(x)) = \int_{-\infty}^{\infty} \delta[g(x)]\phi(x)\,dx = \int_{-\infty}^{\infty} \frac{\delta(\sigma)\phi[x(\sigma)]}{g'[x(\sigma)]}\,d\sigma$$

$$= \frac{\phi(x_0)}{g'(x_0)} = \left(\frac{1}{g'(x_0)}\delta(x - x_0), \phi(x)\right), \tag{7.2.66}$$

where we have used the transformation $\sigma = g(x)$ with $\sigma = 0$ corresponding to $x = x_0$. Thus $\delta[g(x)] = (1/g'(x_0))\delta(x - x_0)$. If $g'(x) < 0$ for all x, we find that the change of variables $\sigma = g(x)$ reverses the order of integration in (7.2.66). As a result, if $g'(x) > 0$ or $g'(x) < 0$ for all x with $g(x_0) = 0$ (and $x = x_0$ as the only zero), we obtain

$$\delta[g(x)] = \frac{1}{|g'(x_0)|}\,\delta(x - x_0). \tag{7.2.67}$$

If $g(x)$ has more than one zero and $g'(x) \neq 0$ at each of the zeros, we have, on applying the foregoing argument in the neighborhood of each zero (which we assume to be isolated),

$$\delta[g(x)] = \sum_n \frac{1}{|g'(x_n)|}\,\delta(x - x_n), \tag{7.2.68}$$

where we sum over all the zeros x_n.

Example 7.4. Properties of the Delta Function. We show that

$$\delta(ax+by)\delta(cx+dy)=\frac{1}{|ad-bc|}\,\delta(x)\delta(y). \tag{7.2.69}$$

Proceeding formally, we would consider the two-dimensional integral representation of the functional $(\delta(ax+by)\delta(cx+dy),\phi(x,y))$ and change the variables to obtain (7.2.69). Instead, we use a direct approach that treats the problem as one involving the product of two delta functions with different arguments.

We first assume that $a=0$ but b and c are not zero. Since $\delta(ax+by)|_{a=0}=\delta(by)=0$ for $y\neq 0$, on using (7.2.14) we find that $\delta(by)\delta(cx+dy)=\delta(by)\delta(cx)$. Then (7.2.10) implies that $\delta(by)\delta(cx)=(1/|bc|)\delta(x)\delta(y)$, and this yields (7.2.69) in the case $a=0$. If $a\neq 0$, we have

$$\begin{aligned}\delta(ax+by)\delta(cx+dy)&=\delta(ax+by)\delta\left[(-bc/a+d)\,y\right]\\&=\delta(ax+by)\frac{|a|}{|ad-bc|}\delta(y)=\frac{|a|}{|ad-bc|}\delta(ax)\delta(y)=\frac{1}{|ad-bc|}\delta(x)\delta(y),\end{aligned} \tag{7.2.70}$$

where we have used (7.2.14) and (7.2.10) several times.

In a similar fashion we may consider the generalized function $\delta[g(x)]$ in the case where $g(x)=0$ only at $x=x_0$ and $g'(x)\neq 0$ for all x. We have [since $g'(x_0)\neq 0$], on using Taylor's theorem with remainder,

$$\delta[g(x)]=\delta\left\{g'(x_0)(x-x_0)\left[1+\frac{1}{2}\frac{g''(\xi)}{g'(x_0)}(x-x_0)\right]\right\}. \tag{7.2.71}$$

Since $\delta[g(x)]$ is singular only at $x=x_0$, we may express it as

$$\delta[g(x)]=\delta[g'(x_0)(x-x_0)]=\frac{1}{|g'(x_0)|}\delta(x-x_0), \tag{7.2.72}$$

which is in agreement with (7.2.67).

Both results obtained show that taking account of the localized (point) singularity of the delta function can lead quickly to simplified expressions in problems that involve the Dirac delta function. The basic property we have used is that the delta function vanishes everywhere except where its argument equals zero. However, this approach must be used with caution when dealing with derivatives of delta functions.

We conclude our discussion of generalized functions by noting that *generalized functions* are often referred to as *distributions*. This is because generalized functions may be thought to represent the distribution of densities of masses, charges, or other physical quantities. In this context, the delta function $\delta(x)$ has infinite density at $x=0$ but $(\delta(x),1)=\int_{-\infty}^{\infty}\delta(x)\,dx=1$, so that the total amount of mass, say, is unity. Similarly, we may consider the derivative of the delta function $\delta'(x)$ to represent the density of a dipole located at $x=0$. The total charge is given as

$(\delta'(x), 1) = -(\delta(x), d(1)/dx) = -(\delta(x), 0) = 0$, so that the total charge vanishes as expected. However, the moment of the dipole around $x = 0$ is given as $\int_{-\infty}^{\infty} x\delta'(x)\,dx = (\delta'(x), x) = -(\delta(x), dx/dx) = -(\delta(x), 1) = -1$. Thus $-\delta'(x)$ is a dipole distribution located at $x = 0$ with a unit moment.

Further properties of one- or higher-dimensional generalized functions will be considered as needed in our discussion.

Exercises 7.2

7.2.1. Consider the sequence of functions $f_k(x, y, z) = \begin{cases} 3k^2/4\pi, & r < 1/k, \\ 0, & r > 1/k, \end{cases}$ where $r^2 = x^2 + y^2 + z^2$. Let $\phi(x, y, z)$ be a test function. Show that

$$\lim_{k\to\infty}(f_k, \phi) = \lim_{k\to\infty} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_k(x, y, z)\phi(x, y, z)\,dx\,dy\,dz = \phi(0, 0, 0),$$

so that the sequence $f_k(x, y, z)$ converges to the delta function $\delta(x, y, z)$ in three dimensions.

7.2.2. Show that the functions $\phi_a(\mathbf{x}) = \begin{cases} \exp[-a^2/(a^2 - |\mathbf{x}|^2)], & |\mathbf{x}| < a, \\ 0, & |\mathbf{x}| \geq a, \end{cases}$ are test functions. (The result is valid in one dimension as well.) That is, they are infinitely differentiable and vanish outside the region $|\mathbf{x}| \leq a$.

7.2.3. (a) Show that the generalized derivative of $|x|$ is sgn x, the signum function, defined as sgn $x = -1$ for $x < 0$ and sgn $x = 1$ for $x > 0$. (b) Show that the generalized derivative of sgn x equals $2\delta(x)$.

7.2.4. Show that $x^n\delta^{(m)}(x) = 0$ if $m < n$, and if $m \geq n$ we have $x^n\delta^{(m)}(x) = (-1)^n[m!/(m - n)!]\delta^{(m-n)}(x)$.

7.2.5. If $f_a(x)$ is the power function defined in (7.2.30), show that $f_a'(x) = af_{a-1}(x)$ and $xf_a(x) = f_{a+1}(x)$.

7.2.6. Show that $xg_1(x) = 1$ and that $xg_n(x) = g_{n-1}(x)$ for $n \geq 2$, where $g_n(x)$ is the generalized power function (7.2.31)–(7.2.32).

7.2.7. Consider the product $xh_n(x)$, with $h_n(x)$ defined as in (7.2.34). Express this product as $(xh_n(x), \phi(x)) = (h_n(x), x\phi(x))$. Use (7.2.35) with $n = 2$ and $\phi(x)$ replaced by $x\phi(x)$ to conclude that $xh_2(x) \neq h_1(x)$ unless we choose two different constants in the definition (7.2.34).

7.2.8. Show that $h_n'(x) = -nh_{n+1}(x)$ for $h_n(x)$ defined by (7.2.34).

7.2.9. Determine the generalized derivative of the following functions: (a) $f(x) = H(x - 1)\sin x$; (b) $f(x) = x^2\delta(x)$; (c) $f(x) = e^x\delta'(x + 3) - H(x)\cos^2 x$; (d) $f(x) = H(x - 1)\log x$; (e) $f(x) = e^{|x|}$.

7.2.10. Obtain the formal Fourier sine and cosine series of the delta function $\delta(x - x_0)$, $\delta(x - x_0) = (2/l)\sum_{k=1}^{\infty} \sin(\pi k x_0/l)\sin(\pi k x/l)$, and $\delta(x - x_0) = (1/l) + (2/l)\sum_{k=1}^{\infty} \cos(\pi k x_0/l)\cos(\pi k x/l)$, where $0 < x, x_0 < l$. Show that these series are convergent in the generalized sense. *Hint*: Consider test functions $\phi(x)$ that vanish outside a closed subinterval of $(0, l)$. Show that the sine and cosine series converge weakly to $\phi(x_0)$.

7.2.11. Express $\delta(x^2 - a^2)$ as a sum of delta functions in the manner of (7.2.68).

7.2.12. Express $\delta(\sin x)$ as a series of delta functions by using (7.2.68).

7.2.13. Show that the sequence of functions $f_\epsilon(x) = (1/\sqrt{4\pi\epsilon})\exp(-x^2/4\epsilon)$ tends to $\delta(x)$ as $\epsilon \to 0$.

7.2.14. Show that $\lim_{\epsilon\to 0} \frac{H(x+\epsilon)-H(x)}{\epsilon} = \delta(x)$ in the sense of weak convergence.

7.2.15. Let $f(x)$ be a positive continuous function in $[a, b]$ that vanishes outside $[a, b]$ and has unit area under the curve. Show that $\lim_{\epsilon\to 0}(1/\epsilon)f(x/\epsilon) = \delta(x)$ in the weak sense.

7.2.16. Show that $\delta'(-x) = -\delta'(x)$.

7.2.17. Let $f(x)$ be defined as $\delta(x)$ in the interval $-\pi < x < \pi$ and extend it as a periodic function of period 2π. Show that the extended function $\hat{f}(x)$ can be expressed as $\hat{f}(x) = \sum_{k=-\infty}^{\infty} \delta[x - 2k\pi]$ and that it has the generalized (complex) Fourier series $\hat{f}(x) = (1/2\pi)\sum_{k=-\infty}^{\infty} \exp(ikx)$.

7.2.18. Express the integral in (7.2.31) as $(1/x, \phi(x)) = \int_0^\infty \frac{\phi(x)-\phi(-x)}{x}\,dx$, and show that the generalized derivative of $g(x) = \log|x|$ is given by the generalized function $g_1(x) = 1/x$.

7.2.19. Noting that $x\delta(x) = 0$, show that $0 = (0, \phi(x)) = (x\delta(x)(1/x), \phi(x)) = (\delta(x), \phi(x)) = \phi(0)$, where $1/x$ stands for $g_1(x) = 1/x$, the generalized function defined in equation (7.2.31). Conclude from this that if the product of the generalized functions $x\delta(x)$ and $1/x$ were defined, we would have the result $0 = \delta(x)$.

7.2.20. Show that the function $F(x)$ in Example 7.3 can be expressed as $F(x) = 1 + \sum_{n=-\infty}^{0} 2(-1)^{n+1}H(nl - x) + \sum_{n=1}^{\infty} 2(-1)^n H(x - nl)$.

7.2.21. Given the divergent integral $\int_0^c x^a\phi(x)\,dx$, $-2 < a < -1$, where $\phi(0) \neq 0$, its *finite part* is defined by Hadamard to be $\int_0^c x^a[\phi(x)-\phi(0)]\,dx + c^{a+1}/(a+1)\phi(0)$, (We do not consider his definition of the finite part for the case where $a < -2$ and nonintegral.) Use the generalized function $f_a(x)$ defined in (7.2.29) to obtain $\int_0^c x^a\phi(x)\,dx = -1/(a+1)\int_0^c x^{a+1}\phi'(x)\,dx + c^{a+1}/(a+1)\,\phi(c)$. Integrate by parts once in Hadamard's finite part formula to show that the two expressions for the value of the divergent integral are equivalent.

7.2.22. Determine the finite part of the integral $\int_0^c (1+x)/x^{3/2}\,dx$.

7.2.23. Given a (piecewise smooth) surface S, consider the generalized function $F(x,y,z)\delta(S)$ that vanishes at all points P not on the surface S and has the property that $(F(x,y,z)\delta(S), \phi(x,y,z)) = \int_S F(x,y,z)\phi(x,y,z)\,ds$. The function $F(x,y,z)\delta(S)$ is called a *single-layer distribution* and represents a generalization of the one-dimensional Dirac delta function. Show that if $f(x,y,z)$ is a piecewise smooth function with a jump discontinuity across the surface S, then $\partial f(x,y,z)/\partial x = \partial f(x,y,z)/\partial x\big|_{P\notin S} + [f(x,y,z)]_S\,\mathbf{n}\cdot\mathbf{i}\,\delta(S)$, where $\mathbf{n}$ is a unit normal to the surface S and $[f(x,y,z)]_S$ is the jump in $f(x,y,z)$ across S, given as the difference between the value of $f(x,y,z)$ on the side of S that $\mathbf{n}$ points to and the side it points away from. Also, obtain the corresponding formula for the other two partial derivatives of $f(x,y,z)$. *Hint*: Use the definition of the generalized (partial) derivative of $f(x,y,z)$ and the divergence theorem.

7.2.24. With the single-layer distribution $F(x,y,z)\delta(S)$ given as in Exercise 7.2.23, the *double-layer distribution* is given as $-\partial/\partial n(F(x,y,z)\delta(S))$, where $\partial/\partial n$ is the normal derivative on S and $(-\partial(F(x,y,z)\delta(S))/\partial n), \phi(x,y,z)) = \int_S F(x,y,z)\partial\phi(x,y,z)/\partial n\,ds$. This corresponds to a generalization of $-\delta'(x)$. Demonstrate that $\nabla^2 f(x,y,z) = \nabla^2 f(x,y,z)\big|_{P\notin S} + \partial([f(x,y,z)]_S\delta(S))/\partial n + [\partial f(x,y,z)/\partial n]_S\,\delta(S)$. Single- and double-layer distributions can be used to describe some of the results in Chapter 6 relating to concentrated source terms in the language of the theory of generalized functions.

7.2.25. Show that (a) $(1/2\pi)\int_{-\infty}^{\infty}\exp(i\lambda x)\exp(-i\hat{\lambda}x)\,dx = \delta(\lambda-\hat{\lambda})$; (b) $(2/\pi)\int_0^{\infty}\cos(\lambda x)\cos(\hat{\lambda}x)\,dx = \delta(\lambda-\hat{\lambda})$, $\lambda, \hat{\lambda} \geq 0$; (c) $(2/\pi)\int_0^{\infty}\sin(\lambda x)\sin(\hat{\lambda}x)dx = \delta(\lambda-\hat{\lambda})$, $\lambda, \hat{\lambda} \geq 0$. Demonstrate that these results can be interpreted as orthogonality conditions for the functions arising in the Fourier transform, the Fourier cosine transform, and the Fourier sine transform. In part (a) the Hermitian inner product must be used.

7.3 GREEN'S FUNCTIONS FOR BOUNDED REGIONS

A general procedure for determining Green's functions for problems given over bounded (spatial) regions is the method of finite Fourier transforms presented in Section 4.6. For the equations (7.1.9), (7.1.23), and (7.1.32) given in Section 7.1, the Green's function is expanded in a series of eigenfunctions of the elliptic operator L defined in (7.1.1). The coefficients in the series of eigenfunctions are specified in the manner shown in Section 4.6. Although this procedure is identical to that given in Section 4.6, the solutions of the given problems expressed in terms of the Green's functions have a somewhat different form than that given earlier. However, the uniqueness theorems guarantee that there can be only one solution for each of these problems. We do not demonstrate in the general case the equivalence of the various solution forms. Our main emphasis in this section lies in the construction of Green's functions.

Green's Functions for Elliptic PDEs

We begin by constructing the Green's function $K(\mathbf{x};\boldsymbol{\xi})$ for the elliptic problem. As shown in Section 7.1, the function $K(\mathbf{x};\boldsymbol{\xi})$ satisfies the equation

$$LK(\mathbf{x};\boldsymbol{\xi}) \equiv -\nabla\cdot(p(\mathbf{x})\nabla K(\mathbf{x};\boldsymbol{\xi})) + q(\mathbf{x})K(\mathbf{x};\boldsymbol{\xi}) = \delta(\mathbf{x}-\boldsymbol{\xi}), \quad \mathbf{x},\ \boldsymbol{\xi}\in G \tag{7.3.1}$$

and the boundary condition

$$\alpha(\mathbf{x})K(\mathbf{x};\boldsymbol{\xi}) + \beta(\mathbf{x})\frac{\partial K(\mathbf{x};\boldsymbol{\xi})}{\partial n}\bigg|_{\partial G} = 0, \tag{7.3.2}$$

where derivatives are taken with respect to the variable $\mathbf{x}$. Let $K(\mathbf{x};\boldsymbol{\xi})$ be the orthonormalized set of eigenfunctions of the operator L; that is,

$$LM_k(\mathbf{x}) = \lambda_k\rho(\mathbf{x})M_k(\mathbf{x}), \qquad k = 1,2,\ldots, \tag{7.3.3}$$

where the $M_k(\mathbf{x})$ are the eigenvalues of L and $\rho(\mathbf{x})$ is a given weight function. The boundary condition for the $M_k(\mathbf{x})$ is (7.3.2) with $K(\mathbf{x};\boldsymbol{\xi})$ replaced by $M_k(\mathbf{x})$.

We express $K(\mathbf{x};\boldsymbol{\xi})$ as a series of eigenfunctions

$$K(\mathbf{x};\boldsymbol{\xi}) = \sum_{k=1}^{\infty} N_k(\boldsymbol{\xi})M_k(\mathbf{x}), \tag{7.3.4}$$

as in (4.6.3) with the (Fourier) coefficients $N_k(\boldsymbol{\xi})$ to be determined. Proceeding as in Section 4.6, we multiply (7.3.1) by $M_k(\mathbf{x})$, $(k = 1,2,\ldots)$ and integrate over the region G. [Note that the operator K introduced in Section 4.6 is unrelated to the Green's function $K(\mathbf{x};\boldsymbol{\xi})$ of the present chapter.] Using the results (4.6.5)–(4.6.8) and noting that both $K(\mathbf{x};\boldsymbol{\xi})$ and $M_k(\mathbf{x})$ satisfy homogeneous boundary conditions of the form (7.3.2), we obtain

$$\lambda_k N_k(\boldsymbol{\xi}) = \iint_G \delta(\mathbf{x}-\boldsymbol{\xi})M_k(\mathbf{x})\,dx = M_k(\boldsymbol{\xi}), \qquad k = 1,2,\ldots, \tag{7.3.5}$$

since $\boldsymbol{\xi}\in G$. Then if all the $\lambda_k > 0$, we have $N_k(\boldsymbol{\xi}) = M_k(\boldsymbol{\xi})/\lambda_k$, and this yields the *bilinear expansion* of the Green's function $K(\mathbf{x};\boldsymbol{\xi})$,

$$K(\mathbf{x};\boldsymbol{\xi}) = \sum_{k=1}^{\infty}\frac{M_k(\mathbf{x})M_k(\boldsymbol{\xi})}{\lambda_k}. \tag{7.3.6}$$

We notice that $K(\mathbf{x};\boldsymbol{\xi}) = K(\boldsymbol{\xi};\mathbf{x})$, so that the Green's function for the elliptic problem is *symmetric*. This result can be proven directly for the Green's function without the use of the bilinear expansion (see Exercise 7.1.1). The symmetry is a consequence of the fact that the operator $(1/\rho)L$ taken together with the boundary conditions is *self-adjoint*. It implies that the interchange of the *source point* $\boldsymbol{\xi}$ and the *observation point* $\mathbf{x}$ does not alter the solution.

It was shown in Section 4.2 that the eigenvalues λ_k for the eigenvalue problem (7.3.3) with the boundary condition (7.3.2) [with $K(\mathbf{x};\boldsymbol{\xi})$ replaced by $M_k(\mathbf{x})$] are nonnegative. It has also been demonstrated in Exercise 4.2.5 that $\lambda_0 = 0$ is an eigenvalue for this problem if and only if $q(\mathbf{x}) = 0$ in the operator L and $\alpha(\mathbf{x}) = 0$ in the boundary condition. The corresponding eigenfunction $M_0(\mathbf{x})$ is clearly a constant in that case. Consequently, unless zero is an eigenvalue, the bilinear expansion (7.3.6) yields a formal solution of the boundary value problem (7.3.1)–(7.3.2) for the Green's function $K(\mathbf{x};\boldsymbol{\xi})$. It may be verified directly that (7.3.6) is a (generalized) solution of (7.3.1)–(7.3.2).

In the one-dimensional case, (7.3.1) is an ODE and the bilinear expansion (7.3.6) represents an expansion of the Green's function $K(x;\xi)$ in terms of the eigenvalues and eigenfunctions of a Sturm-Liouville problem. In that case, it is often preferable to determine the Green's function by solving the differential equation directly. Additionally, for certain higher-dimensional problems it is possible to construct the Green's function in terms of eigenfunctions for lower-dimensional problems. Since the solution of higher-dimensional eigenvalue problems is generally not a simple task, this can lead to a substantial simplification of the Green's function problem, especially if the resulting eigenvalue problems are one-dimensional.

Proceeding as in Section 4.2, we consider the function $u = u(\mathbf{x}, y)$ where $\mathbf{x}$ is a point in the (bounded) region G and y is a scalar variable defined over the interval $0 < y < \hat{l}$. In place of (7.1.1) we now consider the equation

$$\rho(\mathbf{x})u_{yy}(\mathbf{x}, y) - Lu(\mathbf{x}, y) = -\rho(\mathbf{x})F(\mathbf{x}, y), \qquad \mathbf{x} \in G,\ 0 < y < \hat{l}, \tag{7.3.7}$$

and the boundary conditions (7.1.2) and (4.2.8) if G is a two-dimensional region. If G is one-dimensional and is given as $0 < x < l$, $u(x, y)$ satisfies the boundary conditions (7.1.47), where t is replaced by y.

The Green's function for the problem satisfies the equation

$$\rho(\mathbf{x})\frac{\partial^2 K(\mathbf{x}, y; \boldsymbol{\xi}, \eta)}{\partial y^2} - LK(\mathbf{x}, y; \boldsymbol{\xi}, \eta) = -\delta(\mathbf{x}-\boldsymbol{\xi})\delta(y-\eta),\ \mathbf{x}, \boldsymbol{\xi} \in G,\ 0 < y, \eta < \hat{l}, \tag{7.3.8}$$

and the boundary condition (7.3.2)—if G is two-dimensional—as well as a homogeneous form of the boundary condition (4.2.8). The differential operator in (7.3.7) is self-adjoint. The Green's function is expressed in the form $K(\mathbf{x}, y; \boldsymbol{\xi}, \eta)$ and the derivatives in (7.3.8) are taken with respect to the variables $\mathbf{x}$ and y. If the region G is one-dimensional, $K(x, y; \xi, \eta)$ satisfies the boundary conditions (7.1.47) (where t is replaced by y) with $u(x, y)$ replaced by $K(\mathbf{x}, y; \boldsymbol{\xi}, \eta)$ and $g_1(y) = g_2(y) = 0$.

To determine $K(\mathbf{x}, y; \boldsymbol{\xi}, \eta)$ we use the finite Fourier transform procedure of Section 4.6. With the set $\{M_k(\mathbf{x})\}$ $(k = 1, 2, \ldots)$ as the eigenfunctions of the operator L in (7.3.8), we construct the eigenfunction expansion

$$K(\mathbf{x}, y; \boldsymbol{\xi}, \eta) = \sum_{k=1}^{\infty} N_k(y)M_k(\mathbf{x}). \tag{7.3.9}$$

The $N_k(y)$ are the Fourier coefficients of $K(\mathbf{x}, y; \boldsymbol{\xi}, \eta)$ given as

$$N_k(y) = (K(\mathbf{x}, y; \boldsymbol{\xi}, \eta), M_k(\mathbf{x})) = \iint_G \rho(\mathbf{x}) K(\mathbf{x}, y; \boldsymbol{\xi}, \eta) M_k(\mathbf{x})\, d\mathbf{x},\ k \geq 1, \tag{7.3.10}$$

assuming that G is two-dimensional. We multiply (7.3.8) by $M_k(\mathbf{x})$ and integrate over the region G. Again using the results (4.6.5)–(4.6.8), we obtain

$$N_k''(y) - \lambda_k N_k(y) = -M_k(\boldsymbol{\xi})\delta(y - \eta),\ 0 < y,\ \eta < \hat{l},\ k = 1, 2, \ldots. \tag{7.3.11}$$

Since $K(\mathbf{x}, 0; \boldsymbol{\xi}, \eta) = K(\mathbf{x}, \hat{l}; \boldsymbol{\xi}, \eta) = 0$, we must have

$$N_k(0) = 0, \qquad N_k(\hat{l}) = 0, \tag{7.3.12}$$

as the boundary conditions for $N_k(y)$. To determine the $N_k(y)$, we must construct a *Green's function*, given here as $N_k(y)/M_k(\boldsymbol{\xi})$, for an ODE in a finite interval. Although $N_k(y)$ may also be found using eigenfunction expansions, we use a more direct and concrete approach in the following example.

Example 7.5. Green's Function for an Ordinary Differential Equation. Using the notation given in Example 7.1, we consider the Green's function $K(x; \xi)$ that satisfies the equation

$$\frac{\partial^2 K(x;\xi)}{\partial x^2} - c^2 K(x;\xi) = -\delta(x - \xi), \qquad 0 < x,\ \xi < l, \tag{7.3.13}$$

with the homogeneous (Dirichlet) boundary conditions

$$K(0;\xi) = 0, \qquad K(l;\xi) = 0. \tag{7.3.14}$$

From our discussion in Example 7.2 we find that $K(x;\xi)$ satisfies the homogeneous equation

$$\frac{\partial^2 K(x;\xi)}{\partial x^2} - c^2 K(x;\xi) = 0, \qquad x \neq \xi, \tag{7.3.15}$$

and the continuity and jump conditions at $x = \xi$:

$$K(x;\xi) \text{ continuous at } x = \xi, \quad \left[\frac{\partial K(x;\xi)}{\partial x}\right]_{x=\xi} = -1, \tag{7.3.16}$$

where the brackets represent the jump in the first derivative of $K(x;\xi)$ across $x = \xi$.

To solve this problem we denote $K(x;\xi)$ by $K_1(x;\xi)$ for $x < \xi$ and $K_2(x;\xi)$ for $x > \xi$. Both $K_1(x;\xi)$ and $K_2(x;\xi)$ satisfy the homogeneous equation (7.3.15). $K_1(x;\xi)$ vanishes at $x = 0$ while $K_2(x;\xi)$ vanishes at $x = l$. We easily conclude that

$$K_1(x;\xi) = a_1 \sinh(cx), \quad K_2(x;\xi) = a_2 \sinh[c(l - x)], \tag{7.3.17}$$

where a_1 and a_2 are as yet unspecified constants (note that c is assumed to be a constant). The continuity and jump conditions (7.3.16) imply that $a_1 \sinh(c\xi) = a_2 \sinh[c(l-\xi)]$, and $-a_2 c \cosh[c(l-\xi)] - a_1 c \cosh(c\xi) = -1$. The solution of the system for a_1 and a_2 is given as $a_1 = \sinh[c(l-\xi)]/(c\sinh(cl))$, $a_2 = \sinh[c\xi]/(c\sinh(cl))$. Inserting these expressions into (7.3.17), we obtain the Green's function $K(x;\xi)$ as

$$K(x;\xi) = \frac{1}{c\sinh(cl)} \begin{cases} \sinh[c(l-\xi)]\sinh(cx), & 0 < x < \xi, \\ \sinh(c\xi)\sinh[c(l-x)], & \xi < x < l. \end{cases} \tag{7.3.18}$$

We observe that $K(x;\xi) = K(\xi;x)$, so that the Green's function is symmetric.

In the limit as $c \to 0$, we obtain the Green's function $K(x;\xi)$ that satisfies the equation

$$\frac{\partial^2 K(x;\xi)}{\partial x^2} = -\delta(x-\xi), \qquad 0 < x,\ \xi < l, \tag{7.3.19}$$

and (7.3.14). Since $\sinh(x) \approx x$ as $x \to 0$, it is given

$$K(x;\xi) = \frac{1}{l} \begin{cases} x(l-\xi), & 0 < x < \xi, \\ \xi(l-x), & \xi < x < l. \end{cases} \tag{7.3.20}$$

We continue our discussion of (7.3.8) and assume that the eigenvalues λ_k of (7.3.3) are all positive. Then on using the results of the preceding example, we find that the solution of (7.3.11)–(7.3.12) is given as

$$N_k(y) = \frac{M_k(\boldsymbol{\xi})}{\sqrt{\lambda_k}\sinh(\sqrt{\lambda_k}\,\hat{l})} \begin{cases} \sinh[\sqrt{\lambda_k}(\hat{l}-\eta)]\sinh(\sqrt{\lambda_k}\,y), & 0 < y < \eta, \\ \sinh(\sqrt{\lambda_k}\,\eta)\sinh[\sqrt{\lambda_k}(\hat{l}-y)], & \eta < y < \hat{l}. \end{cases} \tag{7.3.21}$$

Inserting (7.3.21) into (7.3.9) completes the formal solution of the problem (7.3.8) for the Green's function $K(\mathbf{x}, y; \boldsymbol{\xi}, \eta)$ for the given special region.

The preceding technique can be generalized to deal with elliptic equations

$$\rho(\mathbf{x})\hat{L}u(\mathbf{x},y) - Lu(\mathbf{x},y) = -\rho(\mathbf{x})F(\mathbf{x},y), \qquad \mathbf{x} \in G,\ 0 < y < \hat{l}, \tag{7.3.22}$$

where

$$\hat{L}u(\mathbf{x},y) = \frac{\partial}{\partial y}\left[a(y)\frac{\partial u(\mathbf{x},y)}{\partial y}\right] + b(y)u(\mathbf{x},y), \tag{7.3.23}$$

so that $\hat{L}$ is a self-adjoint operator. The $(\mathbf{x}, y)$-region may be defined as in the foregoing, and the boundary conditions in the y-variable may be of the more general form (7.1.47), appropriately modified.

The Green's function $K(\mathbf{x}, y; \boldsymbol{\xi}, \eta)$ satisfies the equation

$$\rho(\mathbf{x})\hat{L}K(\mathbf{x},y;\boldsymbol{\xi},\eta) - LK(\mathbf{x},y;\boldsymbol{\xi},\eta) = -\delta(\mathbf{x}-\boldsymbol{\xi})\delta(y-\eta), \tag{7.3.24}$$

where $\mathbf{x}, \boldsymbol{\xi} \in G$, $0 < y, \eta < \hat{l}$, and $K(\mathbf{x}, y; \boldsymbol{\xi}, \eta)$ satisfies a homogeneous version of the boundary conditions for $u(\mathbf{x}, y)$. Expanding $K(\mathbf{x}, y; \boldsymbol{\xi}, \eta)$ as in (7.3.9), we conclude that $N_k(y)$ satisfies

$$\hat{L}N_k(y) - \lambda_k N_k(y) = -M_k(\boldsymbol{\xi})\delta(y - \eta), \ 0 < y, \ \eta < \hat{l}, \quad k = 1, 2, \ldots \quad (7.3.25)$$

and $N_k(y)$ satisfies appropriate boundary conditions at $y = 0$ and $y = \hat{l}$. Thus we again obtain a Green's function problem for a self-adjoint ordinary differential equation. A simple case of the Green's function problem was considered in Example 7.5, and further cases are studied in the exercises. We have further occasion to consider the technique of reducing the Green's function problem to a lower dimensional one later in this chapter.

Example 7.6. Laplace's Equation: Green's Function for a Rectangle. We consider the rectangular region G, given as $0 < x < l$ and $0 < y < \hat{l}$, and construct the Green's function $K(x, y; \xi, \eta)$ that satisfies the equation

$$\frac{\partial^2 K(x, y; \xi, \eta)}{\partial x^2} + \frac{\partial^2 K(x, y; \xi, \eta)}{\partial y^2} = -\delta(x - \xi)\delta(y - \eta), \quad (x, y) \in G, \quad (7.3.26)$$

with $0 < \xi < l$ and $0 < \eta < \hat{l}$ and the Dirichlet boundary condition

$$K(x, y; \xi, \eta) = 0, \qquad (x, y) \in \partial G. \quad (7.3.27)$$

We begin by applying the first of the two methods presented above. Thus, we must solve the following Dirichlet *eigenvalue problem* in the region G:

$$-\nabla^2 M(x, y) = \lambda M(x, y), \ (x, y) \in G, \quad M(x, y) = 0, \ (x, y) \in \partial G. \quad (7.3.28)$$

The eigenvalues and eigenfunctions are determined using the *method of separation of variables*.

Let $M(x, y) = F(x)G(y)$ and insert this expression into (7.3.28). We have $M_{xx}(x, y) + M_{yy}(x, y) + \lambda M(x, y) = F''(x)G(y) + F(x)G''(y) + \lambda F(x)G(y) = 0$. Dividing by $F(x)G(y)$ and separating variables gives $F''(x)/F(x) + \lambda = -G''(y)/G(y) = k^2$, where k^2 is the separation constant. The equations for $F(x)$ and $G(y)$ are

$$F''(x) + (\lambda - k^2)F(x) = 0, \quad G''(y) + k^2 G(y) = 0. \quad (7.3.29)$$

The boundary condition in (7.3.28) implies that

$$F(0) = F(l) = 0, \quad G(0) = G(\hat{l}) = 0. \quad (7.3.30)$$

Consequently, we are led to consider one-dimensional eigenvalue problems for $F(x)$ and $G(y)$ of a type studied in Section 4.3.

For the eigenvalue problem (7.3.29)–(7.3.30) for $G(x)$ we obtain as the eigenvalues and the eigenfunctions

$$k_m^2 = \left(\frac{\pi m}{\hat{l}}\right)^2, \quad G_m(y) = \sin\left(\frac{\pi m y}{\hat{l}}\right), \quad m = 1, 2, \ldots. \tag{7.3.31}$$

For each of these eigenvalues, we have an eigenvalue problem (7.3.29)–(7.3.30) for $F(x)$. As in Section 4.3, we find the eigenvalues

$$\lambda_{nm} - k_m^2 = \left(\frac{\pi n}{l}\right)^2, \qquad m = 1, 2, \ldots, \; n = 1, 2, \ldots, \tag{7.3.32}$$

and the eigenfunctions

$$F_n(x) = \sin\left(\frac{\pi n x}{l}\right), \qquad n = 1, 2, \ldots. \tag{7.3.33}$$

Combining the results obtained, we conclude that the *eigenvalues* and the *eigenfunctions* for the problem (7.3.28) are given as

$$\lambda_{nm} = \left(\frac{\pi n}{l}\right)^2 + \left(\frac{\pi m}{\hat{l}}\right)^2, \quad \hat{M}_{nm}(x, y) = \sin\left(\frac{\pi n x}{l}\right) \sin\left(\frac{\pi m y}{\hat{l}}\right), \tag{7.3.34}$$

with $n, m = 1, 2, \ldots$. The unnormalized eigenfunctions are denoted as $\hat{M}_{nm}(x, y) = F_n(x) G_m(y)$. It may be noted that the eigenvalues λ_{nm} are positive, infinite in number, and tend to infinity as n and m tend to infinity.

The appropriate inner product for the rectangular region G is given as

$$(f(x, y), g(x, y)) = \int_0^{\hat{l}} \int_0^{l} f(x, y) g(x, y) \, dx \, dy. \tag{7.3.35}$$

If we consider the two pairs (n, m) and (j, k) with $(n, m) \neq (j, k)$, we have

$$(\hat{M}_{nm}, \hat{M}_{jk}) = \int_0^{\hat{l}} \sin\left(\frac{\pi m y}{\hat{l}}\right) \sin\left(\frac{\pi k y}{\hat{l}}\right) dy \int_0^{l} \sin\left(\frac{\pi n x}{l}\right) \sin\left(\frac{\pi j x}{l}\right) dx = 0, \tag{7.3.36}$$

since $\{\sin(\pi n x/l)\}$ and $\{\sin(\pi m y/\hat{l})\}$ are orthogonal sets and $m \neq k$ and/or $n \neq j$. Thus eigenfunctions $\{\hat{M}_{nm}(x, y)\}$ form an *orthogonal set* that we now orthonormalize. The square of the norm of $\hat{M}_{nm}(x, y)$ is given as

$$\|\hat{M}_{nm}\|^2 = \int_0^{\hat{l}} \sin^2\left(\frac{\pi m y}{\hat{l}}\right) dy \int_0^{l} \sin^2\left(\frac{\pi n x}{l}\right) dx = \frac{\hat{l}}{2} \frac{l}{2}, \tag{7.3.37}$$

on using Section 4.3. Thus an orthonormal set of eigenfunctions is

$$M_{nm}(x, y) = \frac{2}{\sqrt{l\hat{l}}} \sin\left(\frac{\pi n x}{l}\right) \sin\left(\frac{\pi m y}{\hat{l}}\right), \qquad n, m = 1, 2, \ldots. \tag{7.3.38}$$

Different sets of values of (n, m) do not necessarily yield distinct eigenvalues λ_{nm}. For example, if $l = \hat{l}$, we see that $\lambda_{12} = (\pi/l)^2 + (2\pi/l)^2 = \lambda_{21}$. Nevertheless, the set of eigenvalues λ_{nm} can be arranged in a sequence that corresponds to the positive integers $k = 1, 2, 3, \ldots$ with an equivalent arrangement for the eigenfunctions $M_{nm}(x, y)$. It is then possible to speak of the set of eigenvalues λ_k and the eigenfunctions $M_k(x, y)$, with $k = 1, 2, 3, \ldots$, and even if λ_k is a multiple eigenvalue, the corresponding eigenfunctions are orthogonal, as we have shown. The single-subscript notation was used in our earlier discussions of multidimensional eigenvalue problems.

Given a function $f(x, y)$, defined in the rectangular region $0 < x < l$ and $0 < y < \hat{l}$, that satisfies certain smoothness conditions, we have the expansion

$$f(x, y) = \frac{2}{\sqrt{l\hat{l}}} \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} c_{nm} \sin\left(\frac{\pi n x}{l}\right) \sin\left(\frac{\pi m y}{\hat{l}}\right), \tag{7.3.39}$$

with the *Fourier coefficients* c_{nm} given as

$$c_{nm} = \frac{2}{\sqrt{l\hat{l}}} \int_0^{\hat{l}} \int_0^{l} f(x, y) \sin\left(\frac{\pi n x}{l}\right) \sin\left(\frac{\pi m y}{\hat{l}}\right) dx\, dy. \tag{7.3.40}$$

The expansion (7.3.39) is known as a *double Fourier sine series.*

Having determined the eigenfunctions and eigenvalues for the problem (7.3.28) we are now in a position to construct the Green's function $K(x, y; \xi, \eta)$. In view of (7.3.6), we obtain

$$K(x, y; \xi, \eta) = \frac{4}{l\hat{l}} \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \frac{\sin(\pi n x/l) \sin(\pi n \xi/l) \sin\left(\pi m y/\hat{l}\right) \sin\left(\pi m \eta/\hat{l}\right)}{(\pi n/l)^2 + (\pi m/\hat{l})^2} \tag{7.3.41}$$

as the *eigenfunction expansion* of the Green's function for the region G.

The alternative method presented above for the construction of the Green's function for the rectangular region G has better convergence properties than those of the series (7.3.41), as we now demonstrate. The equation (7.3.26) for the Green's function $K(x, y; \xi, \eta)$ is written as

$$\frac{\partial^2 K(x, y; \xi, \eta)}{\partial y^2} - LK(x, y; \xi, \eta) = -\delta(x - \xi)\delta(y - \eta), \tag{7.3.42}$$

with the operator L given as $L = -(\partial^2/\partial x^2)$. The appropriate eigenvalue problem is

$$LM(x) = -M''(x) = \lambda M(x), \; 0 < x < l, \quad M(0) = M(l) = 0. \tag{7.3.43}$$

This represents a *Sturm-Liouville problem* which was solved in Example 4.4. The eigenvalues and orthonormalized eigenfunctions are

$$\lambda_k = \left(\frac{\pi k}{l}\right)^2, \; M_k(x) = \sqrt{\frac{2}{l}} \sin\left(\frac{\pi k x}{l}\right), \quad k = 1, 2, \ldots. \tag{7.3.44}$$

With $N_k(y)$ defined as in (7.3.21), the Green's function is given by the *eigenfunction expansion*

$$K(x, y; \xi, \eta) = \sum_{k=1}^{\infty} N_k(y) M_k(x). \tag{7.3.45}$$

For y different from η, the hyperbolic functions that occur in $N_k(y)$ can be approximated by exponentials as was done in Example 4.11, and the series can be shown to converge fairly rapidly if $|y - \eta|$ is not small.

Both (7.3.41) and (7.3.45) can be used to solving Dirichlet's problem for Laplace's equation in a rectangle. It has been shown in Example 4.11 that a direct separation of variables approach yields quite satisfactory results.

Modified Green's Functions for Elliptic PDEs

The eigenfunction expansion method for constructing Green's functions for the elliptic equation (7.1.1) with the boundary condition (7.1.2) fails if $\lambda_0 = 0$ is an eigenvalue of the associated eigenvalue problem. As indicated, this occurs for the Green's function problem (7.3.1)–(7.3.2) if and only if $q(\mathbf{x}) = 0$ in (7.3.1) and $\alpha(\mathbf{x}) = 0$ in the boundary condition (7.3.2).

Let $\lambda_0 = 0$ be an eigenvalue of the operator L [see (7.3.3)] and $M_0(\mathbf{x})$ be the eigenfunction corresponding to λ_0. Then (7.3.5) implies that $\lambda_0 N_0 = 0 = M_0(\boldsymbol{\xi})$, and this is not possible since we cannot have $M_0 \equiv 0$. Clearly then, the Green's function cannot be constructed in the given manner if $\lambda_0 = 0$ is an eigenvalue. There are two methods whereby a *modified Green's function* can be constructed in the special case of a zero eigenvalue. For each of the two modified Green's functions it is possible to solve the corresponding boundary value problem for $u(\mathbf{x})$ in a manner similar to that presented for the (ordinary) Green's function $K(\mathbf{x}; \boldsymbol{\xi})$ in Section 7.1, and this will be demonstrated below.

We put $q(\mathbf{x}) = 0$ in (7.1.1) and consider the *elliptic equation*

$$Lu(\mathbf{x}) = -\nabla \cdot (p(\mathbf{x}) \nabla u(\mathbf{x})) = \rho(\mathbf{x}) F(\mathbf{x}), \qquad \mathbf{x} \in G, \tag{7.3.46}$$

with the *Neumann boundary condition*

$$\beta(\mathbf{x}) \left. \frac{\partial u(\mathbf{x})}{\partial n} \right|_{\partial G} = B(\mathbf{x}). \tag{7.3.47}$$

The eigenvalue problem that corresponds to (7.3.46)–(7.3.47) is given as

$$LM(\mathbf{x}) = \lambda \rho(\mathbf{x}) M(\mathbf{x}), \quad \mathbf{x} \in G, \qquad \left. \frac{\partial M(\mathbf{x})}{\partial n} \right|_{\partial G} = 0, \tag{7.3.48}$$

and $\lambda_0 = 0$ is an eigenvalue. It follows from the results of Section 4.6 that

$$\lambda_0 N_0 = 0 = \iint_G \rho(\mathbf{x}) F(\mathbf{x}) M_0(\mathbf{x}) \, d\mathbf{x} + \int_{\partial G} \frac{p(\mathbf{x})}{\beta(\mathbf{x})} M_0(\mathbf{x}) B(\mathbf{x}) \, ds. \tag{7.3.49}$$

Since $M_0(\mathbf{x})$ must be a constant (see Exercise 4.2.5), we conclude that

$$\iint_G \rho(\mathbf{x})F(\mathbf{x})\,d\mathbf{x} + \int_{\partial G} \frac{p(\mathbf{x})}{\beta(\mathbf{x})} B(\mathbf{x})\,ds = 0. \tag{7.3.50}$$

Unless $F(\mathbf{x})$ and $B(\mathbf{x})$ are such that the *compatibility condition* (7.3.50) is satisfied, the boundary value problem (7.3.46)–(7.3.47) has no solution. We assume that (7.3.50) is satisfied and construct the modified Green's function.

In the first method, the *modified Green's function*, which we denote by $\hat{K}(\mathbf{x};\boldsymbol{\xi})$, is expanded in a series of eigenfunctions of the operator L,

$$\hat{K}(\mathbf{x};\boldsymbol{\xi}) = \sum_{k=1}^{\infty} N_k(\boldsymbol{\xi})M_k(\mathbf{x}), \tag{7.3.51}$$

where the $M_k(\mathbf{x})$, $k = 1, 2, \ldots$ correspond to the positive eigenvalues of L. This expansion differs from that given in (7.3.4) since the eigenfunction $M_0(\mathbf{x})$ is absent from the series, even though $\lambda_0 = 0$ is an eigenvalue for this problem. We have removed the term $N_0(\boldsymbol{\xi})M_0(\mathbf{x})$ from the expansion so as to avoid obtaining the contradictory result $\lambda_0 N_0(\boldsymbol{\xi}) = 0 = M_0(\boldsymbol{\xi})$, which was derived in the foregoing.

Since the complete set of eigenfunctions is given as $\{M_k(\mathbf{x})\}$, $k = 0, 1, 2, \ldots$, the eigenfunction expansion of $\delta(\mathbf{x} - \boldsymbol{\xi})$ is

$$\delta(\mathbf{x} - \boldsymbol{\xi}) = \sum_{k=0}^{\infty} (\delta(\mathbf{x} - \boldsymbol{\xi}), M_k(\mathbf{x}))M_k(\mathbf{x}) = \rho(\boldsymbol{\xi}) \sum_{k=0}^{\infty} M_k(\boldsymbol{\xi})M_k(\mathbf{x}). \tag{7.3.52}$$

This series can be expressed as $\delta(\mathbf{x} - \boldsymbol{\xi}) = \rho(\mathbf{x}) \sum_{k=0}^{\infty} M_k(\boldsymbol{\xi})M_k(\mathbf{x})$. Further,

$$L\hat{K}(\mathbf{x};\boldsymbol{\xi}) = \sum_{k=1}^{\infty} N_k(\boldsymbol{\xi})LM_k(\mathbf{x}) = \rho(\mathbf{x}) \sum_{k=1}^{\infty} \lambda_k N_k(\boldsymbol{\xi})M_k(\mathbf{x}). \tag{7.3.53}$$

Comparing (7.3.52) with (7.3.53) shows that if we set $N_k(\boldsymbol{\xi}) = M_k(\boldsymbol{\xi})/\lambda_k$, $k = 1, 2, \ldots$, the modified Green's function satisfies the equation

$$L\hat{K}(\mathbf{x};\boldsymbol{\xi}) = \sum_{k=1}^{\infty} \rho(\mathbf{x})M_k(\boldsymbol{\xi})M_k(\mathbf{x}) = \delta(\mathbf{x} - \boldsymbol{\xi}) - \rho(\mathbf{x})M_0(\boldsymbol{\xi})M_0(\mathbf{x}). \tag{7.3.54}$$

Noting the foregoing discussion, we define the *modified Green's function* $\hat{K}(\mathbf{x};\boldsymbol{\xi})$ to be a solution of the equation

$$L\hat{K}(\mathbf{x};\boldsymbol{\xi}) \equiv -\nabla \cdot (p(\mathbf{x})\nabla \hat{K}(\mathbf{x};\boldsymbol{\xi})) = \delta(\mathbf{x} - \boldsymbol{\xi}) - \rho(\mathbf{x})M_0(\boldsymbol{\xi})M_0(\mathbf{x}),\ \mathbf{x}, \boldsymbol{\xi} \in G \tag{7.3.55}$$

with the Neumann boundary condition

$$\left.\frac{\partial \hat{K}(\mathbf{x};\boldsymbol{\xi})}{\partial n}\right|_{\partial G} = 0. \tag{7.3.56}$$

The bilinear series of eigenfunctions for $\hat{K}(\mathbf{x};\boldsymbol{\xi})$ has the form

$$\hat{K}(\mathbf{x};\boldsymbol{\xi}) = \sum_{k=1}^{\infty} \frac{M_k(\mathbf{x})M_k(\boldsymbol{\xi})}{\lambda_k}. \tag{7.3.57}$$

Given the boundary value problem (7.3.46)–(7.3.47), we now construct a *solution formula* for $u(\mathbf{x})$ in terms of the modified Green's function $\hat{K}(\mathbf{x};\boldsymbol{\xi})$. Following the procedure given in Section 7.1, we set $w = \hat{K}$ in (7.3) and note that (7.1.4) gives

$$\iint_G u(\mathbf{x})L\hat{K}(\mathbf{x};\boldsymbol{\xi})\,dv = u(\boldsymbol{\xi}) - M_0(\boldsymbol{\xi}) \iint_G \rho(\mathbf{x})u(\mathbf{x})M_0(\mathbf{x})\,dv. \tag{7.3.58}$$

The last integral in (7.3.58) is the Fourier coefficient of $u(\mathbf{x})$ with respect to the eigenfunction $M_0(\mathbf{x})$, and it can be expressed as $(u(\mathbf{x}), M_0(\mathbf{x}))$. Then we obtain from (7.1.8) the solution formula

$$u(\boldsymbol{\xi}) = \iint_G \rho F\hat{K}(\mathbf{x};\boldsymbol{\xi})\,dv + \int_{\partial G} \frac{pB}{\beta}\hat{K}(\mathbf{x};\boldsymbol{\xi})\,ds + (u(\mathbf{x}), M_0(\mathbf{x}))M_0(\boldsymbol{\xi}), \tag{7.3.59}$$

where ρ, p, F, B, β are all functions of $\mathbf{x}$. The solution exists only if the compatibility condition (7.3.50) is satisfied, and even then it is determined only up to an arbitrary constant multiple of the constant eigenfunction $M_0(\mathbf{x})$ as seen from the last term in (7.3.59). Consequently, as has been noted in Section 6.8, the solution to the boundary value problem (7.3.46)–(7.3.47) is not unique. Any constant can be added to it.

Similarly, the modified Green's function $\hat{K}(\mathbf{x};\boldsymbol{\xi})$ is determined only up to an arbitrary constant. By expressing $\hat{K}(\mathbf{x};\boldsymbol{\xi})$ in the form (7.3.51) and (7.3.57), we have, in effect, equated the arbitrary constant to zero. As a result, the modified Green's function $\hat{K}(\mathbf{x};\boldsymbol{\xi})$ is symmetric. The modified Green's function $\hat{K}(\mathbf{x};\boldsymbol{\xi})$ differs from the ordinary Green's function $K(\mathbf{x};\boldsymbol{\xi})$ in that $L\hat{K}$ does not equal $\delta(\mathbf{x}-\boldsymbol{\xi})$, but is given as in (7.3.55). However, $\hat{K}(\mathbf{x};\boldsymbol{\xi})$ does satisfy the homogeneous boundary condition (7.3.56).

There is an alternative construction of a *modified Green's function*, which we denote by $\tilde{K}(\mathbf{x};\boldsymbol{\xi})$, associated with the boundary value problem (7.3.46)–(7.3.47), for which we have

$$L\tilde{K}(\mathbf{x};\boldsymbol{\xi}) \equiv -\nabla\cdot(p(\mathbf{x})\nabla\tilde{K}(\mathbf{x};\boldsymbol{\xi})) = \delta(\mathbf{x}-\boldsymbol{\xi}), \qquad \mathbf{x},\ \boldsymbol{\xi} \in G. \tag{7.3.60}$$

An application of the divergence theorem shows that

$$-\iint_G \nabla\cdot(p\,\nabla\tilde{K})\,dv = -\int_{\partial G} p\frac{\partial\tilde{K}}{\partial n}\,ds = \iint_G \delta(\mathbf{x}-\boldsymbol{\xi})\,dv = 1. \tag{7.3.61}$$

Consequently, if we set $L\tilde{K}(\mathbf{x};\boldsymbol{\xi}) = \delta(\mathbf{x}-\boldsymbol{\xi})$, we cannot have $\partial\tilde{K}(\mathbf{x};\boldsymbol{\xi})/\partial n = 0$ on ∂G, as was the case for $\hat{K}(\mathbf{x};\boldsymbol{\xi})$. Instead, we must set $\partial\tilde{K}(\mathbf{x};\boldsymbol{\xi})/\partial n\big|_{\partial G} = b(\mathbf{x})$, where $b(\mathbf{x})$ is any function for which $\int_{\partial G} p(\mathbf{x})b(\mathbf{x})\,ds = -1$, in view of (7.3.61). [Note that $b(\mathbf{x})$ may be taken to be a constant.]

The *solution formula* for $u(\mathbf{x})$ that satisfies (7.3.46)–(7.3.47) is then obtained from the results given at the beginning of Section 7.1. With $w(\mathbf{x}) = \tilde{K}(\mathbf{x};\boldsymbol{\xi})$ we obtain in place of (7.1.7),

$$u(\mathbf{x})\frac{\partial \tilde{K}(\mathbf{x};\boldsymbol{\xi})}{\partial n} - \tilde{K}(\mathbf{x};\boldsymbol{\xi})\frac{\partial u(\mathbf{x})}{\partial n} = u(\mathbf{x})b(\mathbf{x}) - \frac{1}{\beta(\mathbf{x})}B(\mathbf{x})\tilde{K}(\mathbf{x};\boldsymbol{\xi}) \qquad (7.3.62)$$

and the solution is obtained in the form

$$u(\boldsymbol{\xi}) = \iint_G \rho F\tilde{K}(\mathbf{x};\boldsymbol{\xi})\,dv + \int_{\partial G}\frac{pB}{\beta}\tilde{K}(\mathbf{x};\boldsymbol{\xi})\,ds - \int_{\partial G} pbu(\mathbf{x})\,ds, \qquad (7.3.63)$$

where ρ, p, F, B, β, b are all functions of $\mathbf{x}$, on using (7.1.8). The last integral in (7.3.63) is an arbitrary constant since $b(\mathbf{x})$ is arbitrary.

The expansion of $\tilde{K}(\mathbf{x};\boldsymbol{\xi})$ in a series of eigenfunctions is not as straightforward as that for $\hat{K}(\mathbf{x};\boldsymbol{\xi})$ since $\tilde{K}(\mathbf{x};\boldsymbol{\xi})$ does not satisfy a homogeneous boundary condition. Nevertheless, it can be carried out using the finite Fourier transform techniques of Section 4.6. A relation between $\hat{K}(\mathbf{x};\boldsymbol{\xi})$ and $\tilde{K}(\mathbf{x};\boldsymbol{\xi})$ can be established by way of the procedure developed in Chapter 4, whereby inhomogeneous boundary conditions can be transformed to homogeneous conditions. This is considered in the exercises.

Example 7.7. Laplace's Equation: The Modified Green's Function in a Rectangle. Given the rectangular region G defined as $0 < x < l$ and $0 < y < \hat{l}$, we construct the modified Green's function $\hat{K}(x,y;\xi,\eta)$ that satisfies the equation [(see (7.3.55)]

$$\frac{\partial^2 \hat{K}(x,y;\xi,\eta)}{\partial x^2} + \frac{\partial^2 \hat{K}(x,y;\xi,\eta)}{\partial y^2} = -\delta(x-\xi)\delta(y-\eta) + M_0(\boldsymbol{\xi})M_0(\mathbf{x}), \quad (7.3.64)$$

with $(x,y) \in G$ and $(\xi,\eta) \in G$, and the *Neumann condition* $\partial \hat{K}(x,y;\xi,\eta)/\partial n\big|_{\partial G} = 0$, where $\partial \hat{K}(x,y;\xi,\eta)/\partial n$ is an exterior normal derivative.

To solve for $\hat{K}(x,y;\xi,\eta)$, we must determine the eigenvalues and eigenfunctions of the problem

$$-\nabla^2 M(x,y) = \lambda M(x,y),\ (x,y) \in G, \quad \left.\frac{\partial M(x,y)}{\partial n}\right|_{\partial G} = 0. \qquad (7.3.65)$$

Using separation of variables, we set $M(x,y) = F(x)G(y)$ in (7.3.65) and obtain $M_{xx}(x,y) + M_{yy}(x,y) + \lambda M = F''(x)G(y) + F(x)G''(y) + \lambda F(x)G(y) = 0$. Dividing by $F(x)G(y)$ and separating variables gives $F''(x)/F(x) + \lambda = -G''(y)/G(y) = k^2$, where k^2 is the separation constant. Then $F(x)$ satisfies the boundary value problem

$$F''(x) + (\lambda - k^2)F(x) = 0,\ 0 < x < l, \quad F'(0) = F'(l) = 0. \qquad (7.3.66)$$

Also, $G(y)$ satisfies

$$G''(y) + k^2 G(y) = 0,\ 0 < y < \hat{l}, \quad G'(0) = G'(\hat{l}) = 0. \qquad (7.3.67)$$

The eigenvalue problems for $F(x)$ and $G(y)$ were studied in Section 4.3.

The eigenvalues and eigenfunctions for (7.3.67) are

$$k_m^2 = \left(\frac{\pi m}{\hat{l}}\right)^2, \quad G_m(y) = \cos\left(\frac{\pi m y}{\hat{l}}\right), \quad m = 0, 1, 2, \ldots. \tag{7.3.68}$$

For $F(x)$ we find the eigenvalues

$$\lambda_{nm} - k_m^2 = \left(\frac{\pi n}{l}\right)^2, \qquad n = 0, 1, 2, \ldots, \; m = 0, 1, 2, \ldots, \tag{7.3.69}$$

and the eigenfunctions

$$F_n(x) = \cos\left(\frac{\pi n x}{l}\right), \qquad n = 0, 1, 2, \ldots. \tag{7.3.70}$$

Then the *eigenvalues* and the corresponding *eigenfunctions* for (7.3.65) are

$$\lambda_{nm} = \left(\frac{\pi n}{l}\right)^2 + \left(\frac{\pi m}{\hat{l}}\right)^2, \; \hat{M}_{nm}(x, y) = \cos\left(\frac{\pi n x}{l}\right)\cos\left(\frac{\pi m y}{\hat{l}}\right), \tag{7.3.71}$$

with $n, m = 0, 1, 2, \ldots$. The unnormalized eigenfunctions are denoted as $\hat{M}_{nm}(x, y)$. It may be noted that the eigenvalues λ_{nm} are nonnegative, infinite in number, and tend to infinity as n and m tend to infinity.

The inner product for the region G is given as in (7.3.35), and it is easily shown that $\{\hat{M}_{nm}(x, y)\}$ is an orthogonal set using Section 4.3. Also, the square of the norm of $\hat{M}_{nm}(x, y)$ is $||\hat{M}_{nm}(x, y)||^2 = (\hat{M}_{nm}(x, y), \hat{M}_{nm}(x, y)) = l\hat{l}/4$, $n, m = 1, 2, \ldots$. For $\hat{M}_{00}(x, y)$, $||\hat{M}_{00}(x, y)||^2 = l\hat{l}$, and for $\hat{M}_{n0}(x, y)$ and $\hat{M}_{0m}(x, y)$, $||\hat{M}_{n0}(x, y)||^2 = l\hat{l}/2 = ||\hat{M}_{0m}(x, y)||^2$, with $n, m = 1, 2, \ldots$.

Consequently, the orthonormal set of eigenfunctions is given as

$$M_{nm}(x, y) = \begin{cases} \sqrt{1/l\hat{l}}, & m = n = 0, \\ \sqrt{2/l\hat{l}}\cos(\pi n x/l), & m = 0, \; n = 1, 2, \ldots, \\ \sqrt{2/l\hat{l}}\cos\left(\pi m y/\hat{l}\right), & n = 0, \; m = 1, 2, \ldots, \\ \sqrt{4/l\hat{l}}\cos(\pi n x/l)\cos\left(\pi m y/\hat{l}\right), & n, m = 1, 2, \ldots. \end{cases} \tag{7.3.72}$$

We see that $\lambda_{00} = 0$ is an eigenvalue whose eigenfunction $M_{00}(x, y)$ is constant.

Having determined the eigenfunctions $M_{nm}(x, y)$, we expand the modified Green's function in a series of these functions as in (7.3.57). The eigenfunction $M_{00}(x, y)$ must be excluded from the series. We have

$$\hat{K}(x,y;\xi,\eta) = \frac{2}{l\hat{l}}\sum_{m=1}^{\infty}\frac{\cos\left(\pi m y/\hat{l}\right)\cos\left(\pi m\eta/\hat{l}\right)}{(\pi m/\hat{l})^2} + \frac{2}{l\hat{l}}\sum_{n=1}^{\infty}\frac{\cos(\pi n x/l)\cos(\pi n\xi/l)}{(\pi n/l)^2}$$

$$+\frac{4}{l\hat{l}}\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{\cos(\pi n x/l)\cos(\pi n\xi/l)\cos\left(\pi m y/\hat{l}\right)\cos\left(\pi m\eta/\hat{l}\right)}{(\pi n/l)^2+(\pi m/\hat{l})^2}. \qquad (7.3.73)$$

The alternative procedure presented above for the construction of Green's functions in terms of eigenfunctions for lower-dimensional problems can also be used to determine modified Green's functions. Its application to the problem of Example 7.7 is considered in the exercises. We do not construct Green's functions for nonselfadjoint problems because of difficulties that arise in connection with eigenfunction expansions for nonselfadjoint eigenvalue problems.

Green's Functions for Hyperbolic PDEs

The Green's function for the *hyperbolic problem* considered in Section 7.1 is expressed as $K(\mathbf{x},t;\boldsymbol{\xi},\tau)$ and satisfies the equation

$$\rho(\mathbf{x})\frac{\partial^2 K(\mathbf{x},t;\boldsymbol{\xi},\tau)}{\partial t^2} + LK(\mathbf{x},t;\boldsymbol{\xi},\tau) = \delta(\mathbf{x}-\boldsymbol{\xi})\delta(t-\tau),\ \mathbf{x},\boldsymbol{\xi}\in G,\ t,\tau<T, \qquad (7.3.74)$$

with G and T defined as before. In addition, $K(\mathbf{x},t;\boldsymbol{\xi},\tau)$ satisfies the initial conditions $K(\mathbf{x},T;\boldsymbol{\xi},\tau) = \partial K(\mathbf{x},T;\boldsymbol{\xi},\tau)/\partial t = 0$ and the boundary condition

$$\alpha(\mathbf{x})K(\mathbf{x},t;\boldsymbol{\xi},\tau) + \beta(\mathbf{x})\frac{\partial K(\mathbf{x},t;\boldsymbol{\xi},\tau)}{\partial n}\bigg|_{\partial G} = 0, \qquad t<T. \qquad (7.3.75)$$

If the region G is one-dimensional (i.e., $\mathbf{x} = x$ and $0 < x < l$), the boundary condition (7.3.75) is replaced by a one-dimensional form as in (7.1.44).

Proceeding as in Section 4.6, we expand $K(\mathbf{x},t;\boldsymbol{\xi},\tau)$ in a series of eigenfunctions, $K(\mathbf{x},t;\boldsymbol{\xi},\tau) = \sum_{k=1}^{\infty} N_k(t)M_k(\mathbf{x})$, where $M_k(\mathbf{x})$ are the eigenfunctions of the operator L, that satisfy a boundary condition of the form (7.3.75). To determine the $N_k(t)$, (7.3.74) is multiplied by $M_k(\mathbf{x})$ and integrated over G. This yields the equations

$$N_k''(t) + \lambda_k N_k(t) = M_k(\boldsymbol{\xi})\delta(t-\tau), \qquad t,\tau<T,\ k=1,2,\ldots. \qquad (7.3.76)$$

The initial conditions for $N_k(t)$ are $N_k(T) = N_k'(T) = 0$.

On the basis of Example 7.2 we conclude that $N_k(t)$ is continuous at $t=\tau$ and that $N_k'(t)$ has a jump discontinuity $[N_k'(t)]_{t=\tau} = M_k(\boldsymbol{\xi})$. The conditions at $t=T$ imply that $N_k(t) = 0$, $\tau < t \le T$. Additionally, the continuity of $N_k(t)$ and the jump condition on $N_k'(t)$ at $t=\tau$ imply that $N_k(\tau) = 0$, $N_k'(\tau) = -M_k(\boldsymbol{\xi})$. These

serve as initial conditions for $N_k(t)$ in the interval $t < \tau$. Since $\delta(t-\tau) = 0$ for $t < \tau$, $N_k(t)$ satisfies the equation

$$N_k''(t) + \lambda_k N_k(t) = 0, \qquad t < \tau. \tag{7.3.77}$$

The solution of the problem for $N_k(t)$ is easily found and can be given as

$$N_k(t) = \frac{1}{\sqrt{\lambda_k}} \sin\left[\sqrt{\lambda_k}(\tau - t)\right] M_k(\boldsymbol{\xi}) H(\tau - t), \tag{7.3.78}$$

where $H(x)$ is the Heaviside function (7.2.18). The *Green's function* $K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)$ thus has the form

$$K(\mathbf{x}, t; \boldsymbol{\xi}, \tau) = \left[\sum_{k=1}^{\infty} \frac{1}{\sqrt{\lambda_k}} \sin\left[\sqrt{\lambda_k}(\tau - t)\right] M_k(\boldsymbol{\xi}) M_k(\mathbf{x})\right] H(\tau - t). \tag{7.3.79}$$

We observe that $K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)$ is symmetric in $\mathbf{x}$ and $\boldsymbol{\xi}$ but not in t and τ. As noted in the discussion following (7.1.25), the function $S(\mathbf{x}, t; \boldsymbol{\xi}, \tau) = K(\mathbf{x}, -t; \boldsymbol{\xi}, -\tau)$ is the *causal fundamental solution* for the given problem and is given as

$$S(\mathbf{x}, t; \boldsymbol{\xi}, \tau) = \left[\sum_{k=1}^{\infty} \frac{1}{\sqrt{\lambda_k}} \sin\left[\sqrt{\lambda_k}(t - \tau)\right] M_k(\boldsymbol{\xi}) M_k(\mathbf{x})\right] H(t - \tau). \tag{7.3.80}$$

It may be verified directly that (7.3.79) is a generalized solution of (7.3.74). It has been assumed that $\lambda_0 = 0$ is not an eigenvalue. If $\lambda_0 = 0$ is an eigenvalue, an additional term $(\tau - t) M_0(\boldsymbol{\xi}) M_0(\mathbf{x})$ must be added to the series in (7.3.79), as shown in the exercises. Some specific examples of eigenfunction expansions of Green's functions for hyperbolic problems are also considered in the exercises.

We note that the Green's function (7.3.79) is a solution of the homogeneous version of (7.3.74) for $t < \tau$. At $t = \tau$ the Green's function (7.3.79) vanishes, and its time derivative is $-\delta(\mathbf{x} - \boldsymbol{\xi})/\rho(\boldsymbol{\xi})$. Also, the boundary condition (7.3.75) is satisfied by the Green's function for $t < \tau$. These results follow easily from our discussion. Thus, for $t < \tau$ the Green's functions determined by each of the methods given in Section 7.1 are identical.

Green's Functions for Parabolic PDEs

The Green's function $K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)$ for the *parabolic problem* of Section 7.1 satisfies the equation

$$-\rho(\mathbf{x}) \frac{\partial K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)}{\partial t} + LK(\mathbf{x}, t; \boldsymbol{\xi}, \tau) = \delta(\mathbf{x} - \boldsymbol{\xi}) \delta(t - \tau), \ \mathbf{x}, \boldsymbol{\xi} \in G, \ t, \tau < T, \tag{7.3.81}$$

and the initial and boundary conditions $K(\mathbf{x}, T; \boldsymbol{\xi}, \tau) = 0$, $\alpha(\mathbf{x}) K(\mathbf{x}, t; \boldsymbol{\xi}, \tau) + \beta(\mathbf{x}) \partial K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)/\partial n\big|_{\partial G} = 0$, $t < T$, respectively, if G is a two- or three-dimensional region.

As was done for the hyperbolic problem, we expand $K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)$ in a series of eigenfunctions $K(\mathbf{x}, t; \boldsymbol{\xi}, \tau) = \sum_{k=1}^{\infty} N_k(t) M_k(\mathbf{x})$. Multiplying (7.3.81) by $M_k(\mathbf{x})$ and integrating over the region G, we obtain

$$-N_k'(t) + \lambda_k N_k(t) = M_k(\boldsymbol{\xi})\delta(t-\tau), \qquad t, \tau < T, \ k = 1, 2, \ldots. \tag{7.3.82}$$

Since $\delta(t-\tau) = \delta(\tau - t)$, we can write (7.3.82) as

$$\frac{d}{dt}[e^{-\lambda_k t} N_k(t)] = -M_k(\boldsymbol{\xi}) e^{-\lambda_k t}\delta(\tau - t) = -M_k(\boldsymbol{\xi}) e^{-\lambda_k \tau}\delta(\tau - t), \tag{7.3.83}$$

where (7.2.14) was used. Integrating (7.3.83) and using the initial condition $N_k(T) = 0$, we obtain

$$N_k(t) = e^{\lambda_k (t-\tau)} M_k(\boldsymbol{\xi}) H(\tau - t). \tag{7.3.84}$$

[We recall that $dH(x)/dx = \delta(x)$.] Thus the *Green's function* is

$$K(\mathbf{x}, t; \boldsymbol{\xi}, \tau) = \left[\sum_{k=1}^{\infty} e^{\lambda_k (t-\tau)} M_k(\boldsymbol{\xi}) M_k(\mathbf{x})\right] H(\tau - t). \tag{7.3.85}$$

As noted in the discussion following (7.1.33), $S(\mathbf{x}, t; \boldsymbol{\xi}, \tau) = K(\mathbf{x}, -t; \boldsymbol{\xi}, -\tau)$ is the *causal fundamental solution* for the given problem, and has the form

$$S(\mathbf{x}, t; \boldsymbol{\xi}, \tau) = \left[\sum_{k=1}^{\infty} e^{\lambda_k (\tau - t)} M_k(\boldsymbol{\xi}) M_k(\mathbf{x})\right] H(t - \tau). \tag{7.3.86}$$

It is not difficult to show directly (see the exercises) that the series (7.3.85) is a generalized solution of (7.3.81). If $\lambda_0 = 0$ is an eigenvalue, the series (7.3.85) must be modified as shown in the exercises. Some specific examples of eigenfunction expansions of Green's functions for parabolic problems are also considered in the exercises.

We conclude with the observation that the Green's function (7.3.85) is a solution of the homogeneous version of (7.3.81) for $t < \tau$. It takes the value $\delta(\mathbf{x} - \boldsymbol{\xi})/\rho(\boldsymbol{\xi})$ at $t = \tau$. Also, the boundary condition for the problem is satisfied by the Green's function for $t < \tau$. This follows easily from our discussion. Consequently, we find that for $t < \tau$ the Green's functions determined by each of the methods given in Section 7.1 for the parabolic case are identical.

Exercises 7.3

7.3.1. Obtain the Green's function (7.3.20) for the problem (7.3.19) and (7.3.14) directly—that is, not as a limit of the function (7.3.18). Use this Green's function to solve the problem $u''(x) = f(x)$, $0 < x < l$, $u(0) = u(l) = 0$.

7.3.2. Solve the boundary value problem: $u''(x) - c^2 u(x) = -f(x)$, $0 < x < l$, $u(0) = u(l) = 0$ using the Green's function (7.3.18).

7.3.3. Determine the Green's function $K(x;\xi)$ for the problem: $\partial^2 K(x,\xi)/\partial x^2 - c^2 K(x;\xi) = -\delta(x-\xi)$, $0 < x, \xi < l$, $K(0;\xi) = 0$, $\partial K(l;\xi)/\partial x = 0$.

7.3.4. Consider the problem of Exercise 7.3.3 with the boundary condition at $x = 0$ replaced by $\partial K(0;\xi)/\partial x = 0$ and determine the Green's function $K(x;\xi)$. Show that $K(x;\xi)$ has no limit as $c \to 0$ and explain why this limit fails to exist.

7.3.5. Obtain the Green's function $K(x;\xi)$ that satisfies the following problem: $\partial^2 K(x,\xi)/\partial x^2 + c^2 K(x;\xi) = -\delta(x-\xi)$, $0 < x, \xi < l$, $K(0;\xi) = 0$, $K(l;\xi) = 0$.

7.3.6. Find the Green's function $K(x;\xi)$ determined from the following problem: $x\partial^2 K(x,\xi)/\partial x^2 + \partial K(x;\xi)/\partial x = -\delta(x-\xi)$, $0 < x, \xi < 1$, with $K(x;\xi)$ bounded at $x = 0$ and $K(1;\xi) = 0$.

7.3.7. Solve for the Green's function $K(x;\xi)$: $\partial/\partial x\left[(1+x)^2 \partial K(x;\xi)/\partial x\right] - 2K(x;\xi) = -\delta(x-\xi)$, $0 < x, \xi < 1$, $K(0;\xi) = K(1;\xi) = 0$.

7.3.8. Obtain a one-dimensional version of the formula (7.1.8) for the solution of the following boundary value problem, $-(p(x)u'(x))' + q(x)u(x) = \rho(x)F(x)$, $0 < x < l$, $\alpha_1 u(0) - \beta_1 u'(0) = a_1$, $\alpha_2 u(l) + \beta_2 u'(l) = a_2$, where $\alpha_1, \alpha_2, \beta_1, \beta_2$ are nonnegative, $\alpha_1 + \beta_1 > 0$, $\alpha_2 + \beta_2 > 0$, $p(x) > 0$, $q(x) \geq 0$, $\rho(x) > 0$, and $F(x)$ is a given function. *Hint*: The solution formula is given in terms of the Green's function $K(x;\xi)$ determined from (7.1.43)–(7.1.44).

7.3.9. Use the results of Exercises 7.3.3 and 7.3.8 to solve the following problem: $u''(x) - c^2 u(x) = e^x$, $0 < x < l$, $u(0) = 3$, $u'(l) = 10$.

7.3.10. Obtain an eigenfunction expansion for the Green's function $K(x;\xi)$ determined from (7.3.19) and (7.3.14) and show that the result obtained is the Fourier sine series representation of the function $K(x;\xi)$ given in (7.3.20).

7.3.11. Show that the following Green's function problem has no solution: $\partial/\partial x\,[p(x)\,\partial K(x;\xi)/\partial x] = -\delta(x-\xi)$, $0 < x, \xi < l$, $\partial K(0;\xi)/\partial x = \partial K(l;\xi)/\partial x = 0$. *Hint*: Integrate from $x = 0$ to $x = l$.

7.3.12. Obtain the modified Green's function $\hat{K}(x;\xi)$ that satisfies the problem $\partial/\partial x\left[p(x)\partial\hat{K}(x;\xi)/\partial x\right] = -\delta(x-\xi) + 1/l$, $0 < x, \xi < l$, $\partial\hat{K}(0;\xi)/\partial x = \partial\hat{K}(l;\xi)/\partial x = 0$, in the form $\hat{K}(x;\xi) = \begin{cases} \frac{1}{l}\int_\xi^x s/p(s)\,ds, & 0 < x < \xi, \\ \frac{1}{l}\int_\xi^x (s-l)/p(s)\,ds, & \xi < x < l. \end{cases}$

[Note that $\hat{K}(x;\xi)$ is determined up to an arbitrary constant.] Show how $\hat{K}(x;\xi)$ may be used to solve the boundary value problem $-(p(x)u'(x))' = f(x)$, $0 < x < l$, $u'(0) = a$, $u'(l) = b$, assuming that this problem has a solution. Integrate the equation for $u(x)$ from 0 to l to determine a condition for the solution to exist.

7.3.13. Construct a modified Green's function $\tilde{K}(x;\xi)$ that satisfies the problem $\partial/\partial x\left[p(x)\partial\tilde{K}(x;\xi)/\partial x\right] = -\delta(x-\xi)$, $0 < x, \xi < l$, $\partial\tilde{K}(0;\xi)/\partial x = A$, $\partial\tilde{K}(l;\xi)/\partial x = B$. Choose the constants A and B such that the problem for $\tilde{K}(x;\xi)$

has a solution. Then show how $\tilde{K}(x;\xi)$ can be used to solve the boundary value problem for $u(x)$ given in Exercise 7.3.12.

7.3.14. Verify that (7.3.93) is the solution of the problem for (7.3.76).

7.3.15. Verify that (7.3.84) is the solution of (7.3.82) with $N_k(T) = 0$.

7.3.16. Show that if the Green's function $K(\mathbf{x};\boldsymbol{\xi})$ determined from (7.3.1)–(7.3.2) has the eigenfunction expansion (7.3.6), then the Green's function $K(\mathbf{x};\boldsymbol{\xi})$ determined from the problem $LK(\mathbf{x};\boldsymbol{\xi}) - \hat{\lambda}\rho(\mathbf{x})K(\mathbf{x};\boldsymbol{\xi}) = \delta(\mathbf{x}-\boldsymbol{\xi})$, with the boundary condition (7.3.2), is given as $K(\mathbf{x};\boldsymbol{\xi}) = \sum_{k=1}^{\infty} \frac{M_k(\boldsymbol{\xi})M_k(\mathbf{x})}{\lambda_k - \hat{\lambda}}$. (Note that this Green's function is not defined in the preceding form if $\hat{\lambda} = \lambda_k$ for some k.)

7.3.17. Use the result of Exercise 7.3.16 to determine the Green's function for the elliptic equation $\nabla^2 K(x,y;\xi,\eta) + \hat{\lambda}K(x,y;\xi,\eta) = -\delta(x-\xi)\,\delta(y-\eta)$, in the rectangle $0 < x < l$, $0 < y < \hat{l}$ with the Dirichlet boundary condition (7.3.27), from the result (7.3.41).

7.3.18. Construct the Green's function $K(x,y;\xi,\eta)$ for Laplace's equation in the rectangle $0 < x < l$, $0 < y < \hat{l}$, where $K(x,y;\xi,\eta)$ satisfies (7.3.26) and the boundary conditions $K(x,0;\xi,\eta) = K(x,\hat{l};\xi,\eta) = 0$, $0 < x < l$, $K(0,y;\xi,\eta) = \partial K(l,y;\xi,\eta)/\partial x = 0$, $0 < y < \hat{l}$. Use the method presented in the discussion following (7.3.7).

7.3.19. Obtain the Green's function for Laplace's equation in a disk $0 \le r \le R$ with a homogeneous Dirichlet boundary condition on $r = R$. Express the Laplacian in polar coordinates and obtain the equation for the Green's function $K(r,\theta;\rho,\phi)$ in the form

$$\frac{\partial^2 K(r,\theta;\rho,\phi)}{\partial r^2} + \frac{1}{r}\frac{\partial K(r,\theta;\rho,\phi)}{\partial r} + \frac{1}{r^2}\frac{\partial^2 K(r,\theta;\rho,\phi)}{\partial \theta^2} = -\frac{\delta(r-\rho)\delta(\theta-\phi)}{r},$$

where $0 < r,\ \rho < R$ and $0 \le \theta,\ \phi \le 2\pi$, and the right side of the equation is the two-dimensional delta function in polar coordinates. The boundary condition is $K(R,\theta;\rho,\phi) = 0$. Construct an eigenfunction expansion of $K(r,\theta;\rho,\phi)$ based on the eigenvalue problem for the operator $-\partial^2/\partial\theta^2$ with periodic boundary conditions; that is, $-M''(0) = \lambda M(0)$ $(-\pi < 0 < \pi)$ with $M(-\pi) = M(\pi)$ and $M'(-\pi) = M'(\pi)$. Obtain an equation of the general form (7.3.25) in the radial variable and solve it subject to the conditions that the solution is bounded at $r = 0$ and vanishes at $r = R$.

7.3.20. Solve the problem of Exercise 7.3.19 if the disk is replaced by an annular region $R_1 \le r \le R_2$ and $K(r,\theta;\rho,\phi)$ vanishes on the boundary.

7.3.21. Obtain the Green's function in the disk $0 \le r \le R$ for the reduced wave equation with a Dirichlet boundary condition on $r = R$. Proceed as in Exercise 7.3.19 and obtain $\partial^2 K(r,\theta;\rho,\phi)/\partial r^2 + (1/r)\partial K(r,\theta;\rho,\phi)/\partial r + (1/r^2)\partial^2 K(r,\theta;\rho,\phi)/\partial\theta^2 + k^2 K(r,\theta;\rho,\phi) = -\delta(r-\rho)\delta(\theta-\phi)/r$, and the condition $K(R,\theta;\rho,\phi) = 0$. The only difference in the treatment of this problem is that the equation in the radial variable now has the form of Bessel's equation. Two independent solutions of Bessel's

equation of order n are $J_n(z)$ and $Y_n(z)$, the Bessel and Neumann functions, respectively.

7.3.22. Consider the modified Green's functions $\hat{K}(r,\theta;\rho,\phi)$ and $\tilde{K}(r,\theta;\rho,\phi)$ for the Neumann problem for Laplace's equation in the disk $0 \leq r \leq R$. Construct a function $f(r)$ that is smooth in the disk and satisfies the condition $f'(R) = -1/2\pi R$. Use $f(r)$ to transform the problem for $\tilde{K}(r,\theta;\rho,\phi)$ to a new problem with a homogeneous boundary condition. Discuss the relationship between this new problem and that for the function $\hat{K}(r,\theta;\rho,\phi)$.

7.3.23. Obtain the form of the eigenfunction expansion (7.3.79) for $K(\mathbf{x},t;\boldsymbol{\xi},\tau)$ if $\lambda_0 = 0$ is an eigenvalue.

7.3.24. Obtain the form of the expansion (7.3.85) for $K(\mathbf{x},t;\boldsymbol{\xi},\tau)$ if $\lambda_0 = 0$ is an eigenvalue.

7.3.25. Determine the Green's function for the wave equation $u_{tt}(x,y,t) - u_{xx}(x,y,t) - u_{yy}(x,y,t) = 0$ in the rectangle $0 < x < l$, $0 < y < \hat{l}$ with the boundary condition $u(x,y,t) = 0$ on the rectangle. That is, find the solution of (7.3.74)–(7.3.75) with $\rho(\mathbf{x}) = p(\mathbf{x}) = 1$ and $q(\mathbf{x}) = 0$ in (7.3.74) and $\alpha(\mathbf{x}) = 1$, $\beta(\mathbf{x}) = 0$ in the boundary condition (7.3.75). [Here $\mathbf{x}$ corresponds to (x,y).]

7.3.26. Construct the Green's function for the one-dimensional wave equation $u_{tt}(x,t) - u_{xx}(x,t) = 0$ in the interval $0 < x < l$ with the following boundary conditions: (a) $u(0,t) = u(l,t) = 0$, $t > 0$; (b) $u(0,t) = u_x(l,t) = 0$, $t > 0$; (c) $u_x(0,t) = u_x(l,t) = 0$, $t > 0$.

7.3.27. Use the Green's functions determined in Exercise 7.3.26 to write down the solution of the wave equation $u_{tt}(x,t) - u_{xx}(x,t) = 0$ in the interval $0 < x < l$, with the initial data $u(x,0) = \sin x$, $u_t(x,0) = 1$, and the boundary conditions: (a) $u(0,t) = e^{-t}$, $u(l,t) = 0$, $t > 0$; (b) $u(0,t) = 0$, $u_x(l,t) = 1$, $t > 0$; (c) $u(0,t) = \sin(t)$, $u_x(l,t) = 0$, $t > 0$.

7.3.28. Verify that (7.3.85) is a generalized solution of (7.3.81).

7.3.29. Construct the Green's function for the one-dimensional heat equation $u_t(x,t) - u_{xx}(x,t) = 0$ in the interval $0 < x < l$ with the following boundary conditions: (a) $u(0,t) = u(l,t) = 0$, $t > 0$; (b) $u(0,t) = u_x(l,t) = 0$, $t > 0$; (c) $u_x(0,t) = u_x(l,t) = 0$, $t > 0$.

7.3.30. Use the Green's functions obtained in Exercise 7.3.29 to solve the initial and boundary value problems for the heat equation $u_t(x,t) - u_{xx}(x,t) = 0$ in the interval $0 < x < l$ with $u(x,0) = 1$, and the boundary data: (a) $u(0,t) = e^{-t}$, $u(l,t) = 0$, $t > 0$; (b) $u(0,t) = \sin(t)$, $u_x(l,t) = 0$, $t > 0$; (c) $u(0,t) = 1$, $u_x(l,t) = 0$, $t > 0$.

7.3.31. Determine the Green's function for Laplace's equation in the cube $0 < x < l$, $0 < y < l$, $0 < z < l$ with Dirichlet conditions on the boundary. Use the eigenfunctions for the rectangle obtained in Example 7.6.

7.4 GREEN'S FUNCTIONS FOR UNBOUNDED REGIONS

The finite and infinite Fourier transforms introduced in Chapters 4 and 5 can be used to determine Green's functions for problems given over unbounded spatial regions. Rather than give a general discussion for equations of different types, we determine specific Green's functions for problems that involve the *heat, wave, Klein-Gordon, modified telegrapher's*, and *reduced wave equations*. (Each of these equations has constant coefficients.) The main tool will be the Fourier transforms of Sections 5.2 and 5.4, and properties of generalized functions presented in Section 7.2 will be used. Some of the results obtained in Chapter 6 will be reproduced.

Green's Functions for the Heat Equation in an Unbounded Region

We begin by considering the heat equation in one dimension over the infinite interval $-\infty < x < \infty$. The medium is assumed to be homogeneous. In view of Section 7.1, the Green's function $K(x,t;\xi,\tau)$ is a solution of the equation

$$\frac{\partial K(x,t;\xi,\tau)}{\partial t} + c^2 \frac{\partial^2 K(x,t;\xi,\tau)}{\partial x^2} = -\delta(x-\xi)\delta(t-\tau), \quad -\infty < x < \infty,\ t,\tau < T, \tag{7.4.1}$$

and satisfies the initial condition $K(x,T;\xi,\tau) = 0$.

We use the one-dimensional *Fourier transform* in x to solve the initial value problem for $K(x,t;\xi,\tau)$. Let the Fourier transform of $K(x,t;\xi,\tau)$ be denoted by $k(\lambda,t;\xi,\tau)$. Then

$$k(\lambda,t;\xi,\tau) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{i\lambda x} K(x,t;\xi,\tau)\, dx. \tag{7.4.2}$$

To obtain an equation for $k(\lambda,t;\xi,\tau)$, we multiply (7.4.1) by $(1/\sqrt{2\pi})e^{i\lambda x}$ and integrate from $-\infty$ to $+\infty$. Using (5.2.11) we obtain the equation

$$\frac{\partial k(\lambda,t;\xi,\tau)}{\partial t} - (c\lambda)^2 k(\lambda,t;\xi,\tau) = \frac{-1}{\sqrt{2\pi}}\, e^{i\lambda\xi}\, \delta(t-\tau), \qquad t < T. \tag{7.4.3}$$

Also, $k(\lambda,t;\xi,\tau)$ satisfies the condition $k(\lambda,T;\xi,\tau) = 0$. The equation (7.4.3) has the form of equation (7.3.82) with $N_k(t)$ replaced by $k(\lambda,t;\xi,\tau)$, λ_k given as $(c\lambda)^2$, and $M_k(\xi)$ given as $(1/\sqrt{2\pi})\, e^{i\lambda\xi}$. The initial condition at $t = T$ is identical for both functions $N_k(t)$ and $k(\lambda,t;\xi,\tau)$. Therefore, the solution $k(\lambda,t;\xi,\tau)$ is obtained from (7.3.84) as

$$k(\lambda,t;\xi,\tau) = \frac{1}{\sqrt{2\pi}} \exp\left[(c\lambda)^2(t-\tau) + i\lambda\xi\right] H(\tau - t). \tag{7.4.4}$$

The inversion formula for the Fourier transform then gives

$$K(x,t;\xi,\tau) = \frac{1}{2\pi} H(\tau - t) \int_{-\infty}^{\infty} \exp[-i\lambda(x-\xi) - (c\lambda)^2(\tau - t)]\, d\lambda. \tag{7.4.5}$$

This integral was evaluated in Example 5.2. Using the results of that example gives the *Green's function* as

$$K(x,t;\xi,\tau) = \frac{H(\tau-t)}{\sqrt{4\pi c^2(\tau-t)}} \exp\left[-\frac{(x-\xi)^2}{4c^2(\tau-t)}\right]. \tag{7.4.6}$$

We note that $K(x,t;\xi,\tau)$ is symmetric in x and ξ but not in t and τ. In fact, there is a reversal in time in relation to the fundamental solution $G(x-\xi,t-\tau)$ defined in (5.2.39). This results because $K(x,t;\xi,\tau)$ satisfies the *backward heat equation*, whereas $G(x-\xi,t-\tau)$ satisfies the *forward heat equation*.

By using the two-dimensional Fourier transform it is easy to show that the two-dimensional Green's function is

$$K(x,y,t;\xi,\eta,\tau) = \frac{H(\tau-t)}{4\pi c^2(\tau-t)} \exp\left[-\frac{(x-\xi)^2+(y-\eta)^2}{4c^2(\tau-t)}\right]. \tag{7.4.7}$$

Similarly, the three-dimensional Fourier transform yields

$$K(x,y,z,t;\xi,\eta,\zeta,\tau) = \frac{H(\tau-t)}{[4\pi c^2(\tau-t)]^{3/2}} \exp\left[-\frac{(x-\xi)^2+(y-\eta)^2+(z-\zeta)^2}{4c^2(\tau-t)}\right] \tag{7.4.8}$$

as the three-dimensional Green's function. We will refer to (7.4.6)–(7.4.8) as the *free space Green's functions* for the heat equation. In each case we obtain the *causal fundamental solution* for the heat equation by replacing t by $-t$ and τ by $-\tau$. The results agree with (6.7.40).

Given the initial value problem for $u(x,t)$,

$$u_t(x,t) - c^2 u_{xx}(x,t) = F(x,t), \qquad -\infty < x < \infty, \quad t > 0, \tag{7.4.9}$$

with the initial condition $u(x,0) = f(x)$, the Green's function $K(x,t;\xi,\tau)$ may be used to obtain the solution $u(x,t)$. With $-\infty < \xi < \infty$ and $\tau > 0$, the solution formula is

$$u(\xi,\tau) = \int_0^T \int_{-\infty}^{\infty} K(x,t;\xi,\tau)F(x,t)\,dx\,dt + \int_{-\infty}^{\infty} K(x,0;\xi,\tau)f(x)\,dx, \tag{7.4.10}$$

as follows from (7.1.35) in Section 7.1, with $K(x,t;\xi,\tau)$ given in (7.4.6). Now

$$\int_0^T \int_{-\infty}^{\infty} K(x,t;\xi,\tau)F(x,t)\,dx\,dt = \int_0^{\tau} \int_{-\infty}^{\infty} \frac{F(x,t)\,e^{-(x-\xi)^2/4c^2(\tau-t)}}{\sqrt{4\pi c^2(\tau-t)}}\,dx\,dt, \tag{7.4.11}$$

since $H(\tau-t) = 0$ for $t > \tau$. Also,

$$\int_{-\infty}^{\infty} K(x,0;\xi,\tau)f(x)\,dx = \frac{1}{\sqrt{4\pi c^2\tau}} \int_{-\infty}^{\infty} f(x)e^{-(x-\xi)^2/4c^2\tau}\,dx, \tag{7.4.12}$$

since $H(\tau) = 1$ for $\tau > 0$. It is seen immediately that the solution (7.4.10) agrees with that given in Example 5.2.

Green's Functions for the Wave Equation in an Unbounded Region

We construct the *Green's function* $K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)$, where $\mathbf{x} = x, y, z$ and $\boldsymbol{\xi} = \xi, \eta, \zeta$, appropriate for the solution of the Cauchy problem for the wave equation in three dimensions. Using this function, the two-dimensional Green's function is then obtained. In addition, the Green's function for the one-dimensional problem is discussed.

In three space dimensions, the Green's function $K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)$ satisfies the equation

$$\frac{\partial^2 K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)}{\partial t^2} - c^2 \nabla^2 K(\mathbf{x}, t; \boldsymbol{\xi}, \tau) = \delta(\mathbf{x} - \boldsymbol{\xi})\delta(t - \tau), \tag{7.4.13}$$

with $\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$, $\delta(\mathbf{x} - \boldsymbol{\xi}) = \delta(x - \xi)\delta(y - \eta)\delta(z - \zeta)$, $-\infty < x, y, z, \xi, \eta, \zeta < \infty$, and $t, \tau < T$. The conditions at $t = T$ are

$$K(\mathbf{x}, T; \boldsymbol{\xi}, \tau) = \frac{\partial K(\mathbf{x}, T; \boldsymbol{\xi}, \tau)}{\partial t} = 0. \tag{7.4.14}$$

To solve, we use the three-dimensional Fourier transform

$$k(\boldsymbol{\lambda}, t; \boldsymbol{\xi}, \tau) = \frac{1}{\left(\sqrt{2\pi}\right)^3} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{i(\lambda_1 x + \lambda_2 y + \lambda_3 z)} K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)\, dx\, dy\, dz, \tag{7.4.15}$$

where $\boldsymbol{\lambda} = \lambda_1, \lambda_2, \lambda_3$. We obtain an equation for $k(\boldsymbol{\lambda}, t; \boldsymbol{\xi}, \tau)$, by multiplying $1/\left(\sqrt{2\pi}\right)^3 \exp[i(\lambda_1 x + \lambda_2 y + \lambda_3 z)]$ into (7.4.13) and integrating with respect to x, y, and z from $-\infty$ to ∞. Using the properties of the Fourier transform (see Example 5.7) gives (with $|\boldsymbol{\lambda}| = \sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}$)

$$\frac{\partial^2 k(\boldsymbol{\lambda}, t; \boldsymbol{\xi}, \tau)}{\partial t^2} + c^2 |\boldsymbol{\lambda}|^2 k(\boldsymbol{\lambda}, t; \boldsymbol{\xi}, \tau) = \frac{1}{\left(\sqrt{2\pi}\right)^3} \exp[i(\lambda_1 \xi + \lambda_2 \eta + \lambda_3 \zeta)]\, \delta(t - \tau) \tag{7.4.16}$$

and the conditions $k(\boldsymbol{\lambda}, T; \boldsymbol{\xi}, \tau) = \partial k(\boldsymbol{\lambda}, T; \boldsymbol{\xi}, \tau)/\partial t = 0$. This problem is equivalent to that for $N_k(t)$ given in (7.3.76). Using (7.3.78) gives

$$k(\boldsymbol{\lambda}, t; \boldsymbol{\xi}, \tau) = \frac{1}{\left(\sqrt{2\pi}\right)^3 c|\boldsymbol{\lambda}|} \sin[c|\boldsymbol{\lambda}|(\tau - t)] \exp[i(\lambda_1 \xi + \lambda_2 \eta + \lambda_3 \zeta)] H(\tau - t). \tag{7.4.17}$$

Inverting the Fourier transform yields

$$K(\mathbf{x}, t; \boldsymbol{\xi}, \tau) = \frac{H(\tau - t)}{(2\pi)^3} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\sin[c|\boldsymbol{\lambda}|(\tau - t)]}{c|\boldsymbol{\lambda}|} \times \exp\{-i[\lambda_1(x - \xi) + \lambda_2(y - \eta) + \lambda_3(z - \zeta)]\}\, d\lambda_1\, d\lambda_2\, d\lambda_3. \tag{7.4.18}$$

This integral, which is to be interpreted in a generalized sense, can be evaluated by transforming to spherical coordinates. Let $|\boldsymbol{\lambda}|$, θ, and ϕ be *spherical coordinates* with $|\boldsymbol{\lambda}| \geq 0$, $0 \leq \theta \leq 2\pi$, and $0 \leq \phi \leq \pi$. The polar axis (i.e., $\phi = 0$)

is chosen to coincide with the (half) line connecting the origin (i.e., $\lambda_1 = \lambda_2 = \lambda_3 = 0$) to the fixed point $(x-\xi,\ y-\eta,\ z-\zeta)$. Since $(x-\xi,\ y-\eta,\ z-\zeta)$ is on the polar axis, we have $\lambda_1(x-\xi)+\lambda_2(y-\eta)+\lambda_3(z-\zeta) = |\boldsymbol{\lambda}| r\cos(\phi)$, where $r = \sqrt{(x-\xi)^2+(y-\eta)^2+(z-\zeta)^2}$. The volume element is given as $d\lambda_1\, d\lambda_2\, d\lambda_3 = |\boldsymbol{\lambda}|^2 \sin(\phi) d|\boldsymbol{\lambda}|\, d\theta\, d\phi$, and $K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)$ of (7.4.18) takes the form

$$K = \frac{H(\tau-t)}{c(2\pi)^3}\int_0^{\infty}\int_0^{\pi}\int_0^{2\pi} \sin[c|\boldsymbol{\lambda}|(\tau-t)] e^{-i|\boldsymbol{\lambda}| r\cos(\phi)} |\boldsymbol{\lambda}| \sin(\phi)\, d\theta\, d\phi\, d|\boldsymbol{\lambda}|. \tag{7.4.19}$$

The two inner integrals are easily evaluated, and their contribution is found to be $(4\pi/|\boldsymbol{\lambda}|r)\sin[|\boldsymbol{\lambda}|r]$. Then, (7.4.19) can be evaluated as follows:

$$\begin{aligned} K(\mathbf{x}, t; \boldsymbol{\xi}, \tau) &= \frac{H(\tau-t)}{2\pi^2 cr}\int_0^{\infty} \sin[c|\boldsymbol{\lambda}|(\tau-t)]\sin[|\boldsymbol{\lambda}|r]\, d|\boldsymbol{\lambda}| \\ &= \frac{H(\tau-t)}{4\pi^2 cr}\int_0^{\infty} (\cos\{|\boldsymbol{\lambda}|[c(\tau-t)-r]\} - \cos\{|\boldsymbol{\lambda}|[c(\tau-t)+r]\})\, d|\boldsymbol{\lambda}| \\ &= \frac{H(\tau-t)}{8\pi^2 cr}\int_0^{\infty} (\exp\{i|\boldsymbol{\lambda}|[c(\tau-t)-r] - \exp\{i|\boldsymbol{\lambda}|[c(\tau-t)+r]\})\, d|\boldsymbol{\lambda}| \\ &= \frac{H(\tau-t)}{4\pi cr}\{\delta[c(\tau-t)-r] - \delta[c(\tau-t)+r]\}, \end{aligned} \tag{7.4.20}$$

where (7.2.51) has been used. Now $\delta[c(\tau-t)+r] = 0$ since for $\tau - t > 0$ and $r > 0$, $c(\tau-t)+r \neq 0$. For the same reason, $\delta[c(\tau-t)-r]$ vanishes for $\tau - t < 0$ so that there is no need to include the Heaviside function in (7.4.20). Thus $K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)$ is given as

$$K(x, y, z, t; \xi, \eta, \zeta, \tau) = \frac{\delta[c(\tau-t)-r]}{4\pi cr}, \tag{7.4.21}$$

with $r = \sqrt{(x-\xi)^2+(y-\eta)^2+(z-\zeta)^2}$.

To determine the Green's function in two (space) dimensions, we integrate the Green's function (7.4.21) with respect to ζ from $-\infty$ to $+\infty$. The resulting Green's function $K = K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)$, where we now have $\mathbf{x} = x, y$ and $\boldsymbol{\xi} = \xi, \eta$, is independent of z and ζ and satisfies the equation

$$\frac{\partial^2 K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)}{\partial t^2} - c^2\left[\frac{\partial^2 K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)}{\partial x^2} + \frac{\partial^2 K(\mathbf{x}, t; \boldsymbol{\xi}, \tau)}{\partial y^2}\right] = \delta(x-\xi)\delta(y-\eta)\delta(t-\tau) \tag{7.4.22}$$

for $-\infty < x,\ y,\ \xi,\ \eta < \infty$, and $t,\ \tau < T$. In effect, we are considering a line source in three-dimensional space, and because of the homogeneity of the medium (there are constant coefficients in the equation) this corresponds to a point source in two dimensions. If we formally integrate (7.4.13) with respect to ζ and assume that K is independent of z, then (7.4.22) results.

On integrating over ζ, the two-dimensional Green's function becomes

$$K(\mathbf{x}, t; \boldsymbol{\xi}, \tau) = \frac{1}{4\pi c}\int_{-\infty}^{\infty} \frac{\delta[c(\tau-t)-r]}{r}\, d\zeta. \tag{7.4.23}$$

Introducing the change of variables $s = z - \zeta$ gives

$$K(\mathbf{x}, t; \boldsymbol{\xi}, \tau) = \frac{1}{2\pi c} \int_0^\infty \frac{\delta[c(\tau - t) - r]}{r}\, ds, \tag{7.4.24}$$

where $r = \sqrt{(x-\xi)^2 + (y-\eta)^2 + s^2}$ and the integral from 0 to ∞ results since the integrand is an even function of s. A further change of variables $r^2 = \rho^2 + s^2$, with $\rho^2 = (x-\xi)^2 + (y-\eta)^2$, yields $ds/r = dr/s = dr/\sqrt{r^2 - \rho^2}$, and (7.4.24) becomes

$$K(\mathbf{x}, t; \boldsymbol{\xi}, \tau) = \frac{1}{2\pi c} \int_\rho^\infty \frac{\delta[c(\tau - t) - r]}{\sqrt{r^2 - \rho^2}}\, dr. \tag{7.4.25}$$

If $c(\tau - t) < \rho$, the integral vanishes, since the argument of the delta function is negative in that case. If $c(\tau - t) > \rho$, the substitution property of the delta function yields $K(\mathbf{x}, t; \boldsymbol{\xi}, \tau) = (1/2\pi c)(1/\sqrt{c^2(\tau - t)^2 - \rho^2})$. With the use of the Heaviside function, both results can be combined into a single expression,

$$K(x, y, t; \xi, \eta, \tau) = \frac{1}{2\pi c} \frac{H[c(\tau - t) - \rho]}{\sqrt{c^2(\tau - t)^2 - \rho^2}}, \tag{7.4.26}$$

where $\rho^2 = (x-\xi)^2 + (y-\eta)^2$.

In the one-dimensional problem the Green's function $K(x, t; \xi, \tau)$ satisfies the equation

$$\frac{\partial^2 K(x, t; \xi, \tau)}{\partial t^2} - c^2 \frac{\partial^2 K(x, t; \xi, \tau)}{\partial x^2} = \delta(x - \xi)\delta(t - \tau), \tag{7.4.27}$$

with $-\infty < x, \xi, < \infty$, and $t, \tau < T$. Again the conditions at $t = T$ are $K(x, T; \xi, \tau) = \partial K(x, T; \xi, \tau)/\partial t = 0$. It is a simple matter (see the exercises) to adapt the general solution of the Cauchy problem to show that $K(x, t; \xi, \tau)$ is given as

$$K(x, t; \xi, \tau) = \begin{cases} 1/2c, & |x - \xi| < c(\tau - t), \\ 0, & |x - \xi| > c(\tau - t). \end{cases} \tag{7.4.28}$$

$K(x, t; \xi, \tau)$ can be expressed in terms of Heaviside functions as

$$K(x, t; \xi, \tau) = \frac{1}{2c} H[x - ct - (\xi - c\tau)] H[\xi + c\tau - (x + ct)], \tag{7.4.29}$$

and it is straightforward to show that this is a solution of the problem.

A common feature of the Green's functions (7.4.21), (7.4.26), and (7.4.29) is that they are symmetric in $\mathbf{x}$ and $\boldsymbol{\xi}$, where $\mathbf{x}$ is the observation point and $\boldsymbol{\xi}$ is the source point in two or three dimensions and in x and ξ in one dimension. We shall refer to them as the *free space Green's functions* for the wave equation. If we replace t by $-t$ and τ by $-\tau$, we obtain the *causal fundamental solution* for each of the problems. Let the fundamental solutions be denoted by S. In the three-dimensional case we have

$$S = \frac{\delta[c(t - \tau) - r]}{4\pi c r}, \qquad r = \sqrt{(x-\xi)^2 + (y-\eta)^2 + (z-\zeta)^2}. \tag{7.4.30}$$

In two dimensions

$$S = \frac{1}{2\pi c} \frac{H[c(t-\tau)-\rho]}{\sqrt{c^2(t-\tau)^2-\rho^2}}, \qquad \rho = \sqrt{(x-\xi)^2+(y-\eta)^2}, \tag{7.4.31}$$

whereas in one dimension

$$S = \frac{1}{2c} H[c(t-\tau)+(x-\xi)]H[c(t-\tau)-(x-\xi)]. \tag{7.4.32}$$

Each of these functions satisfies the same equation as the Green's function K, except that for each of them we have $S|_{t=0} = \partial S/\partial t|_{t=0} = 0$. Even though they do not all have the same form as the causal fundamental solutions given in (6.7.37)–(6.7.39), they can be shown to be identical on using the properties of the Dirac delta and Heaviside functions.

The fundamental solutions yield a vivid distinction between the nature of wave propagation in two and three dimensions as characterized by solutions of the wave equation. (These differences have already been discussed in Example 5.7.) Given the source point (ξ, η, ζ) and the time τ, the forward characteristic cone for the wave equation is $c(t-\tau) = \sqrt{(x-\xi)^2+(y-\eta)^2+(z-\zeta)^2}, t \geq \tau$. For the two-dimensional problem with the source point (ξ, η) and the time τ we have $c(t-\tau) = \sqrt{(x-\xi)^2+(y-\eta)^2}$, $t \geq \tau$. Noting the behavior of the delta function, (7.4.30) indicates that the disturbance due to a point source at (ξ, η, ζ) acting at the time τ, is concentrated on a sphere of radius $c(t-\tau)$ with center at (ξ, η, ζ) at the later time t and vanishes elsewhere, in the three-dimensional case. For the two-dimensional problem, the disturbance resulting from a point source at (ξ, η) acting the time τ, is distributed throughout a circle of radius $c(t-\tau)$ with center at (ξ, η) at the later time t. Consequently, both disturbances spread out with the speed c in all directions. However, in the three-dimensional case there is a sharp wave front that leaves no wake and the disturbance is only felt instantaneously at any point. In the two-dimensional problem the solution does not return to zero as soon as the wave front passes, but there is a wake that decays like $1/c(t-\tau)$ as t increases. The sharp signals that occur only in three-dimensional wave propagation and not in the two-dimensional case characterize what is known as *Huygens' principle*.

Green's Functions for the Klein-Gordon Equation and the Modified Telegrapher's Equation

In connection with the Cauchy problem for the one-dimensional *Klein-Gordon equation* (5.7.6), we are led to consider the Green's function $K(x,t;\xi,\tau)$ that satisfies the equation

$$\frac{\partial^2 K(x,t;\xi,\tau)}{\partial t^2} - \gamma^2 \frac{\partial^2 K(x,t;\xi,\tau)}{\partial x^2} + c^2 K(x,t;\xi,\tau) = \delta(x-\xi)\delta(t-\tau), \tag{7.4.33}$$

with $-\infty < x, \xi, < \infty$, and $t, \tau < T$. The conditions at $t = T$ are $K(x,T;\xi,\tau) = \partial K(x,T;\xi,\tau)/\partial t = 0$.

On replacing t by $-t$ and τ by $-\tau$ in the causal fundamental solution (6.7.45) and adjusting the parameters, we conclude that the *Green's function* $K(x,t;\xi,\tau)$ should be given as

$$K(x,t;\xi,\tau) = \begin{cases} (1/2\gamma)J_0\left[(c/\gamma)\sqrt{\gamma^2(t-\tau)^2-(x-\xi)^2}\right], & |x-\xi| < \gamma(t-\tau), \\ 0, & |x-\xi| > \gamma(t-\tau), \end{cases} \tag{7.4.34}$$

where J_0 is the Bessel function of order zero. Noting (7.4.29), we can express $K(x,t;\xi,\tau)$ as

$$K = \frac{1}{2\gamma}J_0\left[\frac{c}{\gamma}\sqrt{\gamma^2(t-\tau)^2-(x-\xi)^2}\right]H[x-\xi-\gamma(t-\tau)]H[\xi-x+\gamma(\tau-t)]. \tag{7.4.35}$$

Since $J_0(0) = 1$ we see that (7.4.35) reduces to the Green's function for the one-dimensional wave equation if we set $c = 0$. It is straightforward to verify that (7.4.35) is a solution of the Green's function problem.

It may be noted that with $c = i\hat{c}$ (where $i = \sqrt{-1}$) in the Klein-Gordon equation (7.4.33) we obtain the *modified telegrapher's equation*

$$\frac{\partial^2 K(x,t;\xi,\tau)}{\partial t^2} - \gamma^2\frac{\partial^2 K(x,t;\xi,\tau)}{\partial x^2} - \hat{c}^2 K(x,t;\xi,\tau) = \delta(x-\xi)\delta(t-\tau). \tag{7.4.36}$$

This equation results when the first time derivative term is eliminated from the *telegrapher's equation* (see the exercises). Since $J_0(ix) = I_0(x)$, the modified Bessel function of zero order, we obtain for $K = K(x,t;\xi,\tau)$ in place of (7.4.35) [see also (6.7.46)],

$$K = \frac{1}{2\gamma}I_0\left[\frac{\hat{c}}{\gamma}\sqrt{\gamma^2(t-\tau)^2-(x-\xi)^2}\right]H[x-\xi-\gamma(t-\tau)]H[\xi-x+\gamma(\tau-t)]. \tag{7.4.37}$$

Again, (7.4.35) and (7.4.37) are referred to as *free space Green's functions.*

Given the initial value problem for the *Klein-Gordon equation,*

$$u_{tt}(x,t) - \gamma^2 u_{xx}(x,t) + c^2 u(x,t) = F(x,t), \qquad -\infty < x < \infty,\ t > 0, \tag{7.4.38}$$

with the initial conditions

$$u(x,0) = f(x), \qquad u_t(x,0) = g(x), \qquad -\infty < x < \infty, \tag{7.4.39}$$

the solution at an arbitrary point (ξ,τ) is given in terms of (7.4.35) as

$$u(\xi,\tau) = \int_0^T \int_{-\infty}^{\infty} K(x,t;\xi,\tau)F(x,t)\,dx\,dt + \int_{-\infty}^{\infty}\left[K(x,0;\xi,\tau)g(x) - \frac{\partial K(x,0;\xi,\tau)}{\partial t}f(x)\right]dx, \tag{7.4.40}$$

where $\tau < T$ and (7.1.34) has been modified appropriately.

The solution (7.4.40) can be expressed in a simple form in terms of the data for the problem, as we now demonstrate. In the double integral in (7.4.40) we have $K(x,t;\xi,\tau) = 0$ for $t > \tau$, so that the upper limit in the t integral extends only up to τ. Also, from (7.4.38) we conclude that $K(x,t;\xi,\tau)$ vanishes unless $|x-\xi| < \gamma(\tau - t)$, and this is equivalent to $\xi - \gamma(\tau - t) < x < \xi + \gamma(\tau - t)$. Therefore, we obtain $\int_0^T \int_{-\infty}^{\infty} K(x,t;\xi,\tau)F(x,t)\,dx\,dt = \frac{1}{2\gamma}\int_0^{\tau}\int_{\xi-\gamma(\tau-t)}^{\xi+\gamma(\tau-t)} F(x,t)$ $J_0\left[\frac{c}{\gamma}\sqrt{\gamma^2(t-\tau)^2-(x-\xi)^2}\right]\,dx\,dt$. Further, we have $K(x,0;\xi,\tau) = \frac{1}{2\gamma}$ $J_0\left[\frac{c}{\gamma}\sqrt{\gamma^2\tau^2-(x-\xi)^2}\right]H[x-(\xi-\gamma\tau)]H[\xi+\gamma\tau-x]$, so that $\int_{-\infty}^{\infty} K(x,0;\xi,\tau)$ $g(x)dx = \frac{1}{2\gamma}\int_{\xi-\gamma\tau}^{\xi+\gamma\tau} J_0\left[\frac{c}{\gamma}\sqrt{\gamma^2\tau^2-(x-\xi)^2}\right]g(x)\,dx$, as the Heaviside function product vanishes outside the interval $(\xi-\gamma\tau, \xi+\gamma\tau)$. Finally, $\partial K(x,0;\xi,\tau)/\partial t =$ $-c\tau/2\frac{J_0'\left[c\sqrt{\gamma^2(t-\tau)^2-(x-\xi)^2}/\gamma\right]}{\sqrt{\gamma^2(t-\tau)^2-(x-\xi)^2}}H[x-(\xi-\gamma\tau)]H[\xi+\gamma\tau-x] - \frac{1}{2}\frac{J_0\left[c\sqrt{\gamma^2(t-\tau)^2-(x-\xi)^2}/\gamma\right]}{\sqrt{\gamma^2(t-\tau)^2-(x-\xi)^2}}$ $\delta[x-(\xi-\gamma\tau)]H[\xi+\gamma\tau-x] = -\frac{1}{2}\frac{J_0\left[c\sqrt{\gamma^2(t-\tau)^2-(x-\xi)^2}/\gamma\right]}{\sqrt{\gamma^2(t-\tau)^2-(x-\xi)^2}}H[x-(\xi-\gamma\tau)]\delta[\xi+$ $\gamma\tau-x]$. In view of the substitution property of the delta function, the last two terms in the above reduce to $-\frac{1}{2}\delta[x-(\xi-\gamma\tau)] - \frac{1}{2}\delta[x-(\xi+\gamma\tau)]$ since $J_0(0) = 1$ and $H(2\gamma\tau) = 1$. Therefore, $\int_{-\infty}^{\infty}\partial K(x,0;\xi,\tau)/\partial t f(x)\,dx = -(c\tau/2)\int_{\xi-\gamma\tau}^{\xi+\gamma\tau}$ $\frac{J_0'\left[c\sqrt{\gamma^2t^2-(x-\xi)^2}/\gamma\right]}{\sqrt{\gamma^2\tau^2-(x-\xi)^2}}f(x)\,dx - \frac{1}{2}f(\xi-\gamma\tau) - \frac{1}{2}f(\xi+\gamma\tau)$.

Combining the foregoing results and noting that $-J_0'(x) = J_1(x)$, the Bessel function of order 1, gives the solution $u(x,t)$ of the initial value problem (7.4.38)–(7.4.39) as

$$u(x,t) = \frac{f(x-\gamma t)+f(x+\gamma t)}{2} + \frac{1}{2\gamma}\int_{x-\gamma t}^{x+\gamma t} J_0\left[\frac{c}{\gamma}\sqrt{\gamma^2t^2-(x-\xi)^2}\right]g(\xi)\,d\xi$$

$$-\frac{ct}{2}\int_{x-\gamma t}^{x+\gamma t}\frac{J_1\left[c\sqrt{\gamma^2t^2-(x-\xi)^2}/\gamma\right]}{\sqrt{\gamma^2t^2-(x-\xi)^2}}f(\xi)\,d\xi$$

$$+\frac{1}{2\gamma}\int_0^t\int_{x-\gamma(t-\tau)}^{x+\gamma(t-\tau)} F(\xi,\tau)J_0\left[\frac{c}{\gamma}\sqrt{\gamma^2(t-\tau)^2-(x-\xi)^2}\right]d\xi\,d\tau. \quad (7.4.41)$$

This solution formula reduces to that for the Cauchy problem for the inhomogeneous wave equation if we set $c = 0$.

If we set $c = i\hat{c}$ in (7.4.41) and note that $J_0(iz) = I_0(z)$ and $J_1(iz) = iI_1(z)$, we obtain as the solution of the *modified telegrapher's equation*

$$u_{tt}(x,t) - \gamma^2 u_{xx}(x,t) - \hat{c}^2 u(x,t) = F(x,t), \qquad -\infty < x < \infty,\ t > 0, \quad (7.4.42)$$

with the initial conditions (7.4.39), in the form

$$u(x,t) = \frac{f(x-\gamma t)+f(x+\gamma t)}{2} + \frac{1}{2\gamma}\int_{x-\gamma t}^{x+\gamma t} I_0\left[\frac{\hat{c}}{\gamma}\sqrt{\gamma^2 t^2-(x-\xi)^2}\right] g(\xi)\,d\xi$$

$$+\frac{\hat{c}t}{2}\int_{x-\gamma t}^{x+\gamma t} \frac{I_1\left[\hat{c}\sqrt{\gamma^2 t^2-(x-\xi)^2}/\gamma\right]}{\sqrt{\gamma^2 t^2-(x-\xi)^2}} f(\xi)\,d\xi$$

$$+\frac{1}{2\gamma}\int_0^t \int_{x-\gamma(t-\tau)}^{x+\gamma(t-\tau)} F(\xi,\tau) I_0\left[\frac{\hat{c}}{\gamma}\sqrt{\gamma^2(t-\tau)^2-(x-\xi)^2}\right] d\xi\,d\tau. \quad (7.4.43)$$

Both solutions (7.4.41) and (7.4.43) exhibit the *domains of dependence* and *influence* for the *Klein-Gordon* and *modified telegrapher's equations*. (The telegrapher's equation is discussed in the exercises.) It has already been indicated that these domains, which characterize the maximum speed at which disturbances or signals travel, are determined by the principal parts of the given equations (i.e., the second derivative terms) and do not depend on the lower-order terms. Our results show that these equations and the wave equation have identical domains of dependence and influence.

Green's Functions for Parabolic and Hyperbolic PDEs

Before considering Green's functions for elliptic equations, we examine the relationship between the Green's functions for hyperbolic and parabolic equations defined in terms of the inhomogeneous equations (7.1.23) and (7.1.32) and those defined in terms of the homogeneous equations (7.1.36) and (7.1.37). The relationship between these two determinations of Green's functions for the case of bounded regions was discussed at the end of Section 7.3.

Let the Green's functions $K(\mathbf{x},t;\boldsymbol{\xi},\tau)$ be defined in terms of the (backward) initial value problems (7.1.36)–(7.1.40) for the hyperbolic and parabolic cases with the initial data given at $t=\tau$ and the problem defined over the entire space. Then if we set $\hat{K}(\mathbf{x},t;\boldsymbol{\xi},\tau) = K(\mathbf{x},t;\boldsymbol{\xi},\tau)H(\tau-t)$ with $\tau < T$, it is easy to verify from the properties of $K(\mathbf{x},t;\boldsymbol{\xi},\tau)$ and by direct differentiation that $\hat{K}(\mathbf{x},t;\boldsymbol{\xi},\tau)$ is indeed the solution of the Green's function problem given in terms of the inhomogeneous equations (7.1.23) and (7.1.32) for the hyperbolic and parabolic cases, respectively. Thus for $t<\tau$, both of these formulations lead to the same Green's functions.

In a similar fashion, it can be shown that the *causal fundamental solutions* for hyperbolic and parabolic equations can be characterized in terms of initial value problems. If the instantaneous source point is at $(\mathbf{x},t)=(\boldsymbol{\xi},\tau)$, the initial conditions for the homogeneous form of the hyperbolic equation (7.1.11) are $u(\mathbf{x},t)=0$ and $u_t(\mathbf{x},t)=\delta(\mathbf{x}-\boldsymbol{\xi})/\rho(\mathbf{x})$ at $t=\tau$. For the homogeneous form of the parabolic equation (7.1.26) with the same source point, the initial condition becomes $u(\mathbf{x},t)=\delta(\mathbf{x}-\boldsymbol{\xi})/\rho(\mathbf{x})$ at $t=\tau$. Then the causal fundamental solutions for these problems, which we denote by $S(\mathbf{x},t;\boldsymbol{\xi},\tau)$, can be expressed in terms of solutions $u(\mathbf{x},t)$ of the aforementioned initial value problems in the form $S(\mathbf{x},t;\boldsymbol{\xi},\tau)=u(\mathbf{x},t)H(t-\tau)$.

Consequently, instantaneous point source problems and the related causal fundamental solutions need not necessarily be characterized strictly in terms of inhomogeneous differential equations, as has been done in Section 6.7 or previously in this

section, for instance. We can determine the solutions of these problems in terms of homogeneous equations with appropriate Dirac delta function initial conditions. This approach is used in Section 10.2 when we analyze the propagation of singularities for hyperbolic equations.

Green's Functions for the Reduced Wave Equation: Ocean Acoustics

Since the free space Green's functions for (self-adjoint) elliptic equations with constant coefficients have already been discussed fully in Example 6.13, we restrict our analysis of Green's functions for elliptic problems in unbounded regions to a single boundary value problem involving the reduced wave equation, which is of interest in the theory of *ocean acoustics*.

The ocean, which we assume to be homogeneous, is taken to be of infinite extent in the x and y directions and is assumed to have a constant depth $h > 0$. The plane $z = 0$ represents the surface of the ocean, and the plane $z = -h$ is the (rigid) bottom. Thus, the region G for this problem is given as $-\infty < x,\ y < \infty$, and $-h < z < 0$. A point source is located at $(x,\ y,\ z) = (0,\ 0,\ \zeta)$ with $-h < \zeta < 0$. Then, the Green's function $K(x, y, z; \zeta)$ satisfies the equation

$$\nabla^2 K(x, y, z; \zeta) + k^2 K(x, y, z; \zeta) = -\delta(x)\delta(y)\delta(z - \zeta), \tag{7.4.44}$$

and the boundary conditions

$$K(x, y, 0; \zeta) = 0, \quad \frac{K(x, y, -h; \zeta)}{\partial z} = 0. \tag{7.4.45}$$

The real part of $K(x, y, z; \zeta)e^{-i\omega t}$, where ω is the frequency and $k = \omega/c$ represents the time-harmonic acoustic pressure, satisfies a wave equation with constant wave speed c in the absence of any source terms (see Exercise 2.4.7). Thus the equation for $K(x, y, z; \zeta)$ is known as the *reduced wave equation* or the *Helmholtz equation*. The conditions (7.4.45) are appropriate for the sound pressure at the surface and at the (rigid) bottom, respectively. The homogeneity of the ocean implies that k in (7.4.44) is a constant. This Green's function problem for $K(x, y, z; \zeta)$ does not have a unique solution unless a *radiation condition* is imposed. This is done in our discussion.

To solve for the Green's function $K(x, y, z; \zeta)$ we use the method of *eigenfunction expansions*. We consider an *eigenvalue problem* for the function $M(z)$ associated with the operator $\hat{L} = -\partial^2/\partial z^2 - k^2$ and given over the interval with $-h < z < 0$. Thus, the eigenvalue problem is

$$-M''(z) - k^2 M(z) = \lambda M(z),\ -h < z < 0,\ M(0) = 0,\ M'(-h) = 0, \tag{7.4.46}$$

where λ is the eigenvalue parameter. On comparing (7.4.46) with the Sturm-Liouville problem of Section 4.3, we find that apart from the fact that the coefficient $-k^2$ of $M(z)$ is negative, every condition is the same. All properties of the eigenvalues and

eigenfunctions given in that section are valid for (7.4.46) except that there are a finite number negative eigenvalues, as will be shown.

The general solution of the ODE in (7.4.46) is $M(z) = c_1 \sin\left(z\sqrt{k^2+\lambda}\right) + c_2 \cos\left(z\sqrt{k^2+\lambda}\right)$. The condition $M(0) = 0$ implies that $c_2 = 0$. The condition $M'(-h) = 0$ implies that $\cos\left(h\sqrt{k^2+\lambda}\right) = 0$, so that λ is specified in terms of the zeros of the cosine function. The eigenvalues and eigenfunctions are

$$\lambda_n = \left(n+\frac{1}{2}\right)^2\left(\frac{\pi}{h}\right)^2 - k^2, \quad \hat{M}_n(z) = \sin\left[\left(n+\frac{1}{2}\right)\left(\frac{\pi}{h}\right)z\right] \tag{7.4.47}$$

for $n = 0, 1, 2, \ldots$. We observe that some of the eigenvalues λ_n may be negative, depending on the magnitude of k.

The inner product for this problem is $(\phi(z),\ \psi(z)) = \int_{-h}^{0} \phi(z)\psi(z)\,dz$ and $(\hat{M}_n(z), \hat{M}_m(z)) = 0$ for $m \neq n$, as can easily be shown. Also,

$$(\hat{M}_n(z), \hat{M}_n(z)) = \int_{-h}^{0} \left\{\frac{1}{2} - \frac{1}{2}\cos\left[(2n+1)\left(\frac{\pi}{h}\right)z\right]\right\} dz = \frac{h}{2}. \tag{7.4.48}$$

Thus the normalized set of eigenfunctions are

$$M_n(z) = \sqrt{\frac{2}{h}}\sin\left[\left(n+\frac{1}{2}\right)\left(\frac{\pi}{h}\right)z\right], \qquad n = 0, 1, 2, \ldots. \tag{7.4.49}$$

In terms of the $M_n(z)$, we construct the eigenfunction expansion of the Green's function $K(x, y, z; \zeta) = \sum_{n=0}^{\infty} N_n(x, y) M_n(z)$. The $N_n(x, y)$ are determined by multiplying (7.4.44) by $M_n(z)$ and integrating from $-h$ to 0. Using the procedures of Section 4.6, we obtain, for $n = 0, 1, 2, \ldots$,

$$\nabla^2 N_n(x, y) + \left[k^2 - \left(n+\frac{1}{2}\right)^2\left(\frac{\pi}{h}\right)^2\right] N_n(x, y) = -M_n(\zeta)\delta(x)\delta(y). \tag{7.4.50}$$

To obtain a unique solution of (7.4.50), we must specify the behavior of $N_n(x, y)$ as $x, y \to \infty$. Disregarding the multiplicative constant $M_n(\zeta)$, we are essentially interested in constructing a free space Green's function for (7.4.50). If $k^2 > [n + (1/2)]^2(\pi/h)^2$, (7.4.50) is an inhomogeneous Helmholtz equation, while if $k^2 < [n + (1/2)]^2(\pi/h)^2$, it is a modified Helmholtz equation.

With $r = \sqrt{x^2 + y^2}$, a function $\phi(x, y)$ is said to satisfy the *Sommerfeld radiation condition* at infinity if

$$\lim_{r\to\infty} \sqrt{r}\left[\frac{\partial\phi(x, y)}{\partial r} - i\omega\phi(x, y)\right] = 0, \tag{7.4.51}$$

and $\phi(x, y)$ satisfies the equation

$$\phi_{xx}(x, y) + \phi_{yy}(x, y) + \omega^2\phi(x, y) = 0 \tag{7.4.52}$$

as $r \to \infty$. With $\omega_n^2 = -\lambda_n$ and λ_n defined in (7.4.47), we write (7.4.50) as

$$\nabla^2 N_n(x,y) + \omega_n^2 N_n(x,y) = -M_n(\zeta)\delta(x)\delta(y), \qquad n = 0,1,2,\ldots. \quad (7.4.53)$$

If $\omega_n^2 > 0$, we use Example 6.13 to conclude, in view of (6.7.35), that

$$N_n(x,y) = \frac{i}{4} M_n(\zeta) H_0^{(1)}(\omega_n r), \qquad n = 0,1,\ldots, \quad (7.4.54)$$

where $H_0^{(1)}(z)$ is the zero-order Hankel function of the first kind. With the use of (6.7.38) it is easy to see that (7.4.54) satisfies the radiation condition (7.4.51) if $\omega_n^2 > 0$. If $\omega_n^2 < 0$ (i.e., ω_n is imaginary), the appropriate solution of (7.4.53) must be given in terms of the modified Bessel function $K_0(z)$ as in (6.7.32). However, since $H_0^{(1)}(iz) = (2/\pi i)K_0(z)$ for $z > 0$ we conclude that (7.4.54) is a suitable solution for all values of n. Since $N_n(x,y)$ decays exponentially at infinity when $\omega_n^2 < 0$, in view of (6.7.34), the radiation condition is certainly satisfied.

Collecting our results, we obtain the Green's function $K = K(x,y,z;\zeta)$ in the form

$$K = \frac{i}{2h} \sum_{n=0}^{\infty} \sin\left[\left(n+\frac{1}{2}\right)\frac{\pi}{h}\zeta\right] \sin\left[\left(n+\frac{1}{2}\right)\frac{\pi}{h}z\right] H_0^{(1)}(\omega_n\sqrt{x^2+y^2}), \quad (7.4.55)$$

where $\omega_n^2 = k^2 - [n+(1/2)]^2(\pi/h)^2$. As n increases, the argument of the Hankel function eventually becomes imaginary, so that the terms in the series decay exponentially for large n.

Exercises 7.4

7.4.1. Use the two-dimensional Fourier transform to obtain the Green's function (7.4.7).

7.4.2. Use the three-dimensional Fourier transform to obtain the Green's function (7.4.8).

7.4.3. Apply the Fourier sine transform to obtain the Green's function for the one-dimensional heat equation $u_t(x,t) - c^2 u_{xx}(x,t) = 0$ in the semi-infinite interval $0 < x < \infty$ with $u(x,t)$ prescribed at $x = 0$ and $t = 0$.

7.4.4. Apply the Fourier cosine transform to obtain the Green's function for $u_t(x,t) - c^2 u_{xx}(x,t) = 0$ in $0 < x < \infty$, $t > 0$ with $u(x,0)$ and $u_x(0,t)$ specified.

7.4.5. Use the general solution of the Cauchy problem for the wave equation in one dimension to obtain the Green's function $K(x,t;\xi,\tau)$ as given in (7.4.28).

7.4.6. Apply the formula (7.1.34) to obtain the solution of the Cauchy problem for the wave equation $u_{tt}(\mathbf{x},t) = c^2\nabla^2 u(\mathbf{x},t)$ in three dimensions if the initial data are $u(\mathbf{x},0) = 0$ and $u_t(\mathbf{x},0) = f(\mathbf{x})$. The solution has the form given in (5.4.12).

7.4.7. Obtain the result (5.4.23) for the Cauchy problem for the wave equation $u_{tt}(\mathbf{x},t) = c^2\nabla^2 u(\mathbf{x},t)$ in two dimensions if $u(\mathbf{x},0) = 0$ and $u_t(\mathbf{x},0) = f(\mathbf{x})$ using the Green's function (7.4.26) and the formula (7.1.34).

7.4.8. Use the Green's functions for the wave equation in one, two, and three dimensions given in (7.4.29), (7.4.26), and (7.4.21) to determine domains of dependence for the solutions of the Cauchy problem for the inhomogeneous wave equation in each of the three cases.

7.4.9. The Green's function for the Klein-Gordon equation in three dimensions, $u_{tt} - \gamma^2\nabla^2 u + c^2 u = 0$, is $K(x,y,z,t;\xi,\eta,\zeta,\tau) = \delta[\gamma(\tau-t)-r]/4\pi\gamma r - \left(c/4\pi\gamma^2\right) J_1\left[(c/\gamma)\sqrt{\gamma^2(\tau-t)^2-r^2}\right]/\sqrt{\gamma^2(\tau-t)^2-r^2}\,H[\gamma(\tau-t)-r]$, where $r^2 = (x-\xi)^2+(y-\eta)^2+(z-\zeta)^2$, as can be shown by using the Fourier transform. Let $\hat{r}^2 = \gamma^2(\tau-t)^2 - r^2$ in the Klein-Gordon equation and obtain the equation [see (6.7.42)] $\hat{u}''(\hat{r}) + (3/\hat{r})\hat{u}'(\hat{r}) + (c/\gamma)^2\hat{u}(\hat{r}) = 0$. Show that this equation has the nonsingular solution $\hat{u}_1(\hat{r}) = (1/\hat{r})J_1[c\hat{r}/\gamma]$, where J_1 is the Bessel function of order one. Argue that the preceding expression for the Green's function $K(x,y,z,t;\xi,\eta,\zeta,\tau)$ is reasonable in view of the fact that it reduces to the Green's function for the wave equation when $c = 0$ and has the same singularity on the characteristic cone.

7.4.10. Use the discussion of Exercise 7.4.9 to obtain the Green's function $K(x,y,z,t;\xi,\eta,\zeta,\tau)$ appropriate for the hyperbolic equation $u_{tt}(x,y,z,t) - \gamma^2\nabla^2 u(x,y,z,t) - c^2 u(x,y,z,t) = 0$ in three dimensions. *Hint*: It can be obtained by replacing c by ic in the Green's function for Exercise 7.4.9.

7.4.11. Obtain the Green's function $K(x,y,t;\xi,\eta,\tau)$ for the two-dimensional Klein-Gordon equation $u_{tt}(x,y,z,t) - \gamma^2\nabla^2 u(x,y,z,t) + c^2 u(x,y,z,t) = 0$ as $K(x,y,t;\xi,\eta,\tau) = (1/2\pi\gamma)\cos\left[(c/\gamma)\sqrt{\gamma^2(\tau-t)^2-\rho^2}\right]/\sqrt{\gamma^2(\tau-t)^2-\rho^2}$ $H[\gamma(\tau-t)-\rho]$, where $\rho^2 = (x-\xi)^2+(y-\eta)^2$, by following the procedure given in Exercise 7.4.9.

7.4.12. Obtain the Green's function $K(x,y,t;\xi,\eta,\tau)$ for the two-dimensional hyperbolic equation $u_{tt}(x,y,z,t) - \gamma^2\nabla^2 u(x,y,z,t) - c^2 u(x,y,z,t) = 0$ as $K(x,y,t;\xi,\eta,\tau) = (1/2\pi\gamma)\cosh\left[(c/\gamma)\sqrt{\gamma^2(\tau-t)^2-\rho^2}\right]/\sqrt{\gamma^2(\tau-t)^2-\rho^2}H[\gamma(\tau-t)-\rho]$, where $\rho^2 = (x-\xi)^2+(y-\eta)^2$, by following the method of Exercise 7.4.9.

7.4.13. Obtain the (free space) Green's function for the telegrapher's equation (1.2.29). *Hint*: Eliminate the first derivative term in the equation for the Green's function.

7.4.14. Obtain the Green's function for the Cauchy problem for Schrödinger's equation $i\hbar\,\partial u(x,t)/\partial t + \hbar^2/2m\,\partial^2 u(x,t)/\partial x^2 = 0$, where $\hbar$ is Planck's constant, m is the mass, and $i = \sqrt{-1}$. *Hint*: Use the Fourier transform.

7.4.15. Show that the solution of $\partial^2 K(x;\xi)/\partial x^2 - k^2 K(x;\xi) = -\delta(x-\xi)$, $-\infty < x,\xi < \infty$, $K(x;\xi) \to 0$ as $|x| \to \infty$, is given as $K(x;\xi) = (1/2k)\exp(-k|x-\xi|)$ (see Examples 5.1 and 6.13).

7.4.16. Obtain the solution $K(x;\xi)$ of the problem $\partial^2 K(x;\xi)/\partial x^2 + k^2 K(x;\xi) = -\delta(x-\xi)$, $-\infty < x,\xi < \infty$, $\lim_{|x|\to\infty}[\partial K(x;\xi)/\partial|x| - ikK(x;\xi)] = 0$ [i.e., $K(x;\xi)$ satisfies the radiation condition at infinity] in the form $K(x;\xi) = (i/2k)\exp(ik|x-\xi|)$ (see Example 6.13).

7.4.17. Solve the following problems for $K(x,\xi)$: $\partial^2 K(x;\xi)/\partial x^2 - k^2 K(x;\xi) = -\delta(x-\xi), 0 < x,\xi < \infty$. (a) $K(0;\xi) = 0$, $\lim_{x\to\infty} K(x;\xi) = 0$; (b) $\partial K(0;\xi)/\partial x = 0$, $\lim_{x\to\infty} K(x;\xi) = 0$.

7.4.18. Obtain the Green's function for the following problem: $\partial^2 K(x;\xi)/\partial x^2 + k^2 K(x;\xi) = -\delta(x-\xi)$, $0 < x,\xi < \infty$, with the boundary conditions $K(0;\xi) = 0$, $\lim_{x\to\infty}[\partial K(x;\xi)/\partial x - ikK(x;\xi)] = 0$.

7.4.19. Determine the Green's function $K(x,y;\xi,\eta)$ that satisfies the following problem:

$$\frac{\partial^2 K(x,y;\xi,\eta)}{\partial x^2} + \frac{\partial^2 K(x,y;\xi,\eta)}{\partial y^2} = -\delta(x-\xi)\delta(y-\eta), \quad \begin{cases} 0 < x,\xi < l, \\ -\infty < y,\eta < \infty, \end{cases}$$

$$K(x,y;\xi,\eta) = 0 \text{ at } x = 0 \text{ and } x = l, \quad K(x,y;\xi,\eta) \text{ is bounded as } |y| \to \infty.$$

Hint: Use the eigenfunctions for the interval $0 < x < l$ with zero boundary conditions.

7.4.20. Show that the following problem has no solution.

$$\frac{\partial^2 K(x,y;\xi,\eta)}{\partial x^2} + \frac{\partial^2 K(x,y;\xi,\eta)}{\partial y^2} = -\delta(x-\xi)\delta(y-\eta), \quad \begin{cases} 0 < x,\xi < l, \\ -\infty < y,\eta < \infty, \end{cases}$$

$$\frac{\partial K(x,y;\xi,\eta)}{\partial x} = 0 \text{ at } x = 0 \text{ and } x = l, \quad K(x,y;\xi,\eta) \to 0 \text{ as } |y| \to \infty.$$

Hint: Use the cosine eigenfunctions for $0 < x < l$.

7.4.21. Obtain the Green's function $K(x,y;\xi,\eta)$:

$$\nabla^2 K(x,y;\xi,\eta) - c^2 K(x,y;\xi,\eta) = -\delta(x-\xi)\delta(y-\eta), \quad \begin{cases} 0 < x,\xi < l, \\ -\infty < y,\eta < \infty, \end{cases}$$

$$K(x,y;\xi,\eta) = 0 \text{ at } x = 0 \text{ and } x = l, \quad K(x,y;\xi,\eta) \to 0 \text{ as } |y| \to \infty.$$

Hint: Use the eigenfunctions for the interval $0 < x < l$.

7.4.22. Set up a two-dimensional version of the ocean acoustics problem (i.e., drop the y dependence in the problem) and solve for the Green's function $K(x,z;\zeta)$.

7.4.23. Solve for the Green's function $K(x,y,z;\xi,\eta,\zeta)$: $\nabla^2 K(x,y,z;\xi,\eta,\zeta) - c^2 K(x,y,z;\xi,\eta,\zeta) = -\delta(x-\xi)\delta(y-\eta)\delta(z-\zeta), 0 < x,\xi < l, -\infty < y,\eta,z,\zeta < \infty$, with $K(x,y,z;\xi,\eta,\zeta) = 0$ at $x = 0$ and $x = l$, $K(x,y,z;\xi,\eta,\zeta) \to 0$ as $|y| \to \infty$, $|z| \to \infty$. *Hint*: Use the free space Green's function found in Example 6.13.

7.5 THE METHOD OF IMAGES

The *Green's function* K for any of the PDEs of Section 7.1 can be expressed in the form

$$K = K_F + K_G, \tag{7.5.1}$$

where K_F is the *free space Green's function*. We recall that the free space Green's function satisfies the same differential equation as the Green's function K. In addition, K_F satisfies (backward) causality conditions in the hyperbolic and parabolic cases, and appropriate conditions at infinity in the elliptic case. Consequently, K_G satisfies a homogeneous differential equation with homogeneous end conditions at $t = T$ if these are relevant. The boundary conditions for K_G are no longer homogeneous, however. For example, if K is required to vanish on ∂G, then $K_G = -K_F$ on ∂G. Although it may be possible to use *eigenfunction expansions* or *transform methods* to determine K_G, we do not use these approaches here since K itself can just as easily be determined in the same fashion, as we have seen in Sections 7.3 and 7.4. However, the results for K_G obtained by expansion or transform methods may be better suited for numerical evaluation than those found for K, since K_G is not singular within G.

In this section we construct Green's functions only for the equations of Section 7.1 that have constant coefficients with boundaries ∂G of a special form. We decompose the Green's function K as in (7.5.1) and use the *method of images* to determine K_G. We assume that the boundaries for the given problem are made up of (portions of) lines or planes, or (portions of) circles or spheres. For a prescribed singular point of the Dirac delta function in the equation for K, we consider all possible image points obtained by reflection through lines and planes and inversion through circles and spheres. (The inversion process is defined later.) If none of the resulting image points lies in the interior of the region in which the problem is specified and certain additional conditions are met, the Green's function K can be specified in a simple manner. We do not describe the most general regions and equations for which the method of images works. Instead, we consider a number of problems in the text and the exercises that exhibit the basic features of the method. Clearly, it is necessary to know the free space Green's functions K_F for the given equations in order to apply the method of images, and most of the relevant ones have already been determined.

Laplace's Equation in a Half-Space

We consider Laplace's equation in the half-space $z > 0$. The *Green's function* $K = K(x, y, z; \xi, \eta, \zeta)$ satisfies the equation

$$\nabla^2 K(x, y, z; \xi, \eta, \zeta) = -\delta(x - \xi)\delta(y - \eta)\delta(z - \zeta) \tag{7.5.2}$$

in the region G defined as the half-space $z > 0$. On the boundary ∂G [i.e., the (x, y)-plane $z = 0$ on which $\partial/\partial n = -\partial/\partial z$], we have

$$\alpha\, K(x, y, 0; \xi, \eta, \zeta) - \beta \frac{\partial K(x, y, 0; \xi, \eta, \zeta)}{\partial z} = 0. \tag{7.5.3}$$

We assume that α and β are constants and consider three cases. For the *Dirichlet problem* we set $\alpha = 1$ and $\beta = 0$; for the *Neumann problem* we put $\alpha = 0$ and $\beta = -1$; for the *Robin problem* we set $\alpha = h$ and $\beta = 1$ with $h > 0$. The Green's function $K(x, y, z; \xi, \eta, \zeta)$ is required to vanish at infinity.

We recall that the *free space Green's function* for Laplace's equation is

$$K_F(x, y, z; \xi, \eta, \zeta) = \frac{1}{4\pi} \left[(x - \xi)^2 + (y - \eta)^2 + (z - \zeta)^2\right]^{-1/2}, \qquad (7.5.4)$$

as was shown in Example 6.13. It follows from our discussion in that example that $\nabla^2 K_F(x, y, z; \xi, \eta, \zeta) = -\delta(x - \xi)\delta(y - \eta)\delta(z - \zeta)$. The Green's function $K(x, y, z; \xi, \eta, \zeta)$ is then expressed as $K(x, y, z; \xi, \eta, \zeta) = K_F(x, y, z; \xi, \eta, \zeta) + K_G(x, y, z; \xi, \eta, \zeta)$. As a result, $\nabla^2 K_F(x, y, z; \xi, \eta, \zeta) + \nabla^2 K_G(x, y, z; \xi, \eta, \zeta) = -\delta(x-\xi)\delta(y-\eta)\delta(z-\zeta)$ implies that $\nabla^2 K_G(x, y, z; \xi, \eta, \zeta) = 0$, so that $K_G(x, y, z; \xi, \eta, \zeta)$ is a *harmonic function* (i.e., a solution of Laplace's equation). In view of (7.5.3), the boundary condition for $K_G(x, y, z; \xi, \eta, \zeta)$ is

$$\alpha K_G(x, y, 0; \xi, \eta, \zeta) - \beta \frac{\partial K_G(x, y, 0; \xi, \eta, \zeta)}{\partial z}$$

$$= -\alpha K_F(x, y, 0; \xi, \eta, \zeta) + \beta \frac{\partial K_F(x, y, 0; \xi, \eta, \zeta)}{\partial z}. \qquad (7.5.5)$$

The point (ξ, η, ζ) is the source (or singular) point for the Green's functions $K(x, y, z; \xi, \eta, \zeta)$ and $K_F(x, y, z; \xi, \eta, \zeta)$. Now $K_F(x, y, z; \xi, \eta, \zeta)$ is given in terms of the distance from the observation point (x, y, z) to the source point (ξ, η, ζ). As shown in Figure 7.4, if we introduce the image (source) point $(\xi, \eta, -\zeta)$—that is, the reflection of (ξ, η, ζ) in the plane $z = 0$—then as the observation point (x, y, z) tends to a boundary point $(x, y, 0)$, its distance from (ξ, η, ζ) equals its distance from $(\xi, \eta, -\zeta)$.

Consequently, if we introduce the function

$$\hat{K}_G(x, y, z; \xi, \eta, \zeta) = \frac{1}{4\pi} \left[(x - \xi)^2 + (y - \eta)^2 + (z + \zeta)^2\right]^{-1/2}, \qquad (7.5.6)$$

we have in effect a free space Green's function that corresponds to the image source point $(\xi, \eta, -\zeta)$. At the boundary $z = 0$, $\hat{K}_G(x, y, 0; \xi, \eta, \zeta) = \frac{1}{4\pi}\left[(x - \xi)^2 + (y - \eta)^2 + \zeta^2\right]^{-1/2} = K_F(x, y, 0; \xi, \eta, \zeta)$, $\partial \hat{K}_G(x, y, 0; \xi, \eta, \zeta)/\partial z = -\zeta/(4\pi[(x - \xi)^2 + (y - \eta)^2 + \zeta^2]^{3/2}) = -\partial K_F(x, y, 0; \xi, \eta, \zeta)/\partial z$.

Furthermore, $\nabla^2 \hat{K}_G(x, y, z; \xi, \eta, \zeta) = -\delta(x - \xi)\delta(y - \eta)\delta(z + \zeta)$, and the right-hand side vanishes in the half-space $z > 0$ since $\delta(z + \zeta) = 0$ there. Therefore, if we set $K_G(x, y, z; \xi, \eta, \zeta) = -\hat{K}_G(x, y, z; \xi, \eta, \zeta)$ for the *Dirichlet problem* and $K_G(x, y, z; \xi, \eta, \zeta) = \hat{K}_G(x, y, z; \xi, \eta, \zeta)$ for the *Neumann problem*, the Green's function $K(x, y, z; \xi, \eta, \zeta) = K_F(x, y, z; \xi, \eta, \zeta) + K_G(x, y, z; \xi, \eta, \zeta)$ satisfies (7.5.2) in the half-space $z > 0$ as well as the boundary conditions $K(x, y, 0; \xi, \eta, \zeta) = 0$ and $\partial K(x, y, 0; \xi, \eta, \zeta)/\partial z = 0$, which are appropriate for the Dirichlet and Neumann problems, respectively.

The Green's function for the *Robin problem* cannot be obtained solely in terms of an image source point at $(\xi, \eta, -\zeta)$. Instead, we must introduce an entire line of

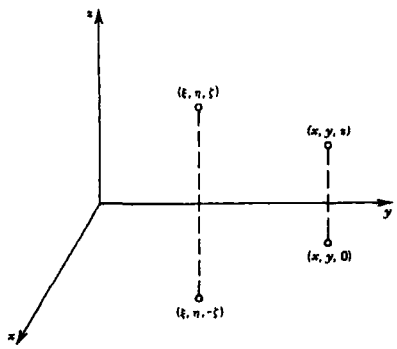

Figure 7.4 The source point (ξ, η, ζ) and its image point.

image sources on the line $x = \xi$ and $y = \eta$ with z extending from $z = -\zeta$ to $z = -\infty$ and a source density function to be determined. Let

$$K_G(x, y, z; \xi, \eta, \zeta) = \frac{1}{4\pi}\left[(x-\xi)^2 + (y-\eta)^2 + (z+\zeta)^2\right]^{-1/2} + \frac{1}{4\pi}\int_{-\infty}^{-\zeta} \frac{\rho(s)}{\left[(x-\xi)^2 + (y-\eta)^2 + (z-s)^2\right]^{1/2}}\, ds, \tag{7.5.7}$$

where $\rho(s)$ is the source density. When we put $h = 0$ in the boundary condition $\partial K(x, y, 0; \xi, \eta, \zeta)/\partial z - hK(x, y, 0; \xi, \eta, \zeta) = 0$, we expect (7.5.7) to reduce to the form appropriate for the Neumann problem, so we must have $\rho(s) = 0$ when $h = 0$. For this reason we have added the free space Green's function corresponding to an image source at $(\xi, \eta, -\zeta)$ to the integral term in (7.5.7).

It is assumed that $\rho(s)$ decays sufficiently rapidly at infinity that the integral in (7.5.7) converges and that differentiation under the integral sign is permitted. We then find that $\nabla^2 K_G(x, y, z; \xi, \eta, \zeta) = 0$ for $z > 0$, since all the singular points in (7.5.7) occur in the lower half-space $z < 0$. Applying the boundary condition at $z = 0$ gives, in view of (7.5.5),

$$\left.\frac{\partial K_G}{\partial z} - hK_G\right|_{z=0} = -\frac{\zeta}{4\pi}\left[(x-\xi)^2 + (y-\eta)^2 + \zeta^2\right]^{-3/2} - \frac{1}{4\pi}\int_{-\infty}^{-\zeta} \rho(s)\frac{\partial}{\partial s}\left[(x-\xi)^2 + (y-\eta)^2 + s^2\right]^{-1/2} ds$$

$$-\frac{h}{4\pi}\left[(x-\xi)^2 + (y-\eta)^2 + \zeta^2\right]^{-1/2} - \frac{h}{4\pi}\int_{-\infty}^{-\zeta} \frac{\rho(s)\, ds}{\left[(x-\xi)^2 + (y-\eta)^2 + s^2\right]^{1/2}}$$

$$= -\left.\frac{\partial K_F}{\partial z} + hK_F\right|_{z=0} = -\frac{\zeta}{4\pi}\left[(x-\xi)^2 + (y-\eta)^2 + \zeta^2\right]^{-3/2} + \frac{h}{4\pi}\left[(x-\xi)^2 + (y-\eta)^2 + \zeta^2\right]^{-1/2}. \tag{7.5.8}$$

The operator $\partial/\partial z$ has the same effect as $-\partial/\partial s$ at $z = 0$, and the use of $\partial/\partial s$ in the integral term enables us to integrate by parts. We have

$$\int_{-\infty}^{-\zeta} \rho(s) \frac{\partial}{\partial s} \left[(x-\xi)^2 + (y-\eta)^2 + s^2\right]^{-1/2} ds \tag{7.5.9}$$

$$= \rho(-\zeta) \left[(x-\xi)^2 + (y-\eta)^2 + \zeta^2\right]^{-1/2} - \int_{-\infty}^{-\zeta} \frac{\rho'(s)\, ds}{\left[(x-\xi)^2 + (y-\eta)^2 + s^2\right]^{1/2}}.$$

Combining results gives

$$\left[-\frac{h}{2\pi} - \frac{\rho(-\zeta)}{4\pi}\right] \left[(x-\xi)^2 + (y-\eta)^2 + \zeta^2\right]^{-1/2}$$

$$+ \frac{1}{4\pi} \int_{-\infty}^{-\zeta} \frac{\rho'(s) - h\rho(s)}{\left[(x-\xi)^2 + (y-\eta)^2 + s^2\right]^{1/2}} ds = 0. \tag{7.5.10}$$

Therefore, the boundary condition is satisfied if we set

$$\rho'(s) - h\rho(s) = 0, \; s < -\zeta, \quad \rho(-\zeta) = -2h. \tag{7.5.11}$$

The solution of the initial value problem (7.5.11) is $\rho(s) = -2he^{h(s+\zeta)}$. We note that $\rho(s)$ vanishes for $h = 0$ and that it decays exponentially as $s \to -\infty$.

The Green's function for the *third boundary value problem* thus has the form

$$K(x,y,z;\xi,\eta,\zeta) = \frac{1}{4\pi} \left[(x-\xi)^2 + (y-\eta)^2 + (z-\zeta)^2\right]^{-1/2}$$

$$+ \frac{1}{4\pi} \left[(x-\xi)^2 + (y-\eta)^2 + (z+\zeta)^2\right]^{-1/2}$$

$$- \frac{h}{2\pi} \int_{-\infty}^{-\zeta} \frac{e^{h(s+\zeta)}}{\left[(x-\xi)^2 + (y-\eta)^2 + (z-s)^2\right]^{1/2}} ds. \tag{7.5.12}$$

With $h = 0$ this reduces to Green's function for the *Neumann problem*. The Green's function for the *Dirichlet problem* is

$$K(x,y,z;\xi,\eta,\zeta) = \frac{1}{4\pi} \left[(x-\xi)^2 + (y-\eta)^2 + (z-\zeta)^2\right]^{-1/2}$$

$$- \frac{1}{4\pi} \left[(x-\xi)^2 + (y-\eta)^2 + (z+\zeta)^2\right]^{-1/2}. \tag{7.5.13}$$

On using the free space Green's function for Laplace's equation in two dimensions, it is easy to obtain the Green's function for the half-plane problem in two dimensions. Furthermore, on using the formulas given in Section 7.1, one can readily obtain the solutions of boundary value problems in the half-space or the half-plane for Laplace's equation, as shown in the exercises.

Hyperbolic Equations in a Semi-Infinite Interval

The free space Green's functions appropriate for the one-dimensional hyperbolic equation (7.1.45) with constant coefficients were obtained in Section 7.4. For the *Klein-Gordon equation*, the free space Green's function $K_F(x,t;\xi,\tau)$ satisfies (7.4.33) and is given as

$$K_F = \frac{1}{2\gamma} J_0\left[\frac{c}{\gamma}\sqrt{\gamma^2(t-\tau)^2-(x-\xi)^2}\right] H[x-\xi-\gamma(t-\tau)]H[\xi-x+\gamma(\tau-t)], \tag{7.5.14}$$

in view of (7.4.35). For the *modified telegrapher's equation* (7.4.36), we have

$$K_F = \frac{1}{2\gamma} I_0\left[\frac{\hat{c}}{\gamma}\sqrt{\gamma^2(t-\tau)^2-(x-\xi)^2}\right] H[x-\xi-\gamma(t-\tau)]H[\xi-x+\gamma(\tau-t)]. \tag{7.5.15}$$

The free space Green's function for the *wave equation* (7.4.27) is

$$K_F(x,t;\xi,\tau) = \frac{1}{2c} H[x-\xi-c(t-\tau)]H[\xi-x+c(\tau-t)]. \tag{7.5.16}$$

For the wave equation, $K_F(x,t;\xi,\tau)$ can also be expressed as

$$K_F(x,t;\xi,\tau) = \frac{1}{2c} H(\tau-t)H[c^2(t-\tau)^2-(x-\xi)^2]. \tag{7.5.17}$$

It is readily seen that (7.5.17) is consistent with (7.4.28) and that the product of the Heaviside functions given in (7.5.14)–(7.5.15), can also be expressed in the form (7.5.17) with appropriately modified constants. In connection with the application of the method of images, it is apparent from the form of $K_F(x,t;\xi,\tau)$ in (7.5.17) that an image source can be introduced at $x=-\xi$ if the Green's function in the interval $x>0$ is to be obtained. It can be shown by direct substitution that (7.5.17) is a solution of (7.4.27).

For each of the hyperbolic equations considered we now obtain *Green's functions* for the semi-infinite interval $x>0$. In the case of *Dirichlet boundary conditions* at $x=0$ [i.e., $K(0,t;\xi,\tau)=0$], the Green's function is given as

$$K(x,t;\xi,\tau) = K_F(x,t;\xi,\tau) - K_F(x,t;-\xi,\tau), \tag{7.5.18}$$

where $K_F(x,t;\xi,\tau)$ is the appropriate free space Green's function (7.5.14), (7.5.15), or (7.5.16). If a *Neumann boundary condition* is given at $x=0$ [i.e., $\partial K(0,t;\xi,\tau)/\partial x = 0$], the Green's function is

$$K(x,t;\xi,\tau) = K_F(x,t;\xi,\tau) + K_F(x,t;-\xi,\tau). \tag{7.5.19}$$

Finally, if a *Robin boundary condition* is assigned (i.e., $\partial K(0,t;\xi,\tau)/\partial x - hK(0,t;\xi,\tau) = 0$ with $h>0$), the Green's function is

$$K(x,t;\xi,\tau) = K_F(x,t;\xi,\tau)+K_F(x,t;-\xi,\tau)-2h\int_{-\infty}^{-\zeta} e^{h(s+\xi)}K_F(x,t;s,\tau)\,ds. \tag{7.5.20}$$

By comparison with the methods used in the preceding example or by direct verification it can be determined that (7.5.18)–(7.5.20) give the required Green's functions for each of the boundary value problems considered.

Heat Equation in a Finite Interval

We construct the Green's function $K(x,t;\xi,\tau)$ for the equation of heat conduction in a finite interval $0 < x < l$. Thus $K(x,t;\xi,\tau)$ satisfies the equation

$$\frac{\partial K(x,t;\xi,\tau)}{\partial t} + c^2\frac{\partial^2 K(x,t;\xi,\tau)}{\partial x^2} = -\delta(x-\xi)\delta(t-\tau),\ 0 < x,\xi < l,\ t,\tau < T. \tag{7.5.21}$$

It is assumed that $K(x,t;\xi,\tau)$ vanishes at the endpoints so that $K(0,t;\xi,\tau) = K(l,t;\xi,\tau) = 0$. In addition, we have the end condition $K(x,T;\xi,\tau) = 0$.

The free space Green's function $K_F(x,t;\xi,\tau)$ for the one-dimensional heat equation was found in Section 7.4 to be [see (7.4.6)]

$$K_F(x,t;\xi,\tau) = \frac{H(\tau-t)}{\sqrt{4\pi c^2(\tau-t)}}\exp\left[-\frac{(x-\xi)^2}{4c^2(\tau-t)}\right]. \tag{7.5.22}$$

We express $K(x,t;\xi,\tau)$ in the form (7.5.1) [i.e., $K(x,t;\xi,\tau) = K_F(x,t;\xi,\tau) + K_G(x,t;\xi,\tau)$] and use the method of images to specify $K_G(x,t;\xi,\tau)$. The source point $x = \xi$ must have an image with respect to $x = 0$ and $x = l$, and each of the image sources, in turn, must also have images with respect to $x = 0$ and $x = l$. Consequently, we are led to consider an infinite sequence of source points $\xi_n = \pm\xi \pm 2nl$, with $n = 0, 1, 2, 3, \ldots$. Some of these points are shown in Figure 7.5.

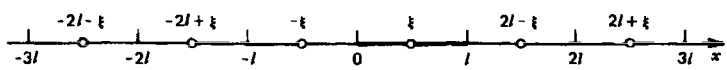

Figure 7.5 The source point and the image sources.

The Green's function $K = K(x,t;\xi,\tau)$ can then be written as an infinite series

$$K = \frac{H(\tau-t)}{\sqrt{4\pi c^2(\tau-t)}}\sum_{n=-\infty}^{\infty}\left[\exp\left[-\frac{(x-\xi-2nl)^2}{4c^2(\tau-t)}\right] - \exp\left[-\frac{(x+\xi-2nl)^2}{4c^2(\tau-t)}\right]\right]. \tag{7.5.23}$$

The term with $n = 0$ and a positive coefficient corresponds to $K_F(x,t;\xi,\tau)$. Clearly, $K(x,T;\xi,\tau) = 0$ since $\tau < T$. It can be shown that the series can be differentiated term by term. Inasmuch as each of the functions in the series except the term that corresponds to $K_F(x,t;\xi,\tau)$ has it source point ξ_n outside the interval $0 < x < l$, we see that (7.5.23) satisfies (7.5.2). Also, it is not difficult to see that at $x = 0$ and $x = l$ there corresponds to each term in the series with a positive coefficient an identical term with a negative coefficient. For example, Figure 7.5 shows that $x = \xi$ and

$x = -\xi$ are images with respect to $x = 0$ and that $x = \xi$ and $x = 2l - \xi$ are images with respect to $x = l$. The terms in the series (7.5.23) that correspond to the points $x = \xi$ and $x = -\xi$ are $\exp[-(x-\xi)^2/4c^2(\tau - t)] - \exp[-(x+\xi)^2/4c^2(\tau - t)]$, and this difference vanishes at $x = 0$. Also, for $x = \xi$ and $x = 2l - \xi$ we have $\exp[-(x-\xi)^2/4c^2(\tau - t)] - \exp[-(x+\xi-2l)^2/4c^2(\tau - t)]$, and this difference vanishes when $x = l$. In fact, if we replace n by $-n$ in the second series, it becomes identical to the first series at $x = 0$ and their difference equals zero. At $x = l$ we replace n by $-n + 1$ in the second series and find that the two series are equal.

The solution of the initial and boundary value problem for the heat equation in a finite interval when the formula (7.1.35) of Section 7.1 is used has a form similar to that given in Example 5.12. As was shown there, this result is expected to be useful for small values of t, whereas that given by the finite Fourier transform method is more useful for t large.

Green's Function for Laplace's Equation in a Sphere

As shown above, the method of images can be applied to equations with constant coefficients of all three types if there are linear or planar boundaries. For Laplace's equation it is possible to extend the image method to problems that involve circular or spherical boundaries as demonstrated below.

We construct the Green's function for the *Dirichlet problem* for *Laplace's equation* in the interior of a sphere using inversion with respect to the sphere. The Green's function $K = K(x, y, z; \xi, \eta, \zeta)$ satisfies the equation

$$\nabla^2 K(x, y, z; \xi, \eta, \zeta) = -\delta(x - \xi)\delta(y - \eta)\delta(z - \zeta) \tag{7.5.24}$$

and the boundary condition $K(x, y, z; \xi, \eta, \zeta) = 0$ on the sphere of radius a with center at the origin. Let the observation point P be denoted by $P = (x, y, z)$. The source point is $P_0 = (\xi, \eta, \zeta)$ and the origin of coordinates (i.e., the center of the sphere) is $O = (0, 0, 0)$.

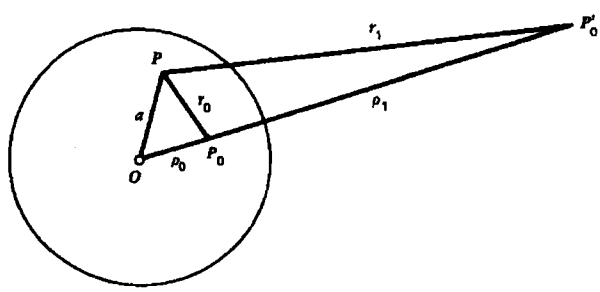

Figure 7.6 Inversion with respect to the sphere.

As shown in Figure 7.6, we introduce the image source point $P_0' = (\xi', \eta', \zeta')$. The point P_0' lies on the radial line extending from the origin O through the source

point P_0. Its distance from the origin equals ρ_1, and the distance of P_0 from the origin is ρ_0. These distances are related by $\rho_0\rho_1 = a^2$, where a is the radius of the sphere. The foregoing process of obtaining P_0' from P_0 is called *inversion with respect to the sphere*. We note that if $a = 1$, then $\rho_1 = 1/\rho_0$, so that the distances ρ_0 and ρ_1 are inverse to one another. The observation point P is assumed to lie on the sphere in Figure 7.6.

The triangles ΔOPP_0 and $\Delta OPP_0'$ in the figure are similar, since they have a common angle $\angle POP_0$ and proportional sides $\overline{OP_0}/\overline{OP} = \overline{OP}/\overline{OP_0'}$. This follows since $\overline{OP} = a$ (the radius of the sphere), $\overline{OP_0} = \rho_0$, $\overline{OP_0'} = \rho_1$, and $\overline{OP_0}\ \overline{OP_0'} = \rho_0\rho_1 = \overline{OP}^{\,2} = a^2$ in view of the above. The similarity of the triangles implies that all three sides are proportional and we have

$$\frac{\rho_0}{a} = \frac{a}{\rho_1} = \frac{r_0}{r_1}, \tag{7.5.25}$$

where $r_0 = \overline{PP_0}$ and $r_1 = \overline{PP_0'}$.

To complete the solution of the problem we set $K = K_F + K_G$, where K_F is the free space Green's function (7.5.4) and K_G is a constant multiple of the free space Green's function with source point at $P_0' = (\xi', \eta', \zeta')$. That is,

$$K(x, y, z; \xi, \eta, \zeta) = \frac{1}{4\pi}\left(\frac{1}{r_0} + \frac{c}{r_1}\right), \tag{7.5.26}$$

where $r_0^2 = (x-\xi)^2 + (y-\eta)^2 + (z-\zeta)^2$ and $r_1^2 = (x-\xi')^2 + (y-\eta')^2 + (z-\zeta')^2$, with the constant c to be specified. Since P_0' lies outside the sphere, the second term in (7.5.26) is a solution of Laplace's equation within the sphere. Thus (7.5.26) is a solution of (7.5.24). On the sphere [i.e., when $P = (x, y, z)$ lies on the sphere] we have, in the notation of Figure 7.6,

$$K(x, y, z; \xi, \eta, \zeta) = \frac{1}{4\pi}\left(\frac{1}{r_0} + \frac{c}{r_1}\right) = \frac{1}{4\pi r_0}\left(1 + \frac{\rho_0 c}{a}\right), \tag{7.5.27}$$

in view of (7.5.26). Thus if $c = -a/\rho_0$, the Green's function K vanishes on the sphere. Therefore, we obtain the Green's function

$$K(x, y, z; \xi, \eta, \zeta) = \frac{1}{4\pi}\left(\frac{1}{r_0} - \frac{a}{\rho_0 r_1}\right). \tag{7.5.28}$$

Let G represent the interior of the sphere $x^2 + y^2 + z^2 = a^2$ with ∂G representing the sphere itself. The Dirichlet boundary value problem for *Poisson's equation*,

$$\nabla^2 u(x, y, z) = -F(x, y, z), \qquad (x, y, z) \in G, \tag{7.5.29}$$

with the boundary condition

$$u(x, y, z)|_{\partial G} = B(x, y, z), \tag{7.5.30}$$

has the solution, with $K = K(x, y, z; \xi, \eta, \zeta)$,

$$u(\xi, \eta, \zeta) = \iiint_G KF(x, y, z)\, dv - \int_{\partial G} \frac{\partial K}{\partial n} B(x, y, z)\, ds, \tag{7.5.31}$$

where (ξ, η, ζ) as an arbitrary point within the sphere, as follows from (7.1.8).

To determine (7.5.31) more explicitly it is necessary to evaluate the exterior normal derivative $\partial K(x, y, z; \xi, \eta, \zeta)/\partial n$ on the surface of the sphere. Now

$$\frac{\partial K}{\partial n} = \frac{1}{4\pi}\left[\frac{\partial}{\partial r_0}\left(\frac{1}{r_0}\right)\frac{\partial r_0}{\partial n} - \frac{a}{\rho_0}\frac{\partial}{\partial r_1}\left(\frac{1}{r_1}\right)\frac{\partial r_1}{\partial n}\right] = \frac{1}{4\pi}\left[-\frac{1}{r_0^2}\frac{\partial r_0}{\partial n} + \frac{a}{\rho_0 r_1^2}\frac{\partial r_1}{\partial n}\right]. \tag{7.5.32}$$

We note that $r_0^2 = (x-\xi)^2 + (y-\eta)^2 + (z-\zeta)^2$ and that $r_1^2 = (x-\xi')^2 + (y-\eta')^2 + (z-\zeta')^2$, with (x, y, z) given as a point on the sphere in the present discussion. On introducing coordinate systems with the origins at (ξ, η, ζ) and (ξ', η', ζ'), we immediately conclude that $\partial r_0/\partial n = \cos(\theta_0)$, $\partial r_1/\partial n = \cos(\theta_1)$. The angle θ_0 is the angle between the exterior unit normal vector to the sphere at the point $P = (x, y, z)$ and the vector extending from the source point $P_0 = (\xi, \eta, \zeta)$ to P. Similarly, θ_1 is the angle between the normal vector at P and the vector from $P_0' = (\xi', \eta', \zeta')$ to P. Referring to Figure 7.6 and using the law of cosines, we obtain $\cos(\theta_0) = (a^2 + r_0^2 - \rho_0^2)/2ar_0$, $\cos(\theta_1) = (a^2 + r_1^2 - \rho_1^2)/2ar_1$. Using (7.5.25) to replace ρ_1 and r_1 by ρ_0 and r_0 in $\cos(\theta_1)$ gives $\cos(\theta_1) = (\rho_0^2 + r_0^2 - a^2)/2\rho_0 r_0$. Combining these results we conclude that $\partial K/\partial n|_{\partial G} = (\rho_0^2 - a^2)/4\pi a r_0^3$. Thus the solution formula (7.5.31) reduces to

$$u(\xi, \eta, \zeta) = \frac{1}{4\pi}\iiint_G \left(\frac{1}{r_0} - \frac{a}{\rho_0 r_1}\right) F\, dv + \frac{1}{4\pi a}\int_{\partial G} \frac{a^2 - \rho_0^2}{r_0^3} B\, ds, \tag{7.5.33}$$

with $F = F(x, y, z)$ and $B = B(x, y, z)$. By transforming to spherical coordinates with center at the origin, the second integral in (7.5.33) can be expressed as, with $\gamma = \angle POP_0$,

$$\frac{1}{4\pi a}\int_{\partial G} \frac{a^2 - \rho_0^2}{r_0^3} B\, ds = \frac{a}{4\pi}\int_0^{2\pi}\int_0^{\pi} \frac{(a^2 - \rho_0^2)B(a, \theta, \phi)\sin(\phi)}{[a^2 - 2a\rho_0\cos(\gamma) + \rho_0^2]^{3/2}}\, d\phi\, d\theta. \tag{7.5.34}$$

This expression is known as *Poisson's integral* for the sphere and can be obtained by using separation of variables for Laplace's equation in the sphere. It represents the solution of the Dirichlet problem for Laplace's equation, that is, (7.5.29)–(7.5.30) with $F(x, y, z) = 0$.

If we put $F(x, y, z) = 0$ in (7.5.33) and evaluate $u(\xi, \eta, \zeta)$ at the origin, we obtain

$$u(0, 0, 0) = \frac{1}{4\pi a^2}\int_{\partial G} B(x, y, z)\, ds, \tag{7.5.35}$$

since $\rho_0 = 0$ and $r_0 = a$ in that case. Now $4\pi a^2$ equals the surface area of the sphere and $B(x, y, z)$ is the value of $u(x, y, z)$ on the surface of the sphere. Thus $u(0, 0, 0)$ evaluated at the center of the sphere equals the average of its values on the surface of the sphere. This *mean value property* is valid for *harmonic functions* (i.e., solutions of Laplace's equation) in both two and three dimensions. It also arises in discrete formulations of Laplace's equation, as was seen in Section 1.3. This property can be derived in a more direct manner using Green's theorem as shown in the exercises.

The method of inversion that we have used to obtain Green's function for the Dirichlet problem does not work for the case of Neumann's problem or the third boundary value problem for the sphere. Similarly, it cannot be applied to other elliptic equations or, for that matter, to hyperbolic or parabolic equations. As an example, we consider the *reduced wave equation* in three dimensions,

$$\nabla^2 u(x, y, z) + k^2 u(x, y, z) = 0, \tag{7.5.36}$$

in a sphere and try to use inversion to find the Green's function for the Dirichlet problem. As shown in Example 6.13, the free space Green's function for this problem is

$$K_F(x, y, z; \xi, \eta, \zeta) = \frac{1}{4\pi r} e^{ikr}, \tag{7.5.37}$$

with $r^2 = (x-\xi)^2 + (y-\eta)^2 + (z-\zeta)^2$. This function satisfies the (three-dimensional) *radiation condition of Sommerfeld* $\lim_{r\to\infty} r\,[\partial K_F/\partial r - ikK_F] = 0$. The Green's function K is sought in the form

$$K = \frac{1}{4\pi}\left(\frac{1}{r_0} e^{ikr_0} + \frac{c}{r_1} e^{ikr_1}\right), \tag{7.5.38}$$

with r_0 and r_1 defined as above and the constant c to be determined. Clearly, c must have the same value given in the foregoing, (i.e., $c = -a/\rho_0$). However, on the sphere we have $r_1 = (a/\rho_0) r_0$, so that the exponentials have different arguments and no cancellation occurs. It is not possible to replace e^{ikr_1} by $e^{ik(\rho_0/a)r_1}$, in which case both exponentials would be equal on the sphere, since $(1/r_1)\exp[ik(\rho_0/a)r_1]$ is not a solution of the reduced wave equation.

There are other analytic techniques for obtaining *Green's functions* that we have not discussed. For example, *conformal mapping methods* are extremely useful in obtaining Green's functions for Laplace's equation in two dimensions. They require a knowledge of the theory of complex variables for their application. In addition, it is often possible to use known Green's functions to convert a differential equation to an *integral equation* and thereby construct other Green's functions. Again, it is rarely a simple matter to solve the integral equations that result, and most often this must be done approximately. Finally, there exist *perturbation* and *asymptotic methods* for finding certain Green's functions. An indication of how to use such methods is given in Chapters 9 and 10, which are devoted to approximation methods.

Exercises 7.5

7.5.1. Use the Green's functions obtained in the subsection on Laplace's equation in a half-space and appropriate forms of the solution formulas obtained in Section 7.1 to construct the solution of Laplace's equation $\nabla^2 u(x,y,z) = 0$ in the half-space $z > 0$ if the following boundary conditions are given: (a) $u(x,y,0) = f(x,y)$, $-\infty < x,y < \infty$; (b) $\partial u(x,y,0)/\partial z = f(x,y)$, $-\infty < x,y < \infty$; (c) $\partial u(x,y,0)/\partial z - hu(x,y,0) = f(x,y)$, $-\infty < x,y < \infty$.

7.5.2. Use the method of images to obtain the Green's function associated with the Dirichlet problem for $\nabla^2 u(x,y) = 0$ in the half-plane $y > 0$, in the form $K(x,y;\xi,\eta) = -(1/2\pi)\log(\rho/\hat{\rho})$, $0 < y,\eta < \infty$, $-\infty < x,\xi < \infty$, where $\rho^2 = (x-\xi)^2 + (y-\eta)^2$ and $\hat{\rho}^2 = (x-\xi)^2 + (y+\eta)^2$. Use this Green's function to obtain the solution (5.2.58) of the boundary value problem (5.2.50)–(5.2.51).

7.5.3. Apply the method of images to obtain the Green's function associated with the Neumann problem for $\nabla^2 u(x,y) = 0$, in the form $K(x,y;\xi,\eta) = -(1/2\pi)\log(\rho\hat{\rho})$, $0 < y,\eta < \infty$, $-\infty < x,\xi < \infty$, where ρ and $\hat{\rho}$ are defined as in Exercise 7.5.2. Use this Green's function to obtain the solution (5.2.66) of Laplace's equation with the boundary condition (5.2.60).

7.5.4. Obtain the Green's function that corresponds to (7.5.12) in the two-dimensional case.

7.5.5. Use the method of images to obtain the Green's function for Laplace's equation in a composite medium in three dimensions. That is, determine $K(x,y,z;\xi,\eta,\zeta)$ that satisfies $p(x,y,z)\nabla^2 K(x,y,z,\xi,\eta,\zeta) = -\delta(x-\xi)\delta(y-\eta)\delta(z-\zeta)$, $-\infty < x,\xi.y,\eta,z,\zeta < \infty$, with $z \neq 0$, the condition $K(x,y,z,\xi,\eta,\zeta) \to 0$ as $|z| \to \infty$, as well as the jump conditions at $z = 0$, $K(x,y,z,\xi,\eta,\zeta)$ and $p(x,y,z)\,\partial K(x,y,z,\xi,\eta,\zeta)/\partial z$ continuous at $z = 0$, where $p(x,y,z) = p_1$ for $z < 0$, $p(x,y,z) = p_2$ for $z > 0$, and p_1 and p_2 are constants. *Hint*: Let $K(x,y,z,\xi,\eta,\zeta) = K_1(x,y,z,\xi,\eta,\zeta)$ for $z < 0$ and $K(x,y,z,\xi,\eta,\zeta) = K_2(x,y,z,\xi,\eta,\zeta)$ for $z > 0$. Since the source point lies in $z > 0$ use the free space Green's function and its image for $z > 0$ and the free space Green's function with a source at (ξ,η,ζ) in $z < 0$. Then apply the matching conditions.

7.5.6. Use the approach that led to (7.5.23) to construct a Green's function for Laplace's equation $\nabla^2 u(x,y,z) = 0$ in the region $-\infty < x < \infty$, $-\infty < y < \infty$, $0 < z < \infty$ for the case of Dirichlet boundary conditions.

7.5.7. Apply the method of images to construct the Green's function for the Helmholtz equation $\nabla^2 u(x,y,z) + k^2 u(x,y,z) = 0$ in the half-space $z > 0$, in the case where $K(x,y,z;\xi,\eta,\zeta)$ satisfies one of the following boundary conditions: (a) $K(x,y,0;\xi,\eta,\zeta) = 0$; (b) $\partial K(x,y,0;\xi,\eta,\zeta)/\partial z = 0$; (c) $\partial K(x,y,0;\xi,\eta,\zeta)/\partial z - hK(x,y,0;\xi,\eta,\zeta) = 0$, with $-\infty < x,y,\xi,\eta < \infty$, and where $K(x,y,z;\xi,\eta,\zeta)$ satisfies the radiation condition at infinity in all cases. *Hint*: Use Example 6.13.

7.5.8. Construct the two-dimensional forms of the Green's functions obtained in Exercise 7.5.7 for the Helmholtz equation $\nabla^2 u(x,y) + k^2 u(x,y) = 0$ in the half-plane $y > 0$.

7.5.9. Construct the Green's function for $\nabla^2 u(x,y,z) - c^2 u(x,y,z) = 0$ in the half-space $z > 0$ if the Green's function $K(x,y,z;\xi,\eta,\zeta)$ satisfies the following boundary conditions: (a) $K(x,y,0;\xi,\eta,\zeta) = 0$; (b) $\partial K(x,y,0;\xi,\eta,\zeta)/\partial z = 0$ for all $-\infty < x, y < \infty$, and $K(x,y,z;\xi,\eta,\zeta) \to 0$ as $|z| \to \infty$. *Hint*: Use Example 6.13 and the method of images.

7.5.10. Carry out the solution of the two-dimensional version of Exercise 7.5.9 in the half-plane $y > 0$.

7.5.11. Use the Green's function (7.5.18) for the wave equation in the semi-infinite interval $0 < x < \infty$, to solve the following initial and boundary value problem: $u_{tt}(x,t) - c^2 u_{xx}(x,t) = 0$, $0 < x < \infty$, $t > 0$, $u(x,0) = u_t(x,0) = 0$, $0 < x < \infty$, $u(0,t) = g(t)$, $t > 0$.

7.5.12. Solve the following problem by use of the Green's function (7.5.19): $u_{tt}(x,t) - \gamma^2 u_{xx}(x,t) + c^2 u(x,t) = 0$, $x > 0$, $t > 0$, $u(x,0) = u_t(x,0) = 0$, $x > 0$, $\partial u(0,t)/\partial x = h(t)$, $t > 0$.

7.5.13. Verify that (7.5.19) is indeed the appropriate Green's function for the modified telegrapher's equation (7.4.36) in the case of Neumann boundary conditions.

7.5.14. Noting the result of Exercise 7.2.17 [i.e., $2\pi \sum_{k=-\infty}^{\infty} \delta[x - 2k\pi] = \sum_{k=-\infty}^{\infty} \exp[ikx]$], multiply both sides by a Fourier transformable function $\phi(x)$, and conclude upon integrating that $\sqrt{2\pi} \sum_{k=-\infty}^{\infty} \phi[2k\pi] = \sum_{k=-\infty}^{\infty} F[k]$, where $F[k]$ is the Fourier transform of $\phi(x)$. This result is known as *Poisson's summation formula.*

7.5.15. Use the Poisson summation formula (Exercise 7.5.14) to show the equivalence of the Green's function expression (7.5.23) for the heat equation and that obtained by the method of eigenfunction expansions.

7.5.16. Use the method of images to construct the Green's function $K(x,t;\xi,\tau)$ for the heat equation in a finite interval if (a) $\partial K(x,t;\xi,\tau)/\partial x = 0$ at $x = 0$ and $x = l$; (b) $\partial K(0,t;\xi,\tau)/\partial x = 0$, $\partial K(l,t;\xi,\tau)/\partial x = 0$.

7.5.17. Apply the method of images to construct a Green's function for the heat equation in the semi-infinite interval $0 < x < \infty$ with Dirichlet boundary conditions and use it to solve the following problem: $u_t(x,t) - c^2 u_{xx}(x,t) = 0$, $0 < x < \infty$, $t > 0$, $u(x,0) = f(x)$, $0 < x < \infty$, $u(0,t) = g(t)$, $t > 0$.

7.5.18. Adapt the inversion process used to determine the Green's function for Laplace's equation in a sphere to obtain the Green's function for the Dirichlet problem for Laplace's equation in a disk $x^2 + y^2 < a^2$.

7.5.19. Verify the result (7.5.36).

7.5.20. Obtain Poisson's integral for the circle [see (7.5.36)] by using the Green's function obtained in Exercise 7.5.18.

7.5.21. Use the method of images to determine the appropriate Green's function $K(x,y,\xi,\eta)$ for Laplace's equation $\nabla^2 u(x,y) = 0$ in the quadrant $0 < x < \infty$, $0 < y < \infty$ with the following boundary conditions: (a) $K(x,y,\xi,\eta) = 0$ on $x = 0$, $y > 0$ and $y = 0$, $x > 0$; (b) $K(x,y,\xi,\eta) = 0$ on $x = 0$, $y > 0$, $\partial K(x,y,\xi,\eta)/\partial y = 0$ on $y = 0$, $x > 0$; (c) $\partial K(x,y,\xi,\eta)/\partial x = 0$ on $x = 0$, $y > 0$, $\partial K(x,y,\xi,\eta)/\partial y = 0$ on $y = 0$, $x > 0$.

7.5.22. Apply the method of images of construct the Green's function for Laplace's equation within the hemisphere $x^2 + y^2 + z^2 < a^2$, $z \geq 0$, with Dirichlet boundary conditions. *Hint*: Use inversion in the sphere and reflection in the plane $z = 0$.

7.5.23. Obtain the Green's function for the Dirichlet problem for Laplace's equation within the quarter-circle $x^2 + y^2 < a^2$, $x > 0$, $y > 0$. *Hint*: Reflect in the x- and y-axis as in the quarter-plane problem and then invert each source and image source in the circle.

7.5.24. Let u be a solution of Laplace's equation $\nabla^2 u(\mathbf{x}) = 0$ in the region G, and let R be the interior of a circle or a sphere centered at P_0 and completely contained within G. (a) Use Green's theorem or the divergence theorem to show that $\int_{\partial R} \partial u(\mathbf{x})/\partial n \, ds = 0$. (b) Let $K(\mathbf{x}; \boldsymbol{\xi})$ be the free space Green's function for Laplace's equation with singular point at P_0. Apply Green's theorem in the region R to the harmonic function $u(\mathbf{x})$ and the Green's function $K(\mathbf{x}; \boldsymbol{\xi})$. Deduce the mean value theorems $u(P_0) = (1/4\pi a^2) \int_{\partial R} u(\mathbf{x}) \, ds$ and $u(P_0) = (1/2\pi a) \int_{\partial R} u(\mathbf{x}) \, ds$ in three and two dimensions, respectively, where a is the radius of the sphere and the circle with center at P_0.

7.6 MAPLE METHODS

Generalized Functions

Maple defines the Heaviside function as *Heaviside* and the Dirac delta function as *Dirac*. That is, $H(x-a) = Heaviside(x-a)$ and $\delta(x-a) = Dirac(x-a)$ in terms of our notation (which we use below). For the Heaviside function we have

$$H(x) = \begin{cases} 0, & x < 0, \\ undefined, & x = 0, \\ 1, & 0 < x. \end{cases} \tag{7.6.1}$$

The value of $H(0)$ is left undefined by Maple and can be assigned. The derivative of the Heaviside function is given as $dH(x)/dx = \delta(x)$. Maple evaluates integrals that contain Heaviside and Dirac delta functions. For example, $\int_{-\infty}^{\infty} f(x)\delta(x-a)\,dx = f(a)$ and $\int_{-\infty}^{\infty} f(x)\delta'(x-a)\,dx = -f'(a)$.

Ordinary differential equations that contain Dirac delta or Heaviside functions can be solved using the *dsolve* procedure. The solution of the boundary value problem

$$y''(x) = -\delta(x - 1/2), \quad 0 < x < 1, \qquad y(0) = y(1) = 0, \tag{7.6.2}$$

is found by Maple, on expressing the result in piecewise form, as

$$y(x) = \begin{cases} x/2, & x < 1/2, \\ -x/2 + 1/2, & 1/2 < x. \end{cases} \tag{7.6.3}$$

Maple carries out integral transforms of Heaviside and Dirac delta functions. The Fourier transform of $H(x)$ is given as $fourier(H(x), x, \lambda) = \pi\delta(\lambda) - i/\lambda$. The Laplace transform of $H(x-1)$ is found to be $laplace(H(x-1), x, \lambda) = e^{-\lambda}/\lambda$. The inverse Laplace transform of 1 is $invlaplace(1, \lambda, x) = \delta(x)$.

Green's Functions for ODEs

Maple can determine Green's functions for ordinary differential equations because of its ability to solve ODEs with Dirac delta function inhomogeneous terms, as we have shown above. As an example, we consider

$$x\frac{\partial^2 K(x;\xi)}{\partial x^2}+\frac{\partial K(x;\xi)}{\partial x}=-\delta(x-\xi), \qquad 0<x,\ \xi<1, \tag{7.6.4}$$

with the boundary conditions

$$K(0,\xi) \text{ bounded}, \qquad K(1;\xi)=0. \tag{7.6.5}$$

The point $x=0$ is a singular point for the ODE so we require that the Green's function be bounded there. On using *dsolve* to solve the problem we can only prescribe the condition at $x=1$. Maple's solution is

$$K(x;\xi)=-\log(x)+H(\xi-x)(\log(x)-\log(\xi))+c\log(x), \tag{7.6.6}$$

with c as an arbitrary constant. Since the solution must be bounded at $x=0$ we set $c=0$. This yields the Green's function

$$K(x;\xi)=\begin{cases}-\ln(\xi), & x\le\xi,\\ -\ln(x), & \xi<x.\end{cases} \tag{7.6.7}$$

We have constructed a procedure $GreensFuncODE$ that automates the construction of Green's functions of self-adjoint ODEs. It takes the form $GreensFuncODE([p(x),p'(x),-q(x)],K(x;\xi),\alpha_1K_x(x;\xi)-\beta_1K(x;\xi)=0,\alpha_2K_x(x;\xi)+\beta_2K(x;\xi)=0,\xi,x=0..l)$. The Green's function $K(x;\xi)$ satisfies (7.2.38) over the finite interval $0<x<l$ with the boundary conditions $\alpha_1K_x(0;\xi)-\beta_1K(0;\xi)=0$ and $\alpha_2K_x(l;\xi)+\beta_2K(l;\xi)=0$. Although we have entered a general boundary condition in the procedure, it is possible to replace the boundary conditions at $x=0$ or $x=l$ by a boundedness condition if the endpoints are singular points for the ODE.

For example, the Green's function problem (7.6.4)–(7.6.5) is solved by the procedure $GreensFuncODE([x,1,0],K(x;\xi),K(x;\xi)=bounded,K(x,\xi)=0,\xi,x=0..1)$ and yields the solution (7.6.7).

Modified Green's functions can also be determined by the procedure. If we invoke $GreensFuncODE([1,0,0],K(x;\xi),K_x(x;\xi)=0,K_x(x;\xi)=0,\xi,x=0..l)$, we obtain the modified Green's function determined from the problem

$$\frac{\partial^2 K(x;\xi)}{\partial x^2}=-\delta(x-\xi)+\frac{1}{l}, \qquad 0<x,\ \xi<l, \tag{7.6.8}$$

with the boundary conditions

$$K_x(0,\xi)=0, \qquad K_x(l;\xi)=0. \tag{7.6.9}$$

The Green's function is found to be

$$K(x;\xi)=\begin{cases}(x^2-\xi^2)/2l, & x<\xi,\\ (x^2-\xi^2-2lx+2l\xi)/2l, & \xi<x.\end{cases} \tag{7.6.10}$$

Adjoint Differential Operators

We have seen the importance of adjoint differential operators and boundary conditions for the construction of Green's functions for ordinary and partial differential equations. Let L^* be the (formal) adjoint operator of L. Then $wLu - uL^*w = \nabla \cdot M$, and an application of the divergence theorem yields an integral over the boundary of the given region. (In the case of an ordinary differential equation, the divergence operator is replaced by a first derivative operator.) The procedure *AdjOp* determines the adjoint differential operators L^* for second order differential operators L and the related operators M. The procedure takes the form $AdjOp(Coeff2D, Coeff1D, Coeffu, u, VarList)$, where u is the dependent variable and $VarList$ is a list of the independent variables. The first two arguments of the procedure are a list of coefficients of the second and first derivative terms, respectively, in the operator L, and the third argument is the coefficient of the undifferentiated term.

If we invoke $AdjOp([[p(x)]], [p'(x)], q(x), u, [x])$, we determine the adjoint of the ODE operator $Lu = p(x)u''(x) + p'(x)u'(x) + q(x)u(x)$. Since L is a self-adjoint operator, the output of the procedure gives the adjoint operator as $L^*w \equiv Lw$, and the operator M is given as $M(u, w) = p(x)[w(x)u'(x) - u(x)w'(x)]$.

The procedure $AdjOp([[\exp(x), x^2/2], [x^2/2, 0]], [0, y], -10, u, [x, y])$ finds the adjoint of the partial differential operator $Lu = e^x u_{xx} + x^2 u_{xy} + yu_y - 10u$, to be $L^*w = e^x w_{xx} + x^2 w_{xy} + 2e^x w_x + (2x - y)w_y + (e^x - 11)w$, We see that L is not self-adjoint. The operator M is given as $M(u, w) = [e^x w u_x - e^x uw - e^x uw_x + x^2 w u_y/2 - x^2 uw_y/2,\ x^2 w u_x/2 - xuw - x^2 uw_x/2 + yuw]$.

Exercises 7.6

7.6.1. Use Maple to solve Exercise 7.2.9.

7.6.2. Determine the Green's function of Example 7.5 using *dsolve*.

7.6.3. Solve the Green's function problem of Exercise 7.3.5 using *dsolve*.

7.6.4. Put $l = 1$ in the boundary value problem of Exercise 7.3.9 and obtain its solution by using *dsolve*.

7.6.5. Solve the ODE in Exercise 7.4.9 using *dsolve*.

7.6.6. Solve the Green's function problem (7.6.4)–(7.6.5) using (a) *dsolve*; (b) the procedure $GreensFuncODE$.

7.6.7. Solve the Green's function problem (7.6.8)–(7.6.9) using (a) *dsolve*; (b) the procedure $GreensFuncODE$.

7.6.8. Obtain the adjoint operator of the general linear second order ordinary differential operator with variable coefficients using the procedure *AdjOp*.

CHAPTER 8

VARIATIONAL AND OTHER METHODS

In this chapter we present a number of methods and results that apply to either problems studied previously or to a new class of problems. We begin with a presentation of a *variational characterization* of the *eigenvalue problems* introduced in Chapter 4 and show how the variational approach enables us to prove some of the properties of eigenvalues and eigenfunctions given in that chapter. An important consequence of this approach is the *Rayleigh-Ritz method,* which yields approximate determinations of eigenvalues and eigenfunctions and this method is considered next.

We continue with a discussion of the classical *method of Riemann* for integrating linear hyperbolic equations of second order. This is followed by a presentation of *maximum and minimum principles* for the diffusion and Laplace's equations, and their consequences for *uniqueness* and *continuous dependence on data* for problems for these equations. In addition, we consider a positivity property for the telegrapher's equation.

The chapter concludes with a discussion of some basic *equations of mathematical physics*. First we consider the fourth order partial differential equations that govern the *vibration of rods and plates*. Next we examine the *equations of fluid dynamics, Maxwell's equations of electromagnetic theory*, and the *equations of elasticity theory*,

Partial Differential Equations of Appplied Mathematics, Third Edition. By Erich Zauderer

each of which is a system of partial differential equations. Techniques for simplifying and solving these equations are considered.

8.1 VARIATIONAL PROPERTIES OF EIGENVALUES AND EIGENFUNCTIONS

It has been demonstrated in the preceding chapters that eigenvalue problems play an important role in various methods for solving problems for PDEs. Previously we have determined the eigenvalues and eigenfunctions by using separation of variables for PDEs and finding general solutions of ODEs for the Sturm-Liouville problem. In this section we show how the *eigenvalues* and *eigenfunctions* can be specified by means of a *variational principle*. Thereby, it is possible to prove some of the properties of the eigenvalues and eigenfunctions given at the beginning of Chapter 4. Furthermore, the variational approach leads to a useful approximate method for determining the first few eigenvalues and eigenfunctions. This method, known as the *Rayleigh-Ritz method*, is discussed in Section 8.2.

Energy Integrals and Rayleigh Quotients

To motivate the use of the variational principle, we begin by considering the (self-adjoint) eigenvalue problem

$$LM(\mathbf{x}) = -\nabla \cdot (p(\mathbf{x})\nabla M(\mathbf{x})) + q(\mathbf{x})M(\mathbf{x}) = \lambda\rho(\mathbf{x})M(\mathbf{x}) \tag{8.1.1}$$

for the function $M(\mathbf{x})$ defined in the bounded region G. As in Sections 4.1 and 4.2, we assume that $p(\mathbf{x}) > 0$ and $\rho(\mathbf{x}) > 0$ in G, while $q(\mathbf{x}) \geq 0$ in G. The eigenfunction $M(\mathbf{x})$ is assumed to satisfy the boundary condition on ∂G,

$$\alpha(\mathbf{x})\, M(\mathbf{x}) + \beta(\mathbf{x})\, \frac{\partial M(\mathbf{x})}{\partial n}\bigg|_{\partial G} = 0, \tag{8.1.2}$$

where $\alpha(\mathbf{x}) \geq 0$ and $\beta(\mathbf{x}) \geq 0$, while $\alpha(\mathbf{x}) + \beta(\mathbf{x}) > 0$ on ∂G. The boundary conditions may be of the first, second, third,, or mixed kind, as described in Section 4.1. (We use notation appropriate for the two- or three-dimensional eigenvalue problem and occasionally indicate the appropriate forms for the one-dimensional Sturm-Liouville problem.)

As discussed in Section 4.2, we assume that there are a countable infinity of eigenvalues λ_k $(k = 1, 2, \ldots)$ that are real valued and nonnegative. They are numbered according to nondecreasing values as

$$0 \leq \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \cdots \leq \lambda_k \leq \cdots. \tag{8.1.3}$$

The associated eigenfunctions $M_k(\mathbf{x})$ are assumed to form a complete orthonormal set. That is,

$$(M_k(\mathbf{x}),\ M_j(\mathbf{x})) = \iint_G \rho(\mathbf{x})M_k(\mathbf{x})M_j(\mathbf{x})\, dv = \delta_{kj}, \tag{8.1.4}$$

where the Kronecker delta $\delta_{kj} = 1$ for $k = j$ and $\delta_{kj} = 0$ for $k \neq j$,with $k,\ j = 1, 2, 3, \ldots$.

Furthermore, any function $u(\mathbf{x})$ that is sufficiently smooth and satisfies the boundary condition (8.1.2)—where $M(\mathbf{x})$ is replaced by $u(\mathbf{x})$—can be expanded in a series of eigenfunctions

$$u(\mathbf{x}) = \sum_{k=1}^{\infty} N_k M_k(\mathbf{x}). \tag{8.1.5}$$

The Fourier coefficients N_k are given as

$$N_k = (u(\mathbf{x}),\ M_k(\mathbf{x})) = \iint_G \rho(\mathbf{x}) u(\mathbf{x}) M_k(\mathbf{x})\, dv. \tag{8.1.6}$$

We assume that $u(\mathbf{x})$ is normalized to unity, so that

$$||u(\mathbf{x})||^2 = (u(\mathbf{x}), u(\mathbf{x})) = \iint_G \rho(\mathbf{x}) u^2(\mathbf{x})\, dv = 1. \tag{8.1.7}$$

To proceed, we consider the integral

$$\iint_G uLu\, dv = \iint_G [p(\mathbf{x})(\nabla u(\mathbf{x}))^2 + q(\mathbf{x})u^2(\mathbf{x})]\, dv - \int_{\partial G} p(\mathbf{x})u(\mathbf{x}) \frac{\partial u(\mathbf{x})}{\partial n}\, ds, \tag{8.1.8}$$

where (4.2.24) has been used. Here $u(\mathbf{x})$ satisfies the boundary condition

$$\alpha(\mathbf{x})\, u(\mathbf{x}) + \beta(\mathbf{x}) \left.\frac{\partial u(\mathbf{x})}{\partial n}\right|_{\partial G} = 0. \tag{8.1.9}$$

As a result, if we note the formulation of the boundary conditions in Section 4.1, the right side of (8.1.8) can be expressed in the following form, which we denote by $E(u(\mathbf{x}))$:

$$E(u(\mathbf{x})) = \iint_G [p(\mathbf{x})(\nabla u(\mathbf{x}))^2 + q(\mathbf{x})u^2(\mathbf{x})]\, dv + \int_{S_3} \frac{\alpha(\mathbf{x})}{\beta(\mathbf{x})} p(\mathbf{x})u^2(\mathbf{x})\, ds. \tag{8.1.10}$$

The expression $E(u(\mathbf{x}))$ represents an *energy integral* on the basis of our discussion in Section 6.8, and $E(u(\mathbf{x}))$ is a nonlinear functional of $u(\mathbf{x})$. It is nonnegative in view of the assumptions on $p(\mathbf{x})$, $\rho(\mathbf{x})$, $q(\mathbf{x})$, $\alpha(\mathbf{x})$, and $\beta(\mathbf{x})$.

On inserting (8.1.5) into (8.1.8), we have

$$\begin{aligned} E(u(\mathbf{x})) &= \iint_G \left[\sum_{k=1}^{\infty} N_k M_k(\mathbf{x}) Lu(\mathbf{x})\right] dv = \sum_{k=1}^{\infty} \left[N_k \iint_G M_k(\mathbf{x}) Lu(\mathbf{x})\, dv\right] \\ &= \sum_{k=1}^{\infty} N_k \iint_G \left[\sum_{j=1}^{\infty} N_j \{\rho(\mathbf{x}) M_k(\mathbf{x}) M_j(\mathbf{x})\}\right] dv \\ &= \sum_{k=1}^{\infty} N_k \left\{\sum_{j=1}^{\infty} \lambda_j N_j (M_k(\mathbf{x}), M_j(\mathbf{x}))\right\} = \sum_{k=1}^{\infty} \lambda_k N_k^2, \end{aligned} \tag{8.1.11}$$

since $LM_j(\mathbf{x}) = \lambda_j \rho M_j(\mathbf{x})$ and $(M_k(\mathbf{x}),\ M_j(\mathbf{x})) = \delta_{kj}$. Also,

$$1 = (u(\mathbf{x}), u(\mathbf{x})) = \sum_{k=1}^{\infty} N_k \iint_G \left[\rho(\mathbf{x}) M_k(\mathbf{x}) \sum_{j=1}^{\infty} N_j M_j(\mathbf{x}) \right] dv$$

$$= \sum_{k=1}^{\infty} N_k \left\{ \sum_{j=1}^{\infty} N_j (M_k(\mathbf{x}), M_j(\mathbf{x})) \right\} = \sum_{k=1}^{\infty} N_k^2, \qquad (8.1.12)$$

since $(M_k(\mathbf{x}),\ M_j(\mathbf{x})) = \delta_{kj}$. We have assumed that the interchanges of summation, differentiation, and integration used in (8.1.11)–(8.1.12) are valid.

Noting (8.1.3) and (8.1.12), we obtain from (8.1.11)

$$E(u(\mathbf{x})) = \sum_{k=1}^{\infty} \lambda_k N_k^2 \geq \sum_{k=1}^{\infty} \lambda_1 N_k^2 = \lambda_1 \sum_{k=1}^{\infty} N_k^2 = \lambda_1. \qquad (8.1.13)$$

Thus for all *admissible functions* $u(\mathbf{x})$ that satisfy the boundary condition (8.1.9) and have unit norm $||u(\mathbf{x})|| = 1$, we have $E(u(\mathbf{x})) \geq \lambda_1$.

Now if $u(\mathbf{x})$ satisfies the additional *constraints*

$$(u(\mathbf{x}), M_k(\mathbf{x})) = \iint_G \rho(\mathbf{x}) u(\mathbf{x}) M_k(\mathbf{x})\, dv = 0, \quad k = 1, 2, \ldots, n-1, \qquad (8.1.14)$$

the first $n-1$ coefficients in the series (8.1.5) (i.e., $N_1, N_2, \ldots, N_{n-1}$) all vanish, in view of (8.1.6). Then (8.1.11) yields

$$E(u(\mathbf{x})) = \sum_{k=n}^{\infty} \lambda_k N_k^2 \geq \lambda_n \sum_{k=n}^{\infty} N_k^2 = \lambda_n \qquad (8.1.15)$$

on using (8.1.3) and (8.1.12). In addition, for all n,

$$E(M_n(\mathbf{x})) = \iint_G M_n(\mathbf{x}) L M_n(\mathbf{x})\, dv = \lambda_n \iint_G \rho(\mathbf{x}) M_n(\mathbf{x})^2\, dv = \lambda_n, \qquad (8.1.16)$$

since the $M_n(\mathbf{x})$ are normalized by assumption.

We have shown the following. The *energy integral* $E(u(\mathbf{x}))$ is greater than or equal to the smallest eigenvalue λ_1 associated with the eigenvalue problem (8.1.1)–(8.1.2), for all admissible $u(\mathbf{x})$. The minimum value λ_1 is assumed if $u(\mathbf{x}) = M_1(\mathbf{x})$, the first eigenfunction. If $u(\mathbf{x})$ satisfies the additional constraints (8.1.14), we have $E(u(\mathbf{x})) \geq \lambda_n$, and the minimum is assumed if $u(\mathbf{x}) = M_n(\mathbf{x})$. Consequently, we may characterize the eigenvalues λ_n and the eigenfunctions $M_n(\mathbf{x})$ in terms of the following *minimum problem.*

The nth *eigenvalue* λ_n *of the problem* (8.1.1)–(8.1.2) *is the minimum value of the energy integral* $E(u(\mathbf{x}))$ *over the set of admissible functions* $u(\mathbf{x})$ *that satisfy the conditions* (8.1.7), (8.1.9), *and* (8.1.14). [*If* $n = 1$, *the conditions* (8.1.14) *are absent.*] *Further, among all admissible functions* $u(\mathbf{x})$, *the minimum value of* $E(u(\mathbf{x}))$ (*i.e,* λ_n) *is assumed if* $u(\mathbf{x})$ *is the normalized eigenfunction* $M_n(\mathbf{x})$ *of* (8.1.1)–(8.1.2).

In a slight variation of this process, we drop the restriction that the admissible functions $u(\mathbf{x})$ are normalized to unity. Then, if $w(\mathbf{x})$ is a function that satisfies the boundary condition (8.1.9) and the constraints (8.1.14) [where $u(\mathbf{x})$ is replaced by $w(\mathbf{x})$], we can set $u(\mathbf{x}) = w(\mathbf{x})/||w(\mathbf{x})||$. The function $u(\mathbf{x})$ satisfies all the conditions stated in the foregoing minimum problem. Its energy integral takes the form

$$E(u(\mathbf{x})) = E\left(\frac{w(\mathbf{x})}{||w(\mathbf{x})||}\right) = \frac{E(w(\mathbf{x}))}{||w(\mathbf{x})||^2} = \frac{E(w(\mathbf{x}))}{(w(\mathbf{x}), w(\mathbf{x}))}. \tag{8.1.17}$$

With $E(w(\mathbf{x}))$ defined as in (8.1.10), the ratio $E(w(\mathbf{x}))/(w(\mathbf{x}), w(\mathbf{x}))$ is known as the *Rayleigh quotient*. Again the minimum values of the Rayleigh quotient determine the eigenvalues, and these minima are assumed by unnormalized eigenfunctions, determined only up to an arbitrary multiplicative constant.

Example 8.1. The Dirichlet Eigenvalue Problem for the Laplacian in a Square. We consider the eigenvalue problem for the *Laplacian operator*,

$$\nabla^2 M(x, y) + \lambda M(x, y) = M_{xx}(x, y) + M_{yy}(x, y) + \lambda M(x, y) = 0 \tag{8.1.18}$$

in the square $0 < x < \pi$ and $0 < y < \pi$, with the *Dirichlet boundary condition*

$$M(x, y)|_{\partial G} = 0. \tag{8.1.19}$$

The region G is the interior of the square, and ∂G is given by the sides of the square. The eigenvalues and eigenfunctions for the problem (8.1.18)–(8.1.19) were determined in Example 7.6. The first eigenvalue λ_{11} was found to be $\lambda_{11} = 2$, and the associated normalized eigenfunction is $M_{11}(x, y) = (2/\pi)\sin(x)\sin(y)$. We use the energy integral to estimate the first eigenvalue.

The appropriate energy integral for this problem is

$$E(u(x, y)) = \int_0^\pi \int_0^\pi [u_x^2(x, y) + u_y^2(x, y)]\, dx\, dy. \tag{8.1.20}$$

An admissible function $u(x, y)$ must vanish on the boundary of the square and must be normalized so that

$$(u(x, y), u(x, y)) = \int_0^\pi \int_0^\pi u^2(x, y)\, dx\, dy = 1. \tag{8.1.21}$$

Now if $u(x, y) = M_{11}(x, y)$, we have

$$(M_{11}(x, y),\ M_{11}(x, y)) = \frac{4}{\pi^2} \int_0^\pi \int_0^\pi \sin^2(x)\sin^2(y)\, dx\, dy = 1, \tag{8.1.22}$$

$$E(M_{11}(x, y)) = \frac{4}{\pi^2} \int_0^\pi \int_0^\pi [\cos^2(x)\sin^2(y) + \sin^2(x)\cos^2(y)]\, dx\, dy = 2, \tag{8.1.23}$$

so that $E(M_{11}(x, y)) = \lambda_{11}$, as was to be expected in view of (8.1.16).

As was shown, if $u(x, y)$ is any admissible function other than $M_{11}(x, y)$, we must have $E(u(x, y)) > \lambda_{11} = 2$. To see this in a particular case we consider the function $u(x, y) = (30/\pi^5)xy(\pi - x)(\pi - y)$. The coefficient $30/\pi^5$ has been prescribed so that (8.1.21) is satisfied and $u(x, y)$ clearly vanishes on the boundary. We readily find that

$$E(u(x,y)) = \frac{900}{\pi^{10}} \int_0^\pi \int_0^\pi [x^2(\pi-x)^2(\pi-2y)^2 + y^2(\pi-y)^2(\pi-2x)^2]\, dx\, dy = 2.03. \tag{8.1.24}$$

Thus $E(u(x, y)) = 2.03 > \lambda_{11} = 2$, as required. Given the arbitrary choice of $u(x, y)$, the value 2.03 is a remarkably good approximation to $\lambda_{11} = 2$.

Although $u(x, y)$ yields an excellent approximation to the eigenvalue λ_{11}, it does not appear to be a good approximation to the eigenfunction $M_{11}(x, y)$. In fact, if we measure the difference between $u(x, y)$ and $M_{11}(x, y)$ in the mean square norm, we obtain $||M_{11}(x, y) - u(x, y)||^2 = \int_0^\pi \int_0^\pi [M_{11}(x, y) - u(x, y)]^2\, dx\, dy = 0.73$, so that $||M_{11}(x, y) - u(x, y)|| \approx 0.85$, which is not small. This indicates that the variational approach may be expected to yield a much better approximation to the eigenvalues than to the eigenfunctions. The reason for this is demonstrated below.

Courant's Maximum-Minimum Principle

For the purpose of comparing the properties of eigenvalues arising in different eigenvalue problems, our formulation of the variational (minimum) problem is unsatisfactory. It characterizes the nth eigenvalue for a given problem in terms of the first $n - 1$ eigenfunctions of the problem, in view of the constraints (8.1.14). That is, we must first determine the eigenfunctions $M_1(\mathbf{x}), \ldots, M_{n-1}(\mathbf{x})$ in order to specify the eigenvalue λ_n. For the same reason, it is a simple matter to estimate the leading eigenvalue for the problem (as seen in Example 8.1) but not the remaining eigenvalues. It is of interest to obtain a variational formulation that permits the determination of any eigenvalue directly without having to solve any other problem. Such an approach was developed by *Courant* and is known as the *maximum-minimum principle.*

Given the eigenvalue problem (8.1.1)–(8.1.2) we consider a collection of (sufficiently smooth) functions $\{m_k(\mathbf{x})\}$, $k = 1, 2, 3, \ldots$, defined in the region G. These functions need not satisfy any specific conditions in G or on the boundary ∂G. To determine the nth eigenvalue λ_n, we introduce a function $w(\mathbf{x})$ that satisfies the boundary condition (8.1.9) and the $n - 1$ constraints

$$(w(\mathbf{x}),\ m_k(\mathbf{x})) = \iint_G \rho(\mathbf{x}) w(\mathbf{x}) m_k(\mathbf{x})\, dv = 0,\ k = 1, 2, \ldots, n-1. \tag{8.1.25}$$

Then we find the minimum of the Rayleigh quotient $E(w(\mathbf{x}))/(w(\mathbf{x}), w(\mathbf{x}))$ by varying over all admissible functions $w(\mathbf{x})$. For each set of functions $m_1(\mathbf{x}), \ldots, m_{n-1}(\mathbf{x})$, we obtain a minimum for the Rayleigh quotient. We then vary these minima over all possible sets of functions $m_1(\mathbf{x}), \ldots, m_{n-1}(\mathbf{x})$. The maximum of these minima is

equal to the nth eigenvalue λ_n. The *maximum-minimum principle* can be stated as

$$\max_{\{m_1,\ldots,m_{n-1}\}} \left[\min_{w(\mathbf{x})} \left(\frac{E(w(\mathbf{x}))}{(w(\mathbf{x}), w(\mathbf{x}))}\right)\right] = \lambda_n. \tag{8.1.26}$$

To prove this result, we note that on putting $w(\mathbf{x}) = \sum_{j=1}^{n} c_j M_j(\mathbf{x})$, where the $M_j(\mathbf{x})$ are eigenfunctions for (8.1.1)–(8.1.2), we may specify the constants c_j so that the $n-1$ constraints (8.1.25) are satisfied. In fact, we have $(w(\mathbf{x}), m_k(\mathbf{x})) = \sum_{j=1}^{n} c_j (M_j(\mathbf{x}), m_k(\mathbf{x}))$, $k = 1, 2, \ldots, n-1$. This is an underdetermined system of homogeneous linear equations for the n constants c_j, and it has a solution with at least one of the c_j remaining arbitrary. Therefore, we can satisfy the additional condition $(w(\mathbf{x}), w(\mathbf{x})) = \sum_{j=1}^{n} c_j^2 = 1$ by choosing the undetermined constant(s) appropriately. Note that $w(\mathbf{x})$ satisfies the boundary condition (8.1.9) since each of the $M_j(\mathbf{x})$ does so. Therefore, $w(\mathbf{x})$ is an admissible function for the minimum problem.

We have, for the $w(\mathbf{x})$ given above,

$$\frac{E(w(\mathbf{x}))}{(w(\mathbf{x}), w(\mathbf{x}))} = \sum_{j=1}^{n} \lambda_j c_j^2 \leq \lambda_n \sum_{j=1}^{n} c_j^2 = \lambda_n \tag{8.1.27}$$

in view of (8.1.11) and the normalization of $w(\mathbf{x})$, since $0 \leq \lambda_j \leq \lambda_n$, $j = 1, 2, \ldots, n-1$. Therefore, for each set of functions $m_1(\mathbf{x}), \ldots, m_{n-1}(\mathbf{x})$, we have found an admissible function $w(\mathbf{x})$ for which the Rayleigh quotient is less than or equal λ_n. Consequently, the maxima of all the minima of the Rayleigh quotient cannot exceed λ_n. However, if $w(\mathbf{x}) = M_n(\mathbf{x})$, the nth eigenfunction for the problem (8.1.1)–(8.1.2), we have

$$\frac{E(M_n(\mathbf{x}))}{(M_n(\mathbf{x}), M_n(\mathbf{x}))} = E(M_n(\mathbf{x})) = \lambda_n, \tag{8.1.28}$$

using (8.1.16), so that the *maximum-minimum principle* (8.1.26) is verified.

Variational Formulation of the Eigenvalue Problem

We have shown that subject to appropriate constraints, the minima of the Rayleigh quotient (8.1.17) are the eigenvalues of (8.1.1)–(8.1.2) and they are assumed when the minimizing functions are eigenfunctions. Reversing this process, we now formulate a *variational problem* for the *Rayleigh quotient* and show directly that the eigenvalues and eigenfunctions for (8.1.1)–(8.1.2) result from the solution of the minimum problem. The variational problem considered depends on the conditions placed on the admissible functions and the boundary conditions for the eigenvalue problem (8.1.1)–(8.1.2). As seen in (8.1.10), the energy integral $E(u)$ varies with the boundary conditions (8.1.9).

In the case of *Dirichlet boundary conditions* [i.e., $M(\mathbf{x}) = 0$ on ∂G] the appropriate *Rayleigh quotient* is

$$\frac{E(w(\mathbf{x}))}{||w(\mathbf{x})||^2} = \frac{1}{||w(\mathbf{x})||^2} \iint_G [\, p(\mathbf{x})(\nabla w(\mathbf{x}))^2 + q(\mathbf{x}) w^2(\mathbf{x})]\, dv. \tag{8.1.29}$$

In the case of *Neumann boundary conditions* [i.e., $\partial M(\mathbf{x})/\partial n = 0$ on ∂G] the Rayleigh quotient again has the form (8.1.29). However, for *Robin boundary conditions* [i.e., $\partial M(\mathbf{x})/\partial n + h(\mathbf{x})M(\mathbf{x}) = 0$ on ∂G with $h(\mathbf{x}) = \alpha(\mathbf{x})/\beta(\mathbf{x})$], we obtain, in view of (8.1.10),

$$\frac{E(w(\mathbf{x}))}{||w(\mathbf{x})||^2} = \frac{1}{||w(\mathbf{x})||^2}\left[\iint_G [\, p(\nabla w(\mathbf{x}))^2 + qw^2(\mathbf{x})]\, dv + \int_{\partial G} h(\mathbf{x})\rho w^2(\mathbf{x})\, ds\right]. \tag{8.1.30}$$

It is assumed that $h(\mathbf{x}) > 0$ and that p, q, ρ are functions of $\mathbf{x}$. In the case of mixed boundary conditions and if S_3 is not an empty set, the Rayleigh quotient has the form (8.1.30) with the surface integral taken over S_3.

The variational problem is given as follows. *The function $w(\mathbf{x})$ that yields a minimum for the variational problem*

$$\frac{E(w(\mathbf{x}))}{||w(\mathbf{x})||^2} = \text{minimum}, \tag{8.1.31}$$

among all functions $w(\mathbf{x})$ that vanish on ∂G for the case of Dirichlet boundary conditions and satisfy the constraints

$$(w(\mathbf{x}), M_1(\mathbf{x})) = (w(\mathbf{x}), M_2(\mathbf{x})) = \cdots = (w(\mathbf{x}), M_{n-1}(\mathbf{x})) = 0, \tag{8.1.32}$$

is—upon normalization—the nth *eigenfunction $M_n(\mathbf{x})$ of the problem* (8.1.1)–(8.1.2) *with $\alpha = 1$ and $\beta = 0$ in* (8.1.2). *The* nth *successive minimum value of the Rayleigh quotient* (8.1.31) *is the eigenvalue λ_n that corresponds to $M_n(\mathbf{x})$. The eigenvalues λ_j are ordered as $0 \le \lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$. For the eigenvalue problem* (8.1.1)–(8.1.2) *with boundary conditions of the second and third kinds, there no restrictions are placed on the boundary values of admissible functions w. In the case of mixed boundary conditions, we require that $w(\mathbf{x}) = 0$ on S_1.*

To demonstrate this result we assume that a sufficiently smooth minimizing function for the variational problem exists. (This is difficult to prove, in general.) Let the minimizing function be denoted by $w_n(\mathbf{x})$ and the admissible functions $w(\mathbf{x})$ be represented as, with ϵ is a constant parameter,

$$w(\mathbf{x}) = w_n(\mathbf{x}) + \epsilon W(\mathbf{x}), \tag{8.1.33}$$

where $W(\mathbf{x})$ satisfies the same admissibility conditions as $w(\mathbf{x})$.

In terms of (8.1.33) the *variational problem* becomes

$$\frac{E(w(\mathbf{x}))}{||w(\mathbf{x})||^2} = \frac{E(w_n(\mathbf{x}) + \epsilon W(\mathbf{x}))}{||w_n(\mathbf{x}) + \epsilon W(\mathbf{x})||^2} = \text{minimum}. \tag{8.1.34}$$

As a function of the parameter ϵ, the Rayleigh quotient in (8.1.34) assumes a minimum value when $\epsilon = 0$ [i.e., when $w(\mathbf{x}) = w_n(\mathbf{x})$]. Consequently, $\epsilon = 0$ must be a critical or stationary point of the function $E(w(\mathbf{x}))/||w^2(\mathbf{x})||$, so that its derivative with respect to ϵ must vanish at $\epsilon = 0$. As a result we have

$$\frac{\partial}{\partial \epsilon}\left\{\frac{E(w_n+\epsilon W)}{||w_n+\epsilon W||^2}\right\}\Bigg|_{\epsilon=0}=0 \tag{8.1.35}$$

$$=\frac{2}{||w_n||^2}\left[\iint_G\left[p\nabla w_n\cdot\nabla W+\left(q-\frac{\rho E(w_n)}{||w_n||^2}\right)w_nW\right]dv+\int_{\partial G}hpw_nW\,ds\right].$$

If we are dealing with the Dirichlet or Neumann problem, the surface integral is absent in (8.1.35). Using (4.2.17) and the divergence theorem gives

$$\iint_G p\nabla w_n\cdot\nabla W\,dv=-\iint_G W\nabla\cdot(p\nabla w_n)\,dv+\iint_G\nabla\cdot(pW\nabla w_n)\,dv$$

$$=-\iint_G W\nabla\cdot(p\nabla w_n)\,dv+\int_{\partial G}pW\frac{\partial w_n}{\partial n}\,ds. \tag{8.1.36}$$

Introducing (8.1.36) into (8.1.35) yields

$$\iint_G\left[-\nabla\cdot(p\nabla w_n)+\left[q-\frac{\rho E(w_n)}{||w_n||^2}\right]w_n\right]W\,dv+\int_{\partial G}\left[\frac{\partial w_n}{\partial n}+hw_n\right]W\,pds=0. \tag{8.1.37}$$

For the eigenvalue problem with *Dirichlet boundary conditions*, the surface integral in (8.1.37) vanishes since $W(\mathbf{x})=0$ on the boundary. From the arbitrariness of $W(\mathbf{x})$ in the region G (see Exercises 8.1.7 and 8.1.8 for some additional details regarding this point) we conclude that the bracketed term in the volume integral must vanish. Therefore, $w_n(\mathbf{x})$ satisfies the equation

$$-\nabla\cdot(p(\mathbf{x})\nabla w_n(\mathbf{x}))+q(\mathbf{x})w_n(\mathbf{x})=\frac{E(w_n(\mathbf{x}))}{||w_n(\mathbf{x})||^2}\rho(\mathbf{x})w_n(\mathbf{x}), \tag{8.1.38}$$

with $w_n(\mathbf{x})=0$ on ∂G. This implies that

$$\frac{E(w_n(\mathbf{x}))}{||w_n(\mathbf{x})||^2}=\lambda_n \tag{8.1.39}$$

and $w_n(\mathbf{x})/||w_n(\mathbf{x})||=M_n(\mathbf{x})$, in view of (8.1.1)–(8.1.2).

In the case of *Neumann boundary conditions* for the eigenvalue problem, we must set $h=0$ in the surface integral in (8.1.35). The arbitrariness of $W(\mathbf{x})$ again implies that the integrals over G and ∂G in (8.1.37) must vanish separately. Thus $w_n(\mathbf{x})$ again is a solution of (8.1.38), but the vanishing of the surface integral leads to the *natural boundary condition*

$$\frac{\partial w_n(\mathbf{x})}{\partial n}\Bigg|_{\partial G}=0. \tag{8.1.40}$$

We conclude that $w_n(\mathbf{x})/||w_n(\mathbf{x})||=M_n(\mathbf{x})$, the nth eigenfunction for the problem (8.1.1) with $\alpha(\mathbf{x})=0$ and $\beta(\mathbf{x})=1$ in (8.1.2), and that (8.1.39) again yields the eigenvalue λ_n.

For the *Robin boundary value problem* [i.e., with $h(\mathbf{x}) \neq 0$], the arbitrariness of $W(\mathbf{x})$ implies that both integrals in (8.1.37) vanish separately. The *natural boundary condition* for $w_n(\mathbf{x})$ becomes

$$\frac{\partial w_n(\mathbf{x})}{\partial n} + h(\mathbf{x}) w_n(\mathbf{x}) \Big|_{\partial G} = 0 \tag{8.1.41}$$

and $w_n(\mathbf{x})$ satisfies (8.1.38). Consequently, $w_n(\mathbf{x})/||w_n(\mathbf{x})|| = M_n(\mathbf{x})$, the nth eigenfunction for (8.1.1), with $\alpha(\mathbf{x})/\beta(\mathbf{x}) = h(\mathbf{x})$ in (8.1.2), and (8.1.39) yields the eigenvalue λ_n.

In the case of *mixed boundary conditions*, we conclude as before that $w_n(\mathbf{x})$ satisfies (8.1.38). On S_1, where Dirichlet conditions are given, we have $w_n(\mathbf{x}) = 0$. On S_2, $w_n(\mathbf{x})$ satisfies the natural boundary condition (8.1.40) (where ∂G is replaced by S_2). On S_3, $w_n(\mathbf{x})$ satisfies the natural boundary condition (8.1.41) (where ∂G is replaced by S_3).

We observe from (8.1.33) that $||w(\mathbf{x}) - w_n(\mathbf{x})|| = O(\epsilon)$, but for the approximate eigenvalue $\lambda = E(w(\mathbf{x}))/||w(\mathbf{x})||^2$, we have $\lambda - \lambda_n = O(\epsilon^2)$, since the $O(\epsilon)$ term in the expansion of λ in powers of ϵ vanishes, noting (8.1.35). This accounts for the fact that we consistently obtain better approximations for the eigenvalues than for the eigenfunctions when using the variational method.

This concludes our direct demonstration that the eigenvalues and eigenfunctions for (8.1.1)–(8.1.2) can be determined from a variational problem. Since (based on our assumptions) the Rayleigh quotient is nonnegative, so is its minimum and we must have $\lambda_1 \geq 0$. The ordering of the eigenvalues is considered in Exercise 8.1.6.

Distribution of the Eigenvalues

The foregoing variational problem requires that the first $n-1$ eigenfunctions must be known if we are to determine the nth eigenvalue and eigenfunction. However, *Courant's maximum-minimum principle* permits the determination of each eigenvalue independently of the other (lesser) eigenvalues. Restricting ourselves to the *variational eigenvalue problem* with *Dirichlet boundary conditions*, we now show that the eigenvalues $\lambda_n \to \infty$ as $n \to \infty$. Although the variational problem shows (see Exercise 8.1.6) that the eigenvalues can be ordered as in (8.1.3), it may happen that the eigenvalues have a finite limit point and that infinitely many (independent) eigenfunctions can be associated with a single eigenvalue. However, the fact that $\lambda_n \to \infty$ implies that each eigenvalue has a *finite multiplicity*; that is, there are only a finite number of linearly independent eigenfunctions associated with each distinct eigenvalue.

The use of the maximum-minimum principle permits the comparison of variational problems with different sets of coefficients in the Rayleigh quotients and problems defined over different regions G. We assume that the coefficients $p(\mathbf{x})$, $\rho(\mathbf{x})$, and $q(\mathbf{x})$ that occur in the Rayleigh quotient have maximum and minimum values in the region G together with its boundary ∂G and continue to require that $p(\mathbf{x}) > 0$, $\rho(\mathbf{x}) > 0$, and $q(\mathbf{x}) \geq 0$ in G. Let the constants p_M, ρ_M, and q_M, represent the maximum values and p_m, ρ_m, and q_m, represent the minimum values of the functions $p(\mathbf{x})$, $\rho(\mathbf{x})$, and

$q(\mathbf{x})$, respectively. (We assume that p_m and ρ_m are not zero.) Further, the admissible functions $w(\mathbf{x})$ will be required to vanish on ∂G and to satisfy the constraints (8.1.25).

On examining the Rayleigh quotient appropriate for the Dirichlet problem,

$$\frac{E(w(\mathbf{x}))}{||w(\mathbf{x})||^2} = \frac{\iint_G [p(\mathbf{x})(\nabla w(\mathbf{x}))^2 + q(\mathbf{x})w^2(\mathbf{x})]\, dv}{\iint_G \rho(\mathbf{x})w^2(\mathbf{x})\, dv}, \tag{8.1.42}$$

it is clear that if we replace $p(\mathbf{x})$, $\rho(\mathbf{x})$, and $q(\mathbf{x})$ by the constants p_M, ρ_m, and q_M, we obtain a new Rayleigh quotient, $E_M(w(\mathbf{x}))/||w(\mathbf{x})||_m^2$, for which

$$\frac{E_M(w(\mathbf{x}))}{||w(\mathbf{x})||_m^2} \geq \frac{E(w(\mathbf{x}))}{||w(\mathbf{x})||^2} \tag{8.1.43}$$

for each admissible function $w(\mathbf{x})$. Now the inner product that occurs in the constraints (8.1.18) for the new problem with constant coefficients can be expressed as

$$(w(\mathbf{x}),\ m_k(\mathbf{x})) = \iint_G \rho_m w(\mathbf{x}) m_k(\mathbf{x})\, dv = \iint_G \rho(\mathbf{x})w(\mathbf{x}) \left(m_k(\mathbf{x}) \frac{\rho_m}{\rho(\mathbf{x})} \right) dv. \tag{8.1.44}$$

Thus the constraints for the new problem must be given in terms of the set of functions $\hat{m}_k(\mathbf{x}) = (\rho_m/\rho(\mathbf{x}))m_k(\mathbf{x})$ if we want the new problem to have the same set of constraints. But in the maximum-minimum principle the set of functions $m_k(\mathbf{x})$ ranges over all sufficiently smooth functions and, therefore, so does the set $\hat{m}_k(\mathbf{x})$. We conclude that the max-min of $E_M(w(\mathbf{x}))/||w(\mathbf{x})||_m^2$ is greater than or equal to the max-min of $E(w(\mathbf{x}))/||w(\mathbf{x})||^2$.

Similarly, replacing $p(\mathbf{x})$, $\rho(\mathbf{x})$, and $q(\mathbf{x})$ in (8.1.42) by the constants p_m, ρ_M, and q_m yields a modified Rayleigh quotient $E_m(w(\mathbf{x}))/||w(\mathbf{x})||_M^2$ for which

$$\frac{E_m(w(\mathbf{x}))}{||w(\mathbf{x})||_M^2} \leq \frac{E(w(\mathbf{x}))}{||w(\mathbf{x})||^2}. \tag{8.1.45}$$

The foregoing procedure may be applied to the constraints (8.1.25), and we conclude that the max-min of $E_m(w(\mathbf{x}))/||w(\mathbf{x})||_M^2$ does not exceed the max-min of $E(w(\mathbf{x}))/||w(\mathbf{x})||^2$. For the set of admissible functions $w(\mathbf{x})$, vanishing on ∂G and satisfying the constraints (8.1.25), each of the aforementioned max-min yields the nth eigenvalue for the given variational problem. Thus

$$\lambda_n^{(m)} \leq \lambda_n \leq \lambda_n^{(M)}, \tag{8.1.46}$$

where $\lambda_n^{(m)}$, λ_n, and $\lambda_n^{(M)}$ are the nth eigenvalues for the variational problem with the Rayleigh quotients $E_m(w(\mathbf{x}))/||w(\mathbf{x})||_M^2$, $E(w(\mathbf{x}))/||w(\mathbf{x})||^2$, and $E_M(w(\mathbf{x}))/||w(\mathbf{x})||_m^2$, respectively.

Next we introduce changes in the region G. Suppose that we consider a subregion $\hat{G}$ contained in G (i.e., $\hat{G} \subset G$, where a portion of the boundary $\partial\hat{G}$ may coincide with ∂G). Let $E(w(\mathbf{x}))/||w(\mathbf{x})||^2$ be the Rayleigh quotient for the variational problem in G. Suppose we require that all admissible functions $w(\mathbf{x})$, in addition to vanishing

on ∂G and satisfying the constraints (8.1.25), must vanish on $\partial\hat{G}$ and in that part of G exterior to $\hat{G}$. Since the additional requirements on the functions $w(\mathbf{x})$ effectively reduces the number of admissible functions for the variational problem, the maximum-minimum for the new problem cannot be smaller than the max-min λ_n attained when the added restrictions are absent. However, the new set of admissible functions is precisely that which is appropriate for the variational problem in the subregion $\hat{G}$. The max-min for the variational problem in $\hat{G}$ is $\hat{\lambda}_n$, the nth eigenvalue for that problem. As a result, we obtain

$$\lambda_n \leq \hat{\lambda}_n. \tag{8.1.47}$$

To complete our discussion, we consider two rectangular regions, R_M and R_m. (In three dimensions these are rectangular boxes.) The region R_M contains the region G within it, whereas the region R_m is completely contained within G, so that $R_m \subset G \subset R_M$. First we consider the max-min variational problem for the Rayleigh quotient $E_M(w(\mathbf{x}))/||w(\mathbf{x})||_m^2$, with the region G replaced by R_m but the problem otherwise unchanged. In view of (8.1.47), we conclude that

$$\lambda_n^{(M)} \leq \hat{\lambda}_n^{(M)}, \tag{8.1.48}$$

where $\lambda_n^{(M)}$ and $\hat{\lambda}_n^{(M)}$ are the nth eigenvalues for the problem associated with $E_M(w(\mathbf{x}))/||w(\mathbf{x})||_m^2$, over the regions G and R_m, respectively.

Next, consider the max-min variational problem for $E_m(w(\mathbf{x}))/||w(\mathbf{x})||_M^2$ with the region G replaced by R_M but the problem otherwise unchanged. Using (8.1.47) gives

$$\hat{\lambda}_n^{(m)} \leq \lambda_n^{(m)}, \tag{8.1.49}$$

where $\hat{\lambda}_n^{(m)}$ and $\lambda_n^{(m)}$ are the nth eigenvalues for the problem associated with $E_m(w(\mathbf{x}))/||w(\mathbf{x})||_M^2$ over the regions R_M and G, respectively.

Combining (8.1.46), (8.1.48), and (8.1.49) yields the set of inequalities

$$\hat{\lambda}_n^{(m)} \leq \lambda_n^{(m)} \leq \lambda_n \leq \lambda_n^{(M)} \leq \hat{\lambda}_n^{(M)}. \tag{8.1.50}$$

The eigenvalues $\hat{\lambda}_n^{(m)}$ and $\hat{\lambda}_n^{(M)}$ can be determined exactly by separation of variables since the equations have constant coefficients and the regions are rectangular. This has already been carried out for a special two-dimensional problem in Example 7.6, and the general case is considered below for two and three dimensions. Since the eigenvalues $\hat{\lambda}_n^{(m)}$ and $\hat{\lambda}_n^{(M)}$ tend to infinity as $n \to \infty$ (as shown below), so do the eigenvalues λ_n for the Dirichlet problem, as was to be shown.

Many additional properties of eigenvalues may be obtained by arguments similar to those given. For example, it can be shown that the eigenvalues λ_n for the *Neumann problem, the Robin boundary value problem,* and *the mixed boundary value problem* also tend to infinity as $n \to \infty$ but we do not prove this here. The property that $\lambda_n \to \infty$ as $n \to \infty$, is used below to prove completeness in the mean square sense of the eigenfunctions for the Dirichlet problem. A similar argument can be used to prove completeness in the mean square sense of the eigenfunctions for the

other boundary value problems referred to above, given the fact that the eigenvalues $\lambda_n \to \infty$ as $n \to \infty$.

Dirichlet Eigenvalue Problems for Elliptic Equations with Constant Coefficients

We consider Dirichlet eigenvalue problems for elliptic equations with constant coefficients. The results obtained serve to verify some of our conclusions and illustrate some of the properties of eigenvalues obtained above.

The *eigenvalue problem*

$$-p\nabla^2 M(\mathbf{x}) + qM(\mathbf{x}) = \lambda\rho M(\mathbf{x}) \tag{8.1.51}$$

in the region G with the *Dirichlet boundary condition*

$$M(\mathbf{x})\big|_{\partial G} = 0, \tag{8.1.52}$$

where p, q, and ρ are assumed to be constants, is associated with the *variational problem*

$$\frac{E(w(\mathbf{x}))}{||w(\mathbf{x})||^2} = \frac{\iint_G [p(\nabla w(\mathbf{x}))^2 + qw^2(\mathbf{x})]\, dv}{\iint_G \rho w^2(\mathbf{x})\, dv} = \text{minimum}, \tag{8.1.53}$$

where all admissible functions $w(\mathbf{x})$ must vanish on ∂G and satisfy appropriate additional constraints.

Now (8.1.51) can be written as

$$\nabla^2 M(\mathbf{x}) + \frac{\lambda\rho - q}{p} M(\mathbf{x}) \equiv \nabla^2 M(\mathbf{x}) + \tilde{\lambda} M(\mathbf{x}) = 0; \tag{8.1.54}$$

that is, $\tilde{\lambda} = (\lambda\rho - q)/p$. Similarly, the variational problem (8.1.53) can be expressed as

$$\frac{E(w(\mathbf{x}))}{||w(\mathbf{x})||^2} = \frac{p}{\rho} \frac{\iint_G (\nabla w(\mathbf{x}))^2\, dv}{\iint_G w^2(\mathbf{x})\, dv} + \frac{q}{\rho} = \text{minimum}. \tag{8.1.55}$$

Both (8.1.54) and (8.1.55) imply that if we can solve a simplified version of (8.1.51)–(8.1.52) or (8.1.53) with $p = \rho = 1$ and $q = 0$ [we just get the negative Laplacian on the left in (8.1.52)] and we denote the eigenvalues of that simplified problem by $\tilde{\lambda}_n$, then the eigenvalues λ_n of (8.1.51)–(8.1.52) or (8.1.53) are given as

$$\lambda_n = \frac{p\,\tilde{\lambda}_n + q}{\rho}, \qquad n = 1, 2, 3, \ldots. \tag{8.1.56}$$

The relationship (8.1.56) between λ_n and $\tilde{\lambda}_n$ demonstrates the following properties discussed in the general case of variable coefficients. If the coefficients p and q are increased, the eigenvalues increase, whereas the reverse is true if p and q are decreased.

Also if ρ is decreased, the eigenvalues are increased, whereas if ρ is increased, the eigenvalues decrease.

Next we determine an explicit set of eigenvalues for several regions G. First we specialize G to be a *rectangle* as given in Example 7.6, with $0 < x < l$ and $0 < y < \hat{l}$. The eigenvalues for (8.1.54) and (8.1.52) are given as

$$\tilde{\lambda}_{nm} = \left(\frac{\pi n}{l}\right)^2 + \left(\frac{\pi m}{\hat{l}}\right)^2, \qquad n,\ m = 1, 2, 3, \ldots, \tag{8.1.57}$$

as shown in that example.

Similarly, on considering the *rectangular box* $0 < x < l$, $0 < y < \hat{l}$, and $0 < z < \tilde{l}$ in three dimensions and separating variables, we readily obtain the eigenvalues for (8.1.54) and (8.1.52) as

$$\tilde{\lambda}_{nmk} = \left(\frac{\pi n}{l}\right)^2 + \left(\frac{\pi m}{\hat{l}}\right)^2 + \left(\frac{\pi k}{\tilde{l}}\right)^2, \qquad n,\ m,\ k = 1, 2, 3, \ldots. \tag{8.1.58}$$

In both cases the eigenvalues can be expressed as a sequence $\{\tilde{\lambda}_n\}$ and we have $\tilde{\lambda}_n \to \infty$ as $n \to \infty$, since $\tilde{\lambda}_{nm} \to \infty$ and $\tilde{\lambda}_{nmk} \to \infty$ as their subscripts tend to infinity. Consequently, in view of (8.1.56), $\lambda_n \to \infty$ as $n \to \infty$, since p and ρ are positive. Furthermore, the rectangular regions R_M and R_m introduced in our earlier discussion may be assumed to have sides parallel to the coordinate lines or planes in the two- or three-dimensional cases, respectively. Since the eigenvalue equation in the case of constant coefficients is invariant under the translation of axes, the eigenvalues are given as in (8.1.57) or (8.1.58). Therefore, we conclude that the eigenvalues $\hat{\lambda}_n^{(M)}$ and $\hat{\lambda}_n^{(m)}$ tend to infinity as $n \to \infty$, and (8.1.50) implies that $\lambda_n \to \infty$ as $n \to \infty$.

It is seen from (8.1.57) and (8.1.58) that as the dimensions of the rectangular region G are changed, the eigenvalues change in a specific manner. As the lengths of the sides are decreased, the eigenvalues $\tilde{\lambda}_{nm}$ and $\tilde{\lambda}_{nmk}$ increase, whereas the reverse is true if the lengths of the sides are increased. This result is consistent with our conclusion relating the eigenvalues of region G to those of a subregion $\hat{G}$.

The following example uses the foregoing results to obtain an upper and lower bound for the first eigenvalue for the Laplacian in the unit circle with Dirichlet boundary conditions.

Example 8.2. The Dirichlet Eigenvalue Problem for the Laplacian in a Circle. We discuss the *Dirichlet eigenvalue problem* for the *unit circle*. First the eigenvalues and eigenfunctions are determined by means of separation of variables. Then the inequalities (8.1.50) are used to bound the leading eigenvalue for the circle in terms of the eigenvalues for inscribed and circumscribed squares.

Expressing the Laplacian operator in polar coordinates, we consider the eigenvalue problem

$$\nabla^2 M(r,\theta) + \lambda M(r,\theta) = M_{rr}(r,\theta) + \frac{1}{r} M_r(r,\theta) + \frac{1}{r^2} M_{\theta\theta}(r,\theta) + \lambda M(r,\theta) = 0 \tag{8.1.59}$$

in the unit disk $0 \leq r < 1$, with the Dirichlet boundary condition

$$M(1, \theta) = 0. \tag{8.1.60}$$

The solutions $M(r, \theta)$ are required to be single valued in the disk, so that

$$M(r, \theta + 2\pi) = M(r, \theta). \tag{8.1.61}$$

To solve for $M(r, \theta)$ we use separation of variables and set $M(r, \theta) = F(r)G(\theta)$. Inserting this expression into (8.1.59) gives

$$r^2 \frac{F''(r) + (1/r)F'(r) + \lambda F(r)}{F(r)} = -\frac{G''(\theta)}{G(\theta)} = m^2, \tag{8.1.62}$$

where m^2 is the separation constant. The equation for $G(\theta)$ is

$$G''(\theta) + m^2 G(\theta) = 0. \tag{8.1.63}$$

The condition (8.1.61) requires $G(\theta)$ to be periodic of period 2π, and this implies that $m = n$, an integer. The equation for $F_n(r)$ (i.e., for each n) is

$$F_n''(r) + \frac{1}{r} F_n'(r) + \left(\lambda - \frac{n^2}{r^2}\right) F_n(r) = 0, \tag{8.1.64}$$

which is Bessel's equation of order n. This equation was discussed in Example 4.7. The solutions $F_n(r)$ that are finite at $r = 0$ are the Bessel functions and are given as $F_n(r) = J_n(\sqrt{\lambda} r)$. The boundary condition (8.1.60) requires that $J_n(\sqrt{\lambda}) = 0$, so that the eigenvalues λ are the squares of the zeros of the Bessel functions. These were denoted by α_{kn} $(k = 1, 2, \ldots)$ in Example 4.7, so that we have

$$\lambda_{kn} = (\alpha_{kn})^2, \qquad k = 1, 2, \ldots, \qquad n = 0, 1, 2, \ldots. \tag{8.1.65}$$

Since with $m = n$, (8.1.63) yields the trigonometric functions $\cos(n\theta)$ and $\sin(n\theta)$, the eigenfunctions $M_{kn}(r, \theta)$ are obtained from the functions $J_0(\alpha_{k0} r)$, $J_n(\alpha_{kn} r)\cos(n\theta)$, $J_n(\alpha_{kn} r)\sin(n\theta)$. It is not difficult to normalize this orthogonal set of eigenfunctions. Here we are mainly concerned with the eigenvalues λ_{kn}, and the first eigenvalue is given in terms of the first zero of the Bessel function $J_0(x)$, that is, $\lambda_{10} = (\alpha_{10})^2 \approx (2.40)^2 = 5.76$.

To obtain an upper and lower bound for the first eigenvalue λ_{10} for the unit circle, we consider inscribed and circumscribed squares of sides $\sqrt{2}$ and 2, respectively, as pictured in Figure 8.1. The leading eigenvalue $\tilde{\lambda}_{11}$ for the circumscribed square is, on using (8.1.57), $\tilde{\lambda}_{11} = 2\,(\pi/2)^2 = \pi^2/2 \approx 4.93$. The first eigenvalue for the inscribed square is $\hat{\lambda}_{11} = 2\left(\pi/\sqrt{2}\right)^2 = \pi^2 \approx 9.87$, so that $\tilde{\lambda}_{11} < \lambda_{10} < \hat{\lambda}_{11}$, consistent with (8.1.50). This is a poor method for approximating the eigenvalue λ_{10} since the squares do not approximate the circular region very well. However, this method can be used to provide bounds for all the zeros of the Bessel functions. In Section 8.2 we use the Rayleigh-Ritz method to approximate the first eigenvalue λ_{10} for the circle.

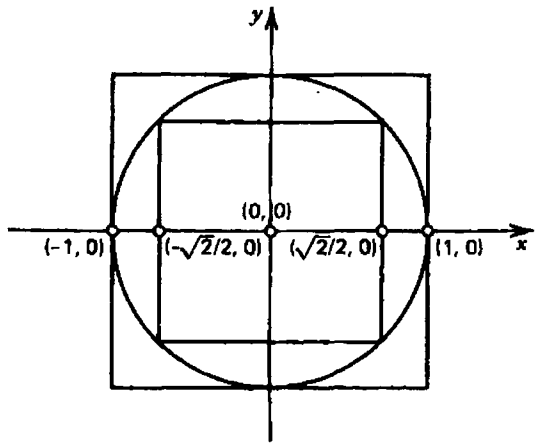

Figure 8.1 Eigenvalues for a circle.

Completeness of the Eigenfunctions

We can now prove the completeness of the orthonormalized set of eigenfunctions $\{M_k(\mathbf{x})\}$ that satisfy (8.1.1) in G and the Dirichlet conditions $M_k(\mathbf{x}) = 0$ on ∂G. Let $u(\mathbf{x})$ be a smooth function defined in G that vanishes on ∂G [i.e., $u(\mathbf{x})$ satisfies the admissibility conditions for the variational problem]. We expand $u(\mathbf{x})$ in a series of eigenfunctions as in (8.1.5) with the Fourier coefficients given in (8.1.6). [$u(\mathbf{x})$ need be normalized as in (8.1.7).] Let the remainder $R_n(\mathbf{x})$ be defined as

$$R_n(\mathbf{x}) = u(\mathbf{x}) - \sum_{k=1}^{n} N_k M_k(\mathbf{x}). \tag{8.1.66}$$

Then we show that

$$\lim_{n\to\infty} ||R_n(\mathbf{x})|| = \lim_{n\to\infty} \left|\left| u(\mathbf{x}) - \sum_{k=1}^{n} N_k M_k(\mathbf{x}) \right|\right| = 0, \tag{8.1.67}$$

so that (8.1.5) converges to $u(\mathbf{x})$ in the *mean square sense,* and the completeness of the eigenfunctions $\{M_k(\mathbf{x})\}$ is demonstrated. The norm $||R_n(\mathbf{x})||$ is defined as $||R_n(\mathbf{x})||^2 = \iint_G \rho(\mathbf{x}) R_n^2(\mathbf{x})\, dv$.

Since $u(\mathbf{x})$ and the eigenfunctions $\{M_k(\mathbf{x})\}$ are admissible for the variational problem, so is the remainder term $R_n(\mathbf{x})$. In addition, for $j = 1, \ldots, n$,

$$(R_n(\mathbf{x}),\ M_j(\mathbf{x})) = (u(\mathbf{x}),\ M_j(\mathbf{x})) - \sum_{k=1}^{n} N_k (M_k(\mathbf{x}),\ M_j(\mathbf{x})) = N_j - N_j = 0, \tag{8.1.68}$$

in view of (8.1.4) and (8.1.6). Consequently, $R_n(\mathbf{x})$ satisfies the constraints for the minimum problem that determines the $(n+1)$st eigenvalue λ_{n+1} and this yields for

the Rayleigh quotient (8.1.42),

$$\frac{E(R_n(\mathbf{x}))}{||R_n(\mathbf{x})||^2} \geq \lambda_{n+1}, \tag{8.1.69}$$

since the minimum is assumed for the eigenfunction $M_{n+1}(\mathbf{x})$. We write (8.1.69) as

$$||R_n(\mathbf{x})||^2 \leq \frac{E(R_n(\mathbf{x}))}{\lambda_{n+1}}. \tag{8.1.70}$$

We already know that $\lambda_{n+1} \to \infty$ as $n \to \infty$, so that we need only show that $E(R_n(\mathbf{x}))$ is bounded as $n \to \infty$ to conclude that $||R_n(\mathbf{x})|| \to 0$ as $n \to \infty$.

Now, as is easily shown,

$$E(u(\mathbf{x})) = E\left[R_n(\mathbf{x}) + \sum_{k=1}^{n} N_k M_k(\mathbf{x})\right] = E(R_n(\mathbf{x})) + E\left[\sum_{k=1}^{n} N_k M_k(\mathbf{x})\right]$$

$$+ 2\sum_{k=1}^{n} N_k \iint_G [p(\mathbf{x})\nabla R_n(\mathbf{x}) \cdot \nabla M_k(\mathbf{x}) + q(\mathbf{x}) R_n(\mathbf{x}) M_k(\mathbf{x})]\, dv. \tag{8.1.71}$$

Using (8.1.36), the integral term in (8.1.71) can be expressed as

$$\iint_G [p(\mathbf{x})\nabla R_n(\mathbf{x}) \cdot \nabla M_k(\mathbf{x}) + q(\mathbf{x}) R_n(\mathbf{x}) M_k(\mathbf{x})]\, dv \tag{8.1.72}$$

$$= \iint_G R_n(\mathbf{x})[-\nabla \cdot (p(\mathbf{x})\nabla M_k(\mathbf{x})) + q(\mathbf{x}) M_k(\mathbf{x})]\, dv = \lambda_k(R_n(\mathbf{x}),\ M_k(\mathbf{x})) = 0,$$

where we have used the fact that $R_n(\mathbf{x}) = 0$ on ∂G, the orthogonality properties (8.1.68), as well as the fact that $M_k(\mathbf{x})$ satisfies (8.1.1) with $\lambda = \lambda_k$. Furthermore, we obtain from (8.1.11),

$$E\left[\sum_{k=1}^{n} N_k M_k(\mathbf{x})\right] = \sum_{j=1}^{n}\sum_{k=1}^{n} \lambda_k N_k N_j (M_k(\mathbf{x}), M_j(\mathbf{x})) = \sum_{k=1}^{n} \lambda_k N_k^2. \tag{8.1.73}$$

Combining these results gives

$$E(R_n(\mathbf{x})) = E(u(\mathbf{x})) - \sum_{k=1}^{n} \lambda_k N_k^2 \leq E(u(\mathbf{x})), \tag{8.1.74}$$

since $E(u(\mathbf{x}))$ is finite by assumption and the eigenvalues λ_k are nonnegative. We conclude from (8.1.74) that $E(R_n(\mathbf{x}))$ is bounded for all n. Therefore, (8.1.70) shows that $||R_n(\mathbf{x})|| \to 0$ as $n \to \infty$, and the *mean square convergence* of the series (8.1.5) to $u(\mathbf{x})$ and the *completeness* of the eigenfunctions $M_k(\mathbf{x})$ is demonstrated. As was indicated earlier, if we can show that the eigenvalues λ_n for the second, third, and mixed boundary value problems tend to infinity, we can use the given argument to show that the corresponding set of eigenfunctions is complete.

For the one-dimensional *Sturm-Liouville eigenvalue problem* discussed in Section 4.3, the appropriate Rayleigh quotient is

$$\frac{E(w(x))}{||w(x)||^2} = \frac{\int_0^l [p(x)(w'(x))^2 + q(x)w^2(x)]\,dx}{\int_0^l \rho(x)w^2(x)\,dx}, \tag{8.1.75}$$

when the admissible functions $w(x)$ satisfy the Dirichlet conditions $w(0) = w(l) = 0$. The appropriate Rayleigh quotients for the second, third, and mixed boundary value problems are presented in the exercises. All the basic results given for the higher dimensional eigenvalue problem carry over to the Sturm-Liouville problem. Thus the completeness of the eigenfunctions (i.e., the sine functions) of Example 4.4, and the convergence of the Fourier sine series in the mean square sense can be demonstrated.

We note that we can drop our assumption that the function $u(\mathbf{x})$ expanded in a series of eigenfunctions must vanish on the boundary ∂G. This follows since any admissible function $u(\mathbf{x})$ that does not vanish on ∂G can be approximated arbitrarily closely, in the mean square norm, by a function that does vanish on ∂G. In this norm the values of functions at specific points or lower dimensional regions do not play a significant role, in general.

Solutions of the *boundary value problems* for the self-adjoint elliptic equations considered in Section 4.1 can also be determined by means of a variational principle. Indeed, *Dirichlet's principle* is a classical variational principle that determines the solution of the Dirichlet problem for Laplace's equation to be the minimum of an energy integral. Dirichlet's principle and more general variational principles are considered in the exercises. We have discussed variational principles for eigenvalue problems in this section because the results are useful in solving problems for equations of all three types as indicated.

Exercises 8.1

8.1.1. Consider the eigenvalue problem of Example 8.1 (see Example 7.6). Show that the eigenvalues can be ordered as in (8.1.3) and that the energy integral 8.1.20) satisfies (8.1.16).

8.1.2. Consider the eigenvalue problem of Example 8.1 with the Dirichlet condition (8.1.19) replaced by the Neumann condition $\partial M(x,y)/\partial n\big|_{\partial G} = 0$ (see Example 7.7). Show that the eigenvalues can be ordered as in (8.1.3) and that the appropriate energy integral [i.e., (8.1.20)] satisfies (8.1.16).

8.1.3. Obtain the appropriate form for the energy integral (8.1.10) for the Sturm-Liouville eigenvalue problem; $-(p(x)u'(x))' + q(x)u(x) = \lambda\rho(x)u(x)$, $0 < x < l$, $u'(0) - h_1u(0) = 0$, $u'(l) + h_2u(l) = 0$, with positive h_1 and h_2.

8.1.4. Show that the energy integral (8.1.10) for the eigenvalue problem $u''(x) + \lambda u(x) = 0$, $0 < x < l$, $u(0) = 0$, $u'(l) + hu(l) = 0$, $h > 0$, is $E(u(x)) = \int_0^l (u'(x))^2\,dx + hu^2(l)$. Let $u_n(x) = \sin(\sqrt{\lambda_n}x)$ be the (unnormalized)

eigenfunctions for this problem with the λ_n determined from $\sqrt{\lambda_n}\cos(\sqrt{\lambda_n}\, l) + h\sin(\sqrt{\lambda_n}\, l) = 0$. Demonstrate that $E(u_n(x)) = \lambda_n(u_n(x), u_n(x))$. *Hint*: Use the eigenvalue equation to simplify the energy integral.

8.1.5. Consider the function $u(x, y) = cxy\sin(x)\sin(y)$. Determine the constant c such that the norm (8.1.21) of this function is unity. Observe that $u(x, y)$ vanishes on the boundary of the square $0 < x < \pi$, $0 < y < \pi$, and evaluate the energy integral (8.1.21) for the function. Compare your result with the eigenvalue $\lambda_{11} = 2$ for the square and evaluate the norm $||M_{11}(x, y) - u(x, y)||$ as in Example 8.1.

8.1.6. Show that the eigenvalues found from the variational principle (8.1.31) with the constraints (8.1.32) can indeed be ordered as $0 \leq \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n \leq \cdots$, by noting that to determine λ_n an additional constraint is added to those already given for the preceding eigenvalues so that the class of admissible functions is reduced.

8.1.7. It is required in the text that $W(\mathbf{x})$ defined in (8.1.33) satisfy the constraints (8.1.32). Show that if we set $W(\mathbf{x}) = \hat{W}(\mathbf{x}) + \sum_{j=1}^{n-1} a_j M_j(\mathbf{x})$, where $\hat{W}(\mathbf{x})$ is arbitrary [apart from having to vanish on ∂G if $W(\mathbf{x})$ vanishes there], we may choose the constants $a_j = -(\hat{W}(\mathbf{x}), M_j(\mathbf{x}))$ so that $W(\mathbf{x})$ satisfies the constraints (8.1.32). Replace $W(\mathbf{x})$ by the preceding expression in the discussion following (8.1.35) and show that (8.1.37) results with $W(\mathbf{x})$ replaced by $\hat{W}(\mathbf{x})$.

8.1.8. Show that if $\iint_G f(\mathbf{x})g(\mathbf{x})\, dv = 0$ for a continuous function $f(\mathbf{x})$ in G and arbitrary functions $g(\mathbf{x})$, we must have $f(\mathbf{x}) = 0$. This result is often called the *fundamental lemma of the calculus of variations*. *Hint*: Assume that $f(\mathbf{x}) \neq 0$ at some point in G, so that it is nonzero in the neighborhood of that point because it is continuous, and then choose $g(\mathbf{x})$ appropriately.

8.1.9. Show that if $\iint_G f(\mathbf{x})\, dv = 0$ for arbitrary regions G within which $f(\mathbf{x})$ is continuous, we must have $f(\mathbf{x}) = 0$. *Hint*: Proceed as in Exercise 8.1.8. This result is known as the *du Bois-Reymond lemma* and is closely related to the lemma of Exercise 8.1.8.

8.1.10. The eigenvalues for the problem $(xu'(x))' + (\lambda/x)u(x) = 0$, $1 < x < 2$, $u(1) = u(2) = 0$ (see Exercise 4.3.3), are given as $\lambda_n = (\pi n/\log 2)^2$, $n = 1, 2, \ldots$. With $p(x) = x$, $q(x) = 0$, $\rho(x) = 1/x$, determine the constants p_m, p_M, ρ_m, and ρ_M for the interval $1 < x < 2$, where the subscripts m and M correspond to the minima and maxima of these functions. Obtain the eigenvalues for the appropriate constant coefficient problems obtained from this problem and verify the result (8.1.46) for the eigenvalues of all three problems.

8.1.11. Obtain upper and lower bounds for the eigenvalues for the following problem: $-\nabla^2 M(x, y) + xyM(x, y) = \lambda M(x, y)$, $0 < x < \pi$, $0 < y < \pi$, $M(0, y) = M(\pi, y) = M(x, 0) = M(x, \pi) = 0$.

8.1.12. Use the method presented in the subsection that deals with Dirichlet eigenvalue problems for elliptic equations with constant coefficients to estimate the lowest eigenvalue for the following problem: $\nabla^2 M(x, y) + \lambda M(x, y) = 0$, $x^2/a^2 + y^2/b^2 < 1$, $a < b$, $M(x, y) = 0$, $x^2/a^2 + y^2/b^2 = 1$.

8.1.13. Obtain an upper and lower bound for the lowest eigenvalue of the following problem in a sphere: $\nabla^2 M(x,y,z) + \lambda M(x,y,z) = 0$, $x^2 + y^2 + z^2 < 1$, $M(x,y,z) = 0$, $x^2 + y^2 + z^2 = 1$, Proceed as in Exercise 8.1.12. The exact eigenvalue will be found in (8.2.18) to equal π^2.

8.1.14. Use the energy integral (8.1.10) to show that the leading eigenvalue λ_1 for the problem (8.1.1)–(8.1.2) increases as α/β increases. (Using the Courant max-min principle, it can be shown that all the eigenvalues λ_n have this property.)

8.1.15. Use separation of variables in polar coordinate to determine the eigenvalues and eigenfunctions for the following problem. $-\nabla^2 M(r,\theta) = \lambda M(r,\theta)$, $0 \le r < a$, $0 < \theta < \phi$, $[\phi \le 2\pi]$, with $M(r,\theta) = 0$ on the boundary. The eigenvalues are the squares of zeros of Bessel functions.

8.1.16. Use the results of Exercise 8.1.15 and the method of the text to obtain an upper and lower bound for the leading eigenvalue for the following problem. $M_{xx}(x,y) + M_{yy}(x,y) + \lambda M(x,y) = 0$, $(x,y) \in G$, with $M(x,y) = 0$ on ∂G, where the G is the interior of the triangle bounded by the lines $y = 0$, $y = x$ and $x = 1$.

8.1.17. Show that the (exact) eigenvalues and eigenfunctions for the problem of Exercise 8.1.16 are given as $\lambda_{nm} = \pi^2(n^2+m^2)$, $n < m$, $n = 1, 2, \ldots$, $m = 2, \ldots$, $M_{nm}(x,y) = \sin(\pi m x)\sin(\pi n y) - \sin(\pi n x)\sin(\pi m y)$. Verify that the bounds on the leading eigenvalue obtained in Exercise 8.1.16 are correct.

8.1.18. Show that the energy integral $E(w(\mathbf{x}))$ [as defined in (8.1.29)] for functions $w(\mathbf{x})$ that equal $f(\mathbf{x})$ on ∂G is minimized by the solution $u(\mathbf{x})$ of the Dirichlet problem $-\nabla\cdot(p(\mathbf{x})\nabla u(\mathbf{x})) + q(\mathbf{x})u(\mathbf{x}) = 0$, $\mathbf{x} \in G$; $u(\mathbf{x}) = f(\mathbf{x})$, $\mathbf{x} \in \partial G$. *Hint*: Let $w(\mathbf{x}) = u(\mathbf{x}) + W(\mathbf{x})$ where $u(\mathbf{x})$ is the (assumed) solution of the problem and $W(\mathbf{x})$ is an arbitrary function that vanishes on G. Show that $E(w(\mathbf{x})) = E(u(\mathbf{x})) + E(W(\mathbf{x})) > E(u(\mathbf{x}))$, if $W(\mathbf{x}) \neq 0$. If $p(\mathbf{x}) = 1$ and $q(\mathbf{x}) = 0$, this is a classical result known as *Dirichlet's principle*.

8.1.19. Adapt the procedure given in the text to show that the solution of the problem $E(w(\mathbf{x})) - \iint_G 2w(\mathbf{x})F(\mathbf{x})\,dv$ = minimum, where $E(w(\mathbf{x}))$ as defined in (8.1.29) or (8.1.30) is a solution of the problem $-\nabla\cdot(p(\mathbf{x})\nabla u(\mathbf{x})) + q(\mathbf{x})u(\mathbf{x}) = F(\mathbf{x})$, with homogeneous boundary conditions of the first, second, third, or mixed kind. In the case of Dirichlet boundary conditions, the admissible functions for the variational problem must vanish on ∂G for the Dirichlet problem or on the subset S_1 of ∂G in the mixed problem. The boundary conditions of the second and third kind are natural and no restriction must be placed on the admissible functions in that case. In the latter cases, the boundary condition that the solution $u(\mathbf{x})$ satisfies depends on the choice of $E(w(\mathbf{x}))$. *Hint*: Let $w(\mathbf{x}) = u(\mathbf{x}) + \epsilon W(\mathbf{x})$, where $u(\mathbf{x})$ is the (assumed) solution of the minimum problem.

8.1.20. Show that the solution of the problem $E(w(\mathbf{x})) - \iint_G 2w(\mathbf{x})F(\mathbf{x})\,dv - \int_{\partial G} 2w(\mathbf{x})p(\mathbf{x})f(\mathbf{x})\,ds$ = minimum, yields the solutions of the boundary value problems $-\nabla\cdot(p(\mathbf{x})\nabla u(\mathbf{x})) + q(\mathbf{x})u(\mathbf{x}) = F(\mathbf{x})$, $\mathbf{x} \in G$ with the boundary conditions $\partial u(\mathbf{x})/\partial n = f(\mathbf{x})$ or $\partial u(\mathbf{x})/\partial n + h(\mathbf{x})u(\mathbf{x}) = f(\mathbf{x})$, $\mathbf{x} \in \partial G$, depending on whether $E(w(\mathbf{x}))$ has the form (8.1.29) or (8.1.30), respectively. The boundary conditions

for the minimum problem are natural. To obtain the solution of the inhomogeneous equation with the Dirichlet boundary condition $u(\mathbf{x}) = f(\mathbf{x})$ on ∂G, we must proceed as in Exercise 8.1.19 except that all admissible functions must equal $f(\mathbf{x})$ on ∂G. Adapt the foregoing results to obtain the appropriate minimum problem in the case of mixed boundary conditions. *Hint*: Let $w(\mathbf{x}) = u(\mathbf{x}) + \epsilon W(\mathbf{x})$, where $u(\mathbf{x})$ is the (assumed) solution of the problem.

8.2 THE RAYLEIGH-RITZ METHOD

The formulation of the eigenvalue problem in variational form has led to some interesting and useful conclusions concerning the general properties of eigenvalues and eigenfunctions, as we have seen. For the most part, we have relied on separation of variables or exact solutions (in the one-dimensional case) to determine the eigenvalues and eigenfunctions for a given problem. The *Rayleigh-Ritz method* is an elegant and useful approximation method for the determination of the first few eigenvalues and eigenfunctions. In fact, it is most effective in approximating the lowest eigenvalue, and this is often the most important one in applications. As we shall see, the method yields an algebraic problem for the determination of the eigenvalues and eigenfunctions.

The *Rayleigh-Ritz method* can also be used to obtain approximate solutions of the variational form of *self-adjoint elliptic boundary value problems*. For *nonself-adjoint elliptic boundary value problems*, the *Galerkin method* can be applied to the Galerkin form of these problems (see Exercise 6.4.8) to yield approximate solutions. These matters are considered in the exercises.

To begin, we select n (basis) functions $\phi_1(\mathbf{x}), \ldots, \phi_n(\mathbf{x})$ that are admissible for the variational problem under consideration. That is, they must be sufficiently smooth and must vanish on ∂G in the case of Dirichlet boundary conditions. For the case of Neumann or Robin boundary conditions, they need not satisfy any conditions on ∂G. These functions are chosen to be linearly independent and are selected, if possible, to be good approximations to the eigenfunctions for the problem.

We form the linear combination

$$w(\mathbf{x}) = \sum_{k=1}^{n} c_k \phi_k(\mathbf{x}) \tag{8.2.1}$$

with as yet undetermined coefficients c_k and insert this sum into the Rayleigh quotient $E(w(\mathbf{x}))/||w(\mathbf{x})||^2$. This yields

$$\frac{E(w(\mathbf{x}))}{||w(\mathbf{x})||^2} = \frac{\sum_{j,k=1}^{n} c_j c_k \left\{ \iint_G [p \nabla \phi_j \cdot \nabla \phi_k + q \phi_j \phi_k]\, dv + \int_{S_3} hp \phi_j \phi_k \, ds \right\}}{\sum_{j,k=1}^{n} c_j c_k \iint_G \rho \phi_j \phi_k \, dv}. \tag{8.2.2}$$

If there are only Dirichlet or Neumann conditions on ∂G, S_3 is an empty set and there is no integral over S_3 in (8.2.2). If there are only Robin boundary conditions, the set S_3 coincides with ∂G. Otherwise, S_3 is a subset of ∂G.

Let the $n \times n$ matrices $A = (a_{jk})$ and $B = (b_{jk})$ have the respective elements

$$a_{jk} = \iint_G [p\nabla\phi_j(\mathbf{x}) \cdot \nabla\phi_k(\mathbf{x}) + q\phi_j(\mathbf{x})\phi_k(\mathbf{x})]\, dv + \int_{S_3} hp\phi_j(\mathbf{x})\phi_k(\mathbf{x})\, ds, \quad (8.2.3)$$

$$b_{jk} = \iint_G \rho\phi_j(\mathbf{x})\phi_k(\mathbf{x})\, dv, \quad (8.2.4)$$

where p, q, h, and ρ are functions of $\mathbf{x}$. We introduce the n-component column vector $\mathbf{c}$ with the components $c_1, c_2, \ldots, c_n$, with $\mathbf{c}^T$ as its transpose. Then (8.2.2) can be expressed as

$$\frac{E(w(\mathbf{x}))}{||w(\mathbf{x})||^2} = \frac{\mathbf{c}^T A\mathbf{c}}{\mathbf{c}^T B\mathbf{c}}, \quad (8.2.5)$$

as is easily verified. We remark that both $\mathbf{c}^T A\mathbf{c}$ and $\mathbf{c}^T B\mathbf{c}$ are *quadratic forms* in the c_i. Further, the (real-valued) matrices A and B are *symmetric*, as follows from (8.2.3)–(8.2.4). The matrix A is positive definite if $q(\mathbf{x}) > 0$, whereas B is always positive definite, as shown in the exercises. This means that if $q(\mathbf{x}) > 0$, $\mathbf{c}^T A\mathbf{c}$ and $\mathbf{c}^T B\mathbf{c}$ are positive scalars for any real vector $\mathbf{c} \neq \mathbf{0}$, and they vanish only if $\mathbf{c} = \mathbf{0}$. If $q(\mathbf{x}) = 0$, the matrix A is, at least, positive semi-definite. That is, it can vanish for $\mathbf{c} \neq \mathbf{0}$ if we take the $\phi_k(\mathbf{x})$ to be constants and a Neumann boundary condition is prescribed on all of ∂G.

The Rayleigh quotient $E(w(\mathbf{x}))/||w(\mathbf{x})||^2$ is now a function of the vector $\mathbf{c}$ and we wish to minimize this expression. To proceed with the Rayleigh-Ritz method, we assume that $\hat{\mathbf{c}} = \boldsymbol{\sigma}$ is the vector that minimizes the expression (8.2.5) and represent the arbitrary vector $\mathbf{c}$ in the form $\mathbf{c} = \boldsymbol{\sigma} + \sum_{k=1}^n \epsilon_k \mathbf{e}_k$. The vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$ are the standard basis vectors for n-component vectors, so that $\mathbf{e}_k$ has 1 as its kth component and zero as its remaining components. The constants ϵ_k characterize the variation of $\mathbf{c}$ around the minimizing vector $\boldsymbol{\sigma}$. Inserting $\mathbf{c}$ into (8.2.5) gives

$$\frac{E(w(\mathbf{x}))}{||w(\mathbf{x})||^2} = \frac{\left[\boldsymbol{\sigma} + \sum_{k=1}^n \epsilon_k \mathbf{e}_k\right]^T A \left[\boldsymbol{\sigma} + \sum_{k=1}^n \epsilon_k \mathbf{e}_k\right]}{\left[\boldsymbol{\sigma} + \sum_{k=1}^n \epsilon_k \mathbf{e}_k\right]^T B \left[\boldsymbol{\sigma} + \sum_{k=1}^n \epsilon_k \mathbf{e}_k\right]}. \quad (8.2.6)$$

To determine the equation satisfied by $\boldsymbol{\sigma}$, we differentiate (8.2.6) with respect to ϵ_k $(k = 1, \ldots, n)$ and then equate all the ϵ_k and each derivative to zero, since the minimum occurs when all the $\epsilon_k = 0$. This yields the n equations (for $k = 1 \ldots, n$)

$$\frac{\partial}{\partial\epsilon_k} \frac{E(w(\mathbf{x}))}{||w(\mathbf{x})||^2}_{[\epsilon_1=0,\ldots,\epsilon_n=0]} = \frac{\mathbf{e}_k^T A\boldsymbol{\sigma} + \boldsymbol{\sigma}^T A\mathbf{e}_k}{\boldsymbol{\sigma}^T B\boldsymbol{\sigma}} - \frac{[\mathbf{e}_k^T B\boldsymbol{\sigma} + \boldsymbol{\sigma}^T B\mathbf{e}_k]\boldsymbol{\sigma}^T A\boldsymbol{\sigma}}{(\boldsymbol{\sigma}^T B\boldsymbol{\sigma})^2} = 0. \quad (8.2.7)$$

Since A and B are symmetric matrices and $\boldsymbol{\sigma}^T A\mathbf{e}_k$ and $\boldsymbol{\sigma}^T B\mathbf{e}_k$ are scalars, we have $(\boldsymbol{\sigma}^T A\mathbf{e}_k)^T = \mathbf{e}_k^T A^T\boldsymbol{\sigma} = \mathbf{e}_k^T A\boldsymbol{\sigma} = \boldsymbol{\sigma}^T A\mathbf{e}_k$, and a similar result with A replaced by B. Thus (8.2.7) reduces to

$$\mathbf{e}_k^T \left(A\boldsymbol{\sigma} - \frac{\boldsymbol{\sigma}^T A\boldsymbol{\sigma}}{\boldsymbol{\sigma}^T B\boldsymbol{\sigma}} B\boldsymbol{\sigma} \right) = 0, \qquad k = 1, \ldots, n, \quad (8.2.8)$$

and the vector in parenthesis in (8.2.8) is orthogonal to the n basis vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$, so that it must be the zero vector. This yields the eigenvalue problem

$$A\boldsymbol{\sigma} = \hat{\lambda} B\boldsymbol{\sigma}, \qquad \hat{\lambda} = \frac{\boldsymbol{\sigma}^T A\boldsymbol{\sigma}}{\boldsymbol{\sigma}^T B\boldsymbol{\sigma}}. \tag{8.2.9}$$

Now if the $\phi_1(\mathbf{x}), \ldots, \phi_n(\mathbf{x})$ form an orthonormal set, we find from (8.2.4) that $b_{jk} = \delta_{jk}$, the Kronecker delta. Then B is a unit matrix, and (8.2.9) reduces to the standard eigenvalue problem for the matrix A with $\boldsymbol{\sigma}$ as the eigenvector and $\hat{\lambda}$ as the eigenvalue. If B is not a unit matrix, we also refer to $\boldsymbol{\sigma}$ as an eigenvector of A with respect to the positive definite matrix B and to $\hat{\lambda}$ as the corresponding eigenvalue.

The characteristic equation for the determination of the eigenvalues of the matrix A is $\det[A - \hat{\lambda}B] = 0$. This is an nth-degree algebraic equation for $\hat{\lambda}$, and the fact that A and B are symmetric matrices implies that all the eigenvalues are real. Since the matrices A and B are positive definite if $q(\mathbf{x}) > 0$, the eigenvalues must be positive. [If $q(\mathbf{x}) = 0$, they are, at least, nonnegative.] Therefore, they can be ordered as $0 \leq \hat{\lambda}_1 \leq \hat{\lambda}_2 \leq \cdots \leq \hat{\lambda}_n$.

For each eigenvalue $\hat{\lambda}_k$ there is at least one independent eigenvector $\boldsymbol{\sigma}^{(k)}$, and there are n linearly independent eigenvectors $\{\boldsymbol{\sigma}^{(k)}\}$, $k = 1, \ldots, n$, even if the n eigenvalues are not all distinct. It is easy to show that eigenvectors corresponding to different eigenvalues are orthogonal with respect to the matrix B. In fact, we have from (8.2.9), $0 = \boldsymbol{\sigma}^{(k)T}[A - \hat{\lambda}_j B]\boldsymbol{\sigma}^{(j)} - \boldsymbol{\sigma}^{(j)T}[A - \hat{\lambda}_k B]\boldsymbol{\sigma}^{(k)} = \boldsymbol{\sigma}^{(k)T} A\boldsymbol{\sigma}^{(j)} - \boldsymbol{\sigma}^{(j)T} A\boldsymbol{\sigma}^{(k)} - \hat{\lambda}_j \boldsymbol{\sigma}^{(k)T} B\boldsymbol{\sigma}^{(j)} + \hat{\lambda}_k \boldsymbol{\sigma}^{(j)T} B\boldsymbol{\sigma}^{(k)} = (\hat{\lambda}_k - \hat{\lambda}_j)\left[\boldsymbol{\sigma}^{(k)T} B\boldsymbol{\sigma}^{(j)}\right]$, where the symmetry of A and B was used. Thus if $\hat{\lambda}_k \neq \hat{\lambda}_j$, we have $\boldsymbol{\sigma}^{(k)T} B\boldsymbol{\sigma}^{(j)} = 0$. If B is the identity matrix, the foregoing is equivalent to the dot product of the two vectors $\boldsymbol{\sigma}^{(k)}$ and $\boldsymbol{\sigma}^{(j)}$ and shows orthogonality. If B is not a unit matrix, we say that the vectors are orthogonal with respect to B. Using a Gram-Schmidt process (see Exercise 8.2.1) it is possible to orthonormalize the set of eigenvectors with respect to B and obtain $\boldsymbol{\sigma}^{(k)T} B\boldsymbol{\sigma}^{(j)} = \delta_{kj}$, $k, j = 1, \ldots, n$, where δ_{kj} is the Kronecker delta.

Each *eigenvector* $\boldsymbol{\sigma}^{(j)}$ determines an *approximate eigenfunction*, $\hat{M}_j(\mathbf{x}) = \sum_{k=1}^n c_k^{(j)} \phi_k(\mathbf{x})$, $j = 1, \ldots, n$, where the $c_k^{(j)}$ are the components of the vector $\boldsymbol{\sigma}^{(j)}$. We observe that the inner product $(\hat{M}_j(\mathbf{x}), \hat{M}_k(\mathbf{x}))$ of the set of approximate eigenfunctions $\hat{M}_j(\mathbf{x})$ is $(\hat{M}_j(\mathbf{x}), \hat{M}_k(\mathbf{x})) = \iint_G \rho(\mathbf{x}) \hat{M}_j(\mathbf{x}) \hat{M}_k(\mathbf{x})\, dv = \sum_{l,m=1}^n c_l^{(j)} c_m^{(j)} b_{lm} = \boldsymbol{\sigma}^{(k)T} B\boldsymbol{\sigma}^{(j)}$. Thus approximate eigenfunctions corresponding to different approximate eigenvalues are *orthogonal*. Furthermore, if the $\{\boldsymbol{\sigma}^{(j)}\}$ are orthonormalized with respect to B, so are the $\{\hat{M}_j(\mathbf{x})\}$ with respect to the inner product above and its (induced) mean square norm.

The foregoing yields a basis for the validity of the Rayleigh-Ritz method. If the set of basis functions $\{\phi_1(\mathbf{x}), \ldots, \phi_n(\mathbf{x})\}$ is a subset of a complete set of functions $\{\phi_k(\mathbf{x})\}$, $k = 1, 2, \ldots$, any admissible function $w(\mathbf{x})$ (including the exact eigenfunctions) can be expanded in an infinite series of the form $w(\mathbf{x}) = \sum_{k=1}^\infty c_k \phi_k(\mathbf{x})$. The original minimum problem is then replaced by a problem where the coefficients c_k in the full expansion must be determined sp as to minimize the Rayleigh quotient.

Proceeding as above gives an infinite system of coupled equations for the c_k. There must be an infinite number of solutions of the system that yield all the eigenvalues and eigenfunctions, because each eigenfunction can be expanded in terms of the $\{\phi_k(\mathbf{x})\}$. The Rayleigh-Ritz method replaces the infinite sum by the finite sum (8.2.1) and obtains n approximate eigenvalues and eigenfunctions for the problem. The approximate eigenfunctions are mutually orthogonal, as are the exact eigenfunctions. Increasing the number of basis functions yields improved results.

The *Courant maximum-minimum principle* enables us to compare the approximate eigenvalues $\hat{\lambda}_k$ with the exact eigenvalues λ_k. In replacing the given variational problem with the Rayleigh-Ritz formulation, we are placing an additional constraint on the problem by requiring all admissible functions to be linear combinations of the functions $\phi_1(\mathbf{x}), \ldots, \phi_n(\mathbf{x})$. Thus every eigenvalue obtained by means of the Rayleigh-Ritz process cannot be smaller than the exact eigenvalue, so that

$$\lambda_k \leq \hat{\lambda}_k, \qquad k = 1, \ldots, n. \tag{8.2.10}$$

Application of the Rayleigh-Ritz Method

In Example 8.1 a single admissible function $u(x, y)$ was used to approximate the lowest eigenvalue and eigenfunction for the Dirichlet problem for the Laplacian in the rectangle. This amounts to using the Rayleigh-Ritz method with a single function $\phi_1(x, y)$ that vanishes on the rectangle. Example 8.3 deals with the Dirichlet problem for the Laplacian in the unit sphere.

Example 8.3. The Dirichlet Eigenvalue Problem for the Laplacian in a Sphere. We consider the eigenvalue problem

$$\nabla^2 M(x, y, z) + \lambda M(x, y, z) = 0, \quad (x, y, z) \in G, \tag{8.2.11}$$

with the Dirichlet boundary condition $M(x, y, z)\big|_{\partial G} = 0$, where G and ∂G are the interior and surface of a unit sphere centered at the origin. Introducing spherical coordinates $(r,\ \theta,\ \phi)$ with $0 \leq r < \infty,\ 0 \leq \theta \leq 2\pi, 0 \leq \phi \leq \pi$, gives

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial M}{\partial r}\right) + \frac{1}{r^2\sin(\phi)}\frac{\partial}{\partial \phi}\left[\sin(\phi)\frac{\partial M}{\partial \phi}\right] + \frac{1}{\sin^2(\phi)}\frac{\partial^2 M}{\partial \theta^2} + \lambda M = 0, \tag{8.2.12}$$

where $M = M(r,\ \theta,\ \phi)$. The boundary condition becomes $M(1,\ \theta,\ \phi) = 0$.

Using separation of variables, we set $M(r,\ \theta,\ \phi) = F(r)Y(\theta, \phi)$, and insert this expression into (8.2.12) to obtain after some simplification

$$\frac{(r^2F'(r))'}{F(r)} + \lambda r^2 = \frac{-1}{Y(\theta,\phi)}\left[\frac{1}{\sin(\phi)}\frac{\partial}{\partial \phi}\left(\sin(\phi)\frac{\partial Y(\theta,\phi)}{\partial \phi}\right) + \frac{1}{\sin^2(\phi)}\frac{\partial^2 Y(\theta,\phi)}{\partial \theta^2}\right] \tag{8.2.13}$$

The left and right sides of (8.2.13) depend on different variables, so they must both be constant. With k^2 as the separation constant, $Y(\theta, \phi)$ satisfies

$$\frac{1}{\sin(\phi)} \frac{\partial}{\partial \phi} \left(\sin(\phi) \frac{\partial Y(\theta, \phi)}{\partial \phi} \right) + \frac{1}{\sin^2(\phi)} \frac{\partial^2 Y(\theta, \phi)}{\partial \theta^2} + k^2 Y(\theta, \phi) = 0. \quad (8.2.14)$$

We require that $Y(\theta, \phi)$ be periodic in θ [i.e., $Y(\theta + 2\pi, \phi) = Y(\theta, \phi)$] and that it be bounded at $\phi = 0$ and $\phi = \pi$, as the coefficients in (8.2.14) are singular there. These conditions determine the values of k^2 to be $k^2 = n(n+1)$, $n = 0, 1, 2, \ldots$, and yield the solutions $Y_n(\theta, \phi)$ known as the *spherical harmonics*. A further separation of variables in (8.2.14) determines them to be products of Legendre functions and sine and cosine functions (see the exercises).

The boundary value problem for $F(r)$ becomes $(1/r^2)d/dr\left[r^2 dF(r)/dr\right] + \left[\lambda - n(n+1)/r^2\right] F(r) = 0$, with $F(r)$ bounded at $r = 0$ and $F(1) = 0$. With $F(r) = f(r)/\sqrt{r}$, the ODE reduces to the Bessel equation of order $n + \frac{1}{2}$,

$$f''(r) + \frac{1}{r} f'(r) + \left[\lambda - \frac{(n+1/2)^2}{r^2} \right] f(r) = 0. \quad (8.2.15)$$

The boundedness condition on $F(r)$ at $r = 0$ implies that (up to a constant multiple) $f(r) = J_{n+1/2}(\sqrt{\lambda}\, r)$ and the boundary condition at $r = 1$ requires that $J_{n+1/2}(\sqrt{\lambda}) = 0$.

If α_{mn} $(m = 1, 2, \ldots)$ are the roots of the Bessel function $J_{n+1/2}(\alpha)$, we obtain the eigenvalues

$$\lambda_{mn} = (\alpha_{mn})^2, \qquad n = 0, 1, 2, \ldots, \quad m = 1, 2, \ldots. \quad (8.2.16)$$

The roots of the Bessel functions are all real and for each n are listed in increasing order. It is well known that $J_{1/2}(x) = \sqrt{2/\pi x} \sin(x)$ and, in fact, all the Bessel functions $J_{n+1/2}(x)$ can be expressed in terms of powers of x and the sine and cosine functions.

We have shown that the eigenvalues for the sphere are given by (8.2.16) and the associated eigenfunctions are

$$\hat{M}_{mn} = \frac{1}{\sqrt{r}} J_{n+1/2}(\alpha_{mn} r) Y_n(\theta, \phi). \quad (8.2.17)$$

(The $\hat{M}_{mn}$ are not normalized.) The lowest (exact) eigenvalue is given by $\lambda_{10} = \alpha_{10}^2$, where α_{10} is the smallest positive zero of the Bessel function $J_{1/2}(x)$. In view of the above, we have

$$\lambda_{10} = \alpha_{10}^2 = \pi^2. \quad (8.2.18)$$

To approximate the eigenvalue λ_{10} using the *Rayleigh-Ritz method*, we introduce the function $\phi_1(r, \theta, \phi) = 1 - r$, which appears to be the simplest function that vanishes at $r = 1$ and is bounded at $r = 0$. Then $w(r, \theta, \phi) = c_1\, \phi_1(r, \theta, \phi)$, is an

admissible function for the variational problem associated with the Dirichlet problem for (8.2.11). The *Rayleigh quotient* is

$$\frac{E(w)}{||w||^2} = \frac{c_1^2 \int_0^{\pi} \int_0^{2\pi} \int_0^1 (\nabla \phi_1(r,\theta,\phi))^2 \, r^2 \sin(\phi) \, dr \, d\theta \, d\phi}{c_1^2 \int_0^{\pi} \int_0^{2\pi} \int_0^1 \phi_1^2(r,\theta,\phi) \, r^2 \sin(\phi) \, dr \, d\theta \, d\phi} = 10. \qquad (8.2.19)$$

Since the exact eigenvalue is $\pi^2 \approx 9.87$, the approximate eigenvalue $\lambda_{10} = 10$ given in (8.2.19) yields excellent agreement. We do not discuss how well the first eigenfunction is approximated nor approximations for higher eigenvalues and eigenfunctions.

Diffusion Process with a Chain Reaction

In the foregoing examples we have been concerned primarily with the determination of the leading or smallest eigenvalue in the problems considered. In a number of applications, a precise determination of this eigenvalue is of the greatest interest, as is the case in the following problem.

Let $u(x, y, z, t)$ be the *concentration* of some diffusing substance, and let the strength of the sources in the substance be proportional to the concentration. According to our discussion in Section 4.1, the equation for u is given as

$$u_t(x, y, z, t) = c^2 \, \nabla^2 u(x, y, z, t) + \gamma u(x, y, z, t), \qquad (8.2.20)$$

where c^2 and γ are positive constants, both of which are characteristic of the substance. Note that $-\gamma$, which corresponds to q in (4.1.4), is now negative rather than nonnegative. An initial and boundary value problem of interest for (8.2.20) occurs if we consider the *diffusion process* within a bounded region G, with the initial and boundary conditions

$$u(x, y, z, 0) = f(x, y, z), \quad u(x, y, z, t)|_{\partial G} = 0. \qquad (8.2.21)$$

Applying the separation of variables method to problem (8.2.20)–(8.2.21), we put $u(x, y, z, t) = N(t)M(x, y, z)$ and insert it into (8.2.20). This leads to

$$-\frac{\nabla^2 M(x, y, z)}{M(x, y, z)} = \frac{-N'(t) + \gamma N(t)}{c^2 N(t)} = \lambda, \qquad (8.2.22)$$

where λ is the separation constant. Thus, we consider the eigenvalue problem

$$\nabla^2 M(x, y, z) + \lambda M(x, y, z) = 0, \; (x, y, z) \in G, \quad M(x, y, z)|_{\partial G} = 0. \quad (8.2.23)$$

The eigenvalues λ_k $(k = 1, 2, \ldots)$ are positive and can be arranged as

$$0 < \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \cdots \leq \lambda_k \leq \cdots, \qquad (8.2.24)$$

as shown previously. We denote the associated eigenfunctions by $M_k(x, y, z)$. For each λ_k we obtain a solution $N_k(t)$ of the separated equation for $N(t)$ in the form

$N_k(t) = a_k \exp[(\gamma - \lambda_k c^2)t]$, $k = 1, 2, \ldots$, with arbitrary constants a_k. Finally, the solution of (8.2.20)–(8.2.21) is given as

$$u(x, y, z, t) = \sum_{k=1}^{\infty} a_k M_k(x, y, z) \exp[(\gamma - \lambda_k c^2)t], \tag{8.2.25}$$

with the a_k given as the Fourier coefficients in the eigenfunction expansion of the initial concentration $f(x, y, z)$.

An important question regarding the solution (8.2.25) is whether the concentration $u(x, y, z, t)$ grows without bound as the time t increases, in which case the problem (8.2.20)–(8.2.21) is *unstable*. Since the concentration at any time t is a result of reactions that occur within the substance as was indicated, we say that when the concentration $u(x, y, z, t)$ increases exponentially, a *chain reaction* is taking place because of the rapid increase in $u(x, y, z, t)$. The solution (8.2.25) shows that exponential growth in t occurs if $\gamma - \lambda_k c^2 > 0$. In view of (8.2.24), we see that the lowest eigenvalue λ_1 is the most significant, and if $\gamma/c^2 > \lambda_1$, a chain reaction takes place in the substance. Now the eigenvalue varies as the region G is changed, and the parameters c^2 and γ can presumably be adjusted for given substances. Therefore, if we fix two of the three constants, say, γ and c^2, we may define a critical value of the third constant, λ_1. If $\lambda_1 > \gamma/c^2$, no chain reaction takes place, whereas if $\lambda_1 < \gamma/c^2$, a chain reaction does occur. Thus the *critical value* of λ_1, which we denote by $\lambda_1^{(c)}$, is given as $\lambda_1^{(c)} = \gamma/c^2$.

For a specific problem in a given region it is important to determine the first eigenvalue (i.e., the lowest) to see if it is greater or less than the critical value. In the case of the unit sphere, we have found that the lowest eigenvalue equals π^2. The Rayleigh-Ritz method yielded a slightly higher value. It is clearly important to be able to approximate the lowest eigenvalue as accurately as possible. In Exercises 8.5.31 and 8.5.32 we consider a waveguide problem for which the determination of the leading eigenvalue is of critical importance.

Rayleigh–Ritz Method for Sturm–Liouville Problems

In the following examples we apply the Rayleigh-Ritz method to two Sturm-Liouville eigenvalue problems.

Example 8.4. A Sturm-Liouville Problem with Mixed Boundary Conditions. The eigenvalue problem with the mixed boundary conditions

$$M''(x) + \lambda M(x) = 0,\ 0 < x < \pi, \quad M'(0) = 0,\ M(\pi) = 0, \tag{8.2.26}$$

has the eigenvalues and the (unnormalized) eigenfunctions

$$\lambda_k = \left(k + \frac{1}{2}\right)^2, \quad \hat{M}_k(x) = \cos\left[\left(k + \frac{1}{2}\right)x\right], \quad k = 0, 1, 2, \ldots. \tag{8.2.27}$$

We use the *Rayleigh-Ritz method* to approximate the two lowest eigenvalues.

The appropriate Rayleigh quotient for this problem is

$$\frac{E(w(x))}{||w(x)||^2} = \frac{\int_0^\pi (w'(x))^2\,dx}{\int_0^\pi w^2(x)\,dx}. \tag{8.2.28}$$

The admissible functions $w(x)$ are required to vanish at $x = \pi$ [i.e., $w(\pi) = 0$]. At $x = 0$ there is a natural boundary condition so that no restrictions need be placed on $w(x)$ at that point.

To apply the Rayleigh-Ritz method, we select two functions $\phi_1(x)$ and $\phi_2(x)$ that satisfy the admissibility conditions; that is, $\phi_1(\pi) = \phi_2(\pi) = 0$. As in (8.2.1), we set

$$w(x) = c_1\phi_1(x) + c_2\phi_2(x), \tag{8.2.29}$$

with $\phi_1(x)$ and $\phi_2(x)$ chosen as

$$\phi_1(x) = \pi^2 - x^2, \qquad \phi_2(x) = \pi^3 - x^3. \tag{8.2.30}$$

We note that $\phi_1(\pi) = \phi_2(\pi) = 0$ and that, in addition, $\phi_1'(0) = \phi_2'(0) = 0$. Although admissible functions need not have vanishing derivatives at $x = 0$, we expect to get improved results with our choices for $\phi_1(x)$ and $\phi_2(x)$ since they are expected to more closely approximate the eigenfunctions (8.2.27), whose derivatives do vanish at $x = 0$.

Evaluating the integrals (8.2.3)–(8.2.4) with $p = \rho = 1$ and $q = 0$ and $\nabla\phi$ replaced by ϕ', we obtain the matrices

$$A = \begin{bmatrix} 4\pi^3/3 & 3\pi^4/2 \\ 3\pi^4/2 & 9\pi^5/5 \end{bmatrix}, \qquad B = \begin{bmatrix} 8\pi^5/15 & 7\pi^6/12 \\ 7\pi^6/12 & 9\pi^7/14 \end{bmatrix}. \tag{8.2.31}$$

The two roots of the characteristic equation $\det(A - \hat{\lambda}B) = 0$ are $\hat{\lambda}_1 = 0.25$, $\hat{\lambda}_2 = 2.39$. We have retained only two decimal places, and $\hat{\lambda}_1$ is, in fact, slightly larger than 0.25. Comparing the approximate eigenvalues $\hat{\lambda}_1$ and $\hat{\lambda}_2$ with the exact eigenvalues $\lambda_1 = 0.25$ and $\lambda_2 = 2.25$ [see (8.2.27)], we observe that there is excellent agreement between $\hat{\lambda}_1$ and λ_1 but that $\hat{\lambda}_2$ is not as close to λ_2. This is the norm for the Rayleigh-Ritz procedure, in that the higher eigenvalues are more poorly approximated than the lower ones. We do observe that we have $\lambda_1 < \hat{\lambda}_1$ and $\lambda_2 < \hat{\lambda}_2$, as required by (8.2.10).

Given the approximate eigenvalues $\hat{\lambda}_1$ and $\hat{\lambda}_2$ we can determine approximate eigenfunctions by specifying two sets of coefficients c_1 and c_2 in (8.2.29). We do not carry out this calculation.

Example 8.5. A Sturm-Liouville Problem with a Variable Coefficient. We consider the problem

$$M''(x) + (\lambda - \epsilon x)M(x) = 0,\ 0 < x < \pi, \quad M(0) = 0,\ M(\pi) = 0. \quad (8.2.32)$$

We assume that $0 < \epsilon \ll 1$, so that $q(x) = \epsilon x$ is uniformly small in the interval $0 \le x \le \pi$ and is nonnegative. With $\epsilon = 0$, (8.2.32) reduces to an exactly solvable problem. In any case, (8.2.32) can be solved exactly in terms of *Airy functions*, which are solutions of the equation $y''(x) - x\,y(x) = 0$. Nevertheless, we do not use the exact solutions of (8.2.32) but use approximate techniques to determine the lowest eigenvalue for the Sturm-Liouville problem (8.2.32).

First we apply the *Rayleigh-Ritz method.* In view of (8.1.75), the Rayleigh quotient appropriate for (8.2.32) is

$$\frac{E(w(x))}{||w(x)||^2} = \frac{\int_0^\pi [(w'(x))^2 + \epsilon x w^2(x)]\,dx}{\int_0^\pi w^2(x)\,dx}. \quad (8.2.33)$$

The admissible functions $w(x)$ are required to vanish at $x = 0$ and $x = \pi$. Since ϵ is small and for $\epsilon = 0$ the lowest eigenvalue for (8.2.32) is $\lambda_1 = 1$ and the corresponding eigenfunction is $M_1(x) = \sqrt{(2/\pi)}\sin(x)$, we pick $\phi_1(x)$ in the Rayleigh-Ritz approximation (8.2.1) to be $\phi_1(x) = \sin(x)$. Clearly, $\phi_1(x)$ is admissible for the variational problem.

With $w(x) = c_1\phi_1(x)$ we have as the leading approximate eigenvalue

$$\frac{E(w(x))}{||w(x)||^2} = \frac{\int_0^\pi [\cos^2(x) + \epsilon x \sin^2(x)]\,dx}{\int_0^\pi \sin^2(x)\,dx} = 1 + \frac{\epsilon\pi}{2} = \hat{\lambda}_1. \quad (8.2.34)$$

Now for $0 \le x \le \pi$, $q(x) = \epsilon x$ satisfies the inequalities $0 \le \epsilon x \le \epsilon\pi$, with $q_m = 0$ and $q_M = \epsilon\pi$ representing the minimum and maximum value of $q(x)$, respectively. Then (8.1.46) implies that $\lambda_1^{(m)} \le \lambda_1 \le \lambda_1^{(M)}$, where $\lambda_1^{(m)} = 1$ and $\lambda_1^{(M)} = 1 + \epsilon\pi$, as is easily seen on replacing $\lambda - \epsilon x$ by λ and by $\lambda - \epsilon\pi$ to determine $\lambda_1^{(m)}$ and $\lambda_1^{(M)}$, respectively. The Rayleigh-Ritz procedure implies that

$$\lambda_1 \le \hat{\lambda}_1 = 1 + \frac{\epsilon\pi}{2}, \quad (8.2.35)$$

and we have $\hat{\lambda}_1 < \lambda_1^{(M)}$. The smaller ϵ is, the better the approximation to λ_1.

Alternatively, a *perturbation method* may be applied to the problem (8.2.32). The solutions $M(x)$ and the eigenvalues λ are functions of ϵ; that is, $M(x) = M(x;\epsilon)$ and $\lambda = \lambda(\epsilon)$. Consequently, we expand both $M(x;\epsilon)$ and $\lambda(\epsilon)$ as

$$M(x;\,\epsilon) = M^{(0)}(x) + \epsilon M^{(1)}(x) + \cdots, \quad \lambda(\epsilon) = \lambda^{(0)} + \epsilon\,\lambda^{(1)} + \cdots. \quad (8.2.36)$$

We insert (8.2.36) into (8.2.32) to obtain

$$M''(x;\epsilon) + (\lambda - \epsilon x)M(x;\epsilon) = \left[\frac{d^2M^{(0)}(x)}{dx^2} + \lambda^{(0)}M^{(0)}(x)\right]$$
$$+\epsilon\left[\frac{d^2M^{(1)}(x)}{dx^2} + \lambda^{(0)}M^{(1)}(x) + (\lambda^{(1)} - x)M^{(0)}(x)\right] + \cdots = 0, \qquad (8.2.37)$$
$$M^{(0)}(0) + \epsilon M^{(1)}(0) + \cdots = M^{(0)}(\pi) + \epsilon M^{(1)}(\pi) + \cdots = 0. \qquad (8.2.38)$$

Equating coefficients of like powers of ϵ to zero in (8.2.37)–(8.2.38) yields

$$\frac{d^2M^{(0)}(x)}{dx^2} + \lambda^{(0)}M^{(0)}(x) = 0, \qquad M^{(0)}(0) = M^{(0)}(\pi) = 0 \qquad (8.2.39)$$

for $M^{(0)}(x)$. The leading eigenvalue and normalized eigenfunction for this problem are

$$\lambda_1^{(0)} = 1, \qquad M_1^{(0)}(x) = \sqrt{\frac{2}{\pi}}\sin(x). \qquad (8.2.40)$$

The problem for $M^{(1)}(x)$ is

$$\frac{d^2M^{(1)}(x)}{dx^2} + \lambda^{(0)}M^{(1)}(x) = -(\lambda^{(1)} - x)M^{(0)}(x), \; M^{(1)}(0) = M^{(1)}(\pi) = 0. \qquad (8.2.41)$$

With $\lambda^{(0)} = \lambda_1^{(0)} = 1$, the first eigenvalue for (8.2.39), the inhomogeneous ODE in (8.2.41) has no solution unless the inhomogeneous term is orthogonal to the eigenfunction $M_1^{(0)}(x)$. This follows on multiplying the ODE in (8.2.41) by $M_1^{(0)}(x)$ and integrating from 0 to π. We have

$$\int_0^\pi M_1^{(0)}\left[\frac{d^2M^{(1)}}{dx^2} + \lambda_1^{(0)}M^{(1)}\right]dx = \int_0^\pi M^{(1)}\left[\frac{d^2M_1^{(0)}}{dx^2} + \lambda_1^{(0)}M_1^{(0)}\right]dx = 0 \qquad (8.2.42)$$

on integrating by parts and using (8.2.39) and (8.2.41). This implies that $\int_0^\pi (\lambda^{(1)} - x)\sin^2(x)\,dx = 0$, from which we conclude that

$$\lambda^{(1)} = \lambda_1^{(1)} = \frac{\pi}{2}. \qquad (8.2.43)$$

Having determined $\lambda_1^{(1)}$, we can specify $M_1^{(1)}(x)$, but this is not carried out.

Inserting (8.2.40) and (8.2.43) in (8.2.36) gives

$$\lambda_1(\epsilon) \approx 1 + \frac{\epsilon\pi}{2}. \qquad (8.2.44)$$

The approximation (8.2.44) to the lowest eigenvalue is identical to that given by the Rayleigh-Ritz method. This appears to be due to our choice of $\phi_1(x)$ in the Rayleigh-Ritz method to equal the leading eigenfunction for (8.2.32) with $\epsilon = 0$. However, the perturbation method readily yields approximations to the higher eigenvalues, whereas this is not so easily accomplished by way of the Rayleigh-Ritz method, as we have seen (see also Section 9.2).

Exercises 8.2

8.2.1. *The Gram–Schmidt Process.* Let $(\phi(\mathbf{x}), \psi(\mathbf{x}))$ represent a (weighted) inner product for either vectors or scalars. Assume that the set of functions $\{\phi_k(\mathbf{x})\}$ $(k = 1, 2, \ldots)$ is linearly independent. Show that if $||\phi(\mathbf{x})||^2 = (\phi(\mathbf{x}), \phi(\mathbf{x}))$, the set of functions $\{\hat{\phi}_k(\mathbf{x})\}$ given as, $\hat{\phi}_1(\mathbf{x}) = \phi_1(\mathbf{x})/||\phi_1(\mathbf{x})||$, $\tilde{\phi}_2(\mathbf{x}) = \phi_2(\mathbf{x}) - (\phi_2(\mathbf{x}), \hat{\phi}_1(\mathbf{x}))\hat{\phi}_1(\mathbf{x})$, $\hat{\phi}_2(\mathbf{x}) = \tilde{\phi}_2(\mathbf{x})/||\tilde{\phi}_2(\mathbf{x})||$, $\tilde{\phi}_3(\mathbf{x}) = \phi_3(\mathbf{x}) - (\phi_3(\mathbf{x}), \hat{\phi}_1(\mathbf{x}))\hat{\phi}_1(\mathbf{x}) - (\phi_3(\mathbf{x}), \hat{\phi}_2(\mathbf{x}))\hat{\phi}_2(\mathbf{x})$, $\hat{\phi}_3(\mathbf{x}) = \tilde{\phi}_3(\mathbf{x})/||\tilde{\phi}_3(\mathbf{x})||$, $\tilde{\phi}_4(\mathbf{x}) = \phi_4(\mathbf{x}) - (\phi_4(\mathbf{x}), \hat{\phi}_1(\mathbf{x}))\hat{\phi}_1(\mathbf{x}) - (\phi_4(\mathbf{x}), \hat{\phi}_2(\mathbf{x}))\hat{\phi}_2(\mathbf{x}) - (\phi_4(\mathbf{x}), \hat{\phi}_3(\mathbf{x}))\hat{\phi}_3(\mathbf{x})$, and so on, is an orthonormal set.

8.2.2. Let $(\phi(x), \psi(x)) = \int_0^1 \phi(x)\psi(x)\,dx$ and put $\phi_1(x) = 1$, $\phi_2(x) = x$, $\phi_3(x) = x^2$, $\phi_4(x) = x^3$. Construct an orthonormal set $\hat{\phi}_1(x)$, $\hat{\phi}_2(x)$, $\hat{\phi}_3(x)$, and $\hat{\phi}_4(x)$ using the Gram-Schmidt process.

8.2.3. Given the three (independent) row vectors $\boldsymbol{\phi}_1^T = [1, 2, -1]$, $\boldsymbol{\phi}_2^T = [0, 1, 3]$, $\boldsymbol{\phi}_3^T = [-1, 1, 0]$, form an orthonormal set of vectors $\hat{\boldsymbol{\phi}}_1$, $\hat{\boldsymbol{\phi}}_2$, and $\hat{\boldsymbol{\phi}}_3$ from the preceding set using the Gram-Schmidt method and the dot product of the vectors as an inner product.

8.2.4. Conclude from the fact that $E(w(\mathbf{x}))$ with $q(\mathbf{x}) > 0$ and $||w(\mathbf{x})||$ are both positive if $w(\mathbf{x})$ vanishes on ∂G and is not identically zero that the matrices A and B defined by (8.2.3)–(8.2.4) must be positive definite in view of (8.2.5).

8.2.5. Separate variables in equation (8.2.14) for the spherical harmonics with $k^2 = n(n+1)$, and let $Y_n(\theta, \phi) = H(\phi)G(\theta)$ to obtain the equations

$$\frac{1}{\sin(\phi)}\frac{d}{d\phi}\left[\sin(\phi)\frac{dH(\phi)}{d\phi}\right] + \left[n(n+1) - \frac{\mu^2}{\sin^2(\phi)}\right]H(\phi) = 0,$$

and $G''(\theta) + \mu^2 G(\theta) = 0$, where μ^2 is the separation constant. Conclude that since $G(\theta)$ is periodic of period 2π, $\mu = m$ is an integer so that the $G(\theta)$ are trigonometric functions. Let $t = \cos\phi$ in the equation for $H(\phi)$ and obtain $d/dt\left[(1-t^2)d\hat{H}(t)/dt\right]$ $+ \left[n(n+1) - m^2/(1-t^2)\right]\hat{H}(t) = 0$ in the interval $-1 < t < 1$, where $\hat{H}(t) = \hat{H}(\cos\phi) = H(\phi)$. This is the *associated Legendre equation.* The solutions $H(\phi)$ must be bounded at $\phi = 0$ and $\phi = \pi$. Consequently, $\hat{H}(t)$ must be bounded at $t = \pm 1$. If $m = 0$, the bounded solutions are the *Legendre polynomials* $P_n(\cos\theta)$ discussed previously. If $m \neq 0$, bounded solutions are obtained only if $m \leq n$ and are given as the *associated Legendre functions* $P_n^m(\cos\phi)$. The functions $P_n^m(t)$ can be defined as $P_n^m(t) = (-1)^m(1-t^2)^{m/2}d^m/dt^m[P_n(t)]$ in terms of the Legendre polynomials $P_n(t)$.

8.2.6. Differentiate the Legendre equation $d/dt\left[(1-t^2)dP_n(t)/dt\right] + n(n+1)P_n(t)$ $= 0$, m times and show that $H(t) = (1-t^2)^{m/2}d^m P_n(t)/dt^m$ is a solution of the associated Legendre equation of Exercise 8.2.5. Determine that $P_n^m(t) = 0$ if $m > n$ from its definition in terms of $P_n(t)$, and show that $P_n^m(t)$ is bounded at $t = \pm 1$.

8.2.7. Determine the coefficients c_1 and c_2 corresponding to each eigenvalue $\hat{\lambda}_1$ and $\hat{\lambda}_2$ for the problem of Example 8.4, and show that the resulting functions are orthogonal.

8.2.8. Use the Rayleigh-Ritz method to approximate the leading eigenvalue for the problem $\nabla^2 M(x,y) + \lambda M(x,y) = 0,\ x^2 + y^2 < 1, \quad M(x,y) = 0,\ x^2 + y^2 = 1.$ Use the approximating function $\phi_1(r) = \cos(\pi r/2)$, where $r^2 = x^2 + y^2$. Compare the result with that given in Example 8.2.

8.2.9. Construct a function $\phi_1(x,y)$ that vanishes on the triangle given in Exercise 8.1.16 and use the Rayleigh-Ritz method to approximate the leading eigenvalue for the problem as given in Exercise 8.1.17.

8.2.10. Approximate the lowest eigenvalue for the problem in Example 8.4 by using only the function $\phi_1(x)$ as defined in (8.2.30). Compare the approximate result with that obtained in Example 8.4.

8.2.11. Using the appropriate Rayleigh quotient, approximate the first eigenvalue for the following problem: $M''(x) + \lambda M(x) = 0,\ 0 < x < 1,\ M(0) = 0,\ M'(1) + M(1) = 0$. Let $\phi_1(x) = 3x - 2x^2$.

8.2.12. Solve the problem of Example 8.5 for the leading eigenvalue by letting $w(x) = c_1\phi_1(x) + c_2\phi_2(x)$, with the basis functions $\phi_1(x) = \sin(x)$ and $\phi_2(x) = x\cos(x/2)$.

8.2.13. Explain why if the approximation functions $\phi_k(\mathbf{x})$ in (8.2.1) are chosen to be elements of a complete set $\{\phi_k(\mathbf{x})\}$, the more terms one takes in the series (8.2.1), the better the approximation to the leading eigenvalue and eigenfunction.

8.2.14. Use the Rayleigh-Ritz method to approximate the leading eigenvalue for the problem in Exercise 8.1.10. Let $\phi_1(x) = (x-1)(x-2)$.

8.2.15. Show how the Rayleigh-Ritz method can be applied to the variational problems given in Exercise 8.1.19. *Hint*: Express $w(\mathbf{x})$ as in (8.2.1) with $w(\mathbf{x})$ required to vanish on ∂G for the Dirichlet problem or on S_1 for the mixed problem, but arbitrary on ∂G or on a subset thereof where boundary conditions of the second or third kind are assigned.

8.2.16. Use the Rayleigh-Ritz method (see Exercise 8.2.15) to approximately solve the following Dirichlet problem for Poisson's equation within the square: $\nabla^2 u(x,y) = -\sin(x),\ 0 < x < \pi,\ 0 < y < \pi, \quad u(x,y) = 0$ on the boundary. Use a one-term Rayleigh-Ritz approximation.

8.2.17. Show how the Rayleigh-Ritz method can be applied to the variational problems of Exercise 8.1.20. *Hint*: See Exercise 8.2.15. For Dirichlet conditions, let $w(\mathbf{x}) = \phi_0(\mathbf{x}) + \sum_{k=1}^{n} c_k\phi_k(\mathbf{x})$, where $\phi_0(\mathbf{x}) = f(\mathbf{x})$ on ∂G or S_1 and the $\phi_k(\mathbf{x})$ vanish on ∂G or S_1. If boundary conditions of the second or third kind are given on all or part of ∂G, put $\phi_0(\mathbf{x}) = 0$ and let the $\phi_k(\mathbf{x})$ be unrestricted there.

8.2.18. Solve the Dirichlet problem $\nabla^2 u(x,y) = -1$, $x^2 + y^2 < 1$, $u(x,y) = x^2$, $x^2 + y^2 = 1$, using the Rayleigh-Ritz method. [See Exercise 8.2.17 and let $\phi_1(x,y) = x^2$ and $\phi_2(x,y) = \cos(\pi r/2)$ as in Exercise 8.2.8.]

8.2.19. Consider the nonself-adjoint elliptic equation and the integral relation of Galerkin form given in Exercise 6.4.8. Depending on the boundary conditions assigned for $u(\mathbf{x})$ on G, expand $u(\mathbf{x})$ in the same form as the function $w(\mathbf{x})$ in Exercise 8.2.17. Then choose n independent functions $v_1(\mathbf{x}), v_2(\mathbf{x}), \ldots, v_n(\mathbf{x})$ that are required to vanish on whatever part of ∂G that $u(\mathbf{x})$ itself is prescribed on, but are otherwise arbitrary. Insert the expansion of $u(\mathbf{x})$ into the Galerkin integral relation and successively set $v(\mathbf{x})$ equal to $v_1(\mathbf{x}), v_2(\mathbf{x}), \ldots, v_n(\mathbf{x})$. This yields n equations for the coefficients $c_1, c_2, \ldots, c_n$. This technique for constructing an approximate solution to the boundary value problem is known as the *Galerkin method.* (a) Determine the system of equations for the c_k. (b) Show that the system obtained in part (a) is identical to the system determined from the Rayleigh-Ritz method if $v_k(\mathbf{x}) = \phi_k(\mathbf{x})$, $k = 1, \ldots, n$ (see Exercise 8.2.17) and the vector $\mathbf{b}(\mathbf{x}) = \mathbf{0}$, so that the original equation is self-adjoint.

8.2.20. Use the Galerkin method to solve the Dirichlet problem $\nabla^2 u(x,y) - u_x(x,y) + u_y(x,y) = -1$, $x^2 + y^2 < 1$, where $u(x,y)$ satisfies the same boundary condition and is expanded as in Exercise 8.2.18, and where we set $v(x,y) = 1 - x^2 - y^2$.

8.3 RIEMANN'S METHOD

Riemann's method is a classical technique for solving the Cauchy problem for hyperbolic linear partial differential equations in two independent variables. Although this method does not yield closed-form solutions except in a limited number of cases, it does provide useful information about domains of dependence and influence for solutions in general. It was shown in Section 3.1 that second order hyperbolic equations in two variables can be brought into one of the two canonical forms (3.1.19) or (3.1.21), and the Riemann method is generally applied to one of these forms. We do not consider this method in its full generality, but we also do not use either of the two canonical forms for hyperbolic equations in order to show how to deal with curvilinear characteristics.

We consider the equation

$$u_{tt}(x,t) - \gamma^2(x)u_{xx}(x,t) + \alpha(x)u_x(x,t) + \beta(x)u_t(x,t) + c(x)u(x,t) = F(x,t), \tag{8.3.1}$$

with $\gamma(x) > 0$, given for $-\infty < x < \infty$, $t > 0$. The initial conditions given at $t = 0$ are

$$u(x,0) = f(x), \qquad u_t(x,0) = g(x), \qquad -\infty < x < \infty. \tag{8.3.2}$$

With the operator M defined as

$$Mu = u_{tt}(x,t) - \gamma^2(x)u_{xx}(x,t) + \alpha(x)u_x(x,t) + \beta(x)u_t(x,t) + c(x)u(x,t), \tag{8.3.3}$$

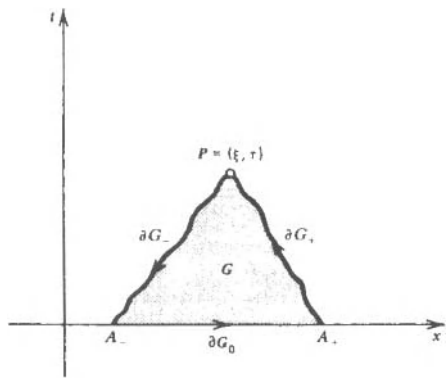

Figure 8.2 The characteristic triangle.

the adjoint operator M^* is given as

$$M^*w = w_{tt}(x,t) - (\gamma^2(x)w(x,t))_{xx} - (\alpha(x)w(x,t))_x - \beta(x)w_t(x,t) + c(x)w(x,t). \tag{8.3.4}$$

Introducing a bounded but as yet arbitrary region G in (x,t)-space with boundary ∂G, we obtain on using (3.6.5)–(3.6.6)

$$\iint_G \{wMu - uM^*w\}\, dx\, dt = \int_{\partial G} [-\gamma^2 w u_x + u(\gamma^2 w)_x + \alpha uw]\, dt + [uw_t - wu_t - \beta uw]\, dx, \tag{8.3.5}$$

with integration over ∂G carried out in the positive direction.

The characteristic curves for (8.3.1) are solutions of the equations $dt/dx = \mp 1/\gamma(x)$, so that they are given by

$$t \pm \int^x \frac{ds}{\gamma(s)} = \text{constant}. \tag{8.3.6}$$

Let $(\xi,\ \tau)$ be a point in the (x,t)-plane with $\xi > 0$. The two characteristic curves (8.3.6) that pass through that point are given as $t \pm \int_\xi^x ds/\gamma(s) = \tau$. If the curves are extended backward in t until they intersect the x-axis, we obtain a *characteristic triangle* whose base is a segment of the x-axis and whose other sides are characteristic curves. The interior of the triangle is denoted as G and the triangle itself by ∂G, as shown in Figure 8.2. The boundary ∂G is divided into three parts, ∂G_0, ∂G_+, and ∂G_-, and we apply (8.3.5) to this region. The figure displays the positive direction of integration.

On the line segment ∂G_0 the line integral in (8.3.5) becomes

$$\int_{\partial G_0} [uw_t - wu_t - \beta uw]\, dx = \int_{A_-}^{A_+} [f(x)w_t(x,0) - (g(x) + \beta(x)f(x))w(x,0)]\, dx, \tag{8.3.7}$$

since $dt = 0$, $u(x,0) = f(x)$, and $u_t(x,0) = g(x)$. On the characteristic segments $\partial G_\pm$ we have, since $\gamma(x) > 0$, $dt = \mp 1/\gamma(x)\,dx$, and for any function $v(x,t)$ defined and differentiable along the characteristics, $dv/dx = v_x + dt/dx\, v_t = v_x \mp 1/\gamma(x)\, v_t$. Thus

$$\int_{\partial G_\pm} [-\gamma^2 w u_x + u(\gamma^2 w)_x + \alpha u w]\,dt + [u w_t - w u_t - \beta u w]\,dx$$

$$= \int_{\partial G_\pm} \left\{ \pm \frac{d}{dx}(\gamma w u) \mp \gamma u \left[2\frac{dw}{dx} + 3\frac{\gamma'}{\gamma} w + \frac{1}{\gamma^2}\alpha w \pm \frac{1}{\gamma}\beta w \right] \right\} dx$$

$$= \gamma u w|_{A\pm}^{P} \mp \int_{\partial G_\pm} \gamma u \left[2\frac{dw}{dx} + 3\frac{\gamma'}{\gamma} w + \frac{1}{\gamma^2}\alpha w \pm \frac{1}{\gamma}\beta w \right] dx. \tag{8.3.8}$$

Collecting the results, we obtain

$$2\gamma u w|_P = \{\gamma f w|_{A_-} + \gamma f w|_{A_+}\} - \int_{A_-}^{A_+} [f w_t - (g + \beta f) w]\,dx$$
$$+ \int_{A_+}^{P} \gamma u \left[2\frac{dw}{dx} + 3\frac{\gamma'}{\gamma} w + \frac{1}{\gamma^2}\,\alpha w + \frac{1}{\gamma}\,\beta w \right] dx$$
$$+ \int_{A_-}^{P} \gamma u \left[2\frac{dw}{dx} + 3\frac{\gamma'}{\gamma} w + \frac{1}{\gamma^2}\alpha w - \frac{1}{\gamma}\beta w \right] dx + \iint_G (w M[u] - u M^*[w])\,dx\,dt. \tag{8.3.9}$$

To obtain an expression for the solution $u(x,t)$ at the point $P = (\xi, \tau)$, we require that $w(x,t)$ satisfy the following conditions. First we set

$$M^*[w(x,t)] = 0 \tag{8.3.10}$$

in the characteristic triangle G. To eliminate the integrals over the characteristics ∂G_+ and ∂G_-, we set

$$2\,\frac{dw}{dx} + 3\,\frac{\gamma'(x)}{\gamma(x)} w + \frac{1}{\gamma^2(x)}\,\alpha(x) w \pm \frac{1}{\gamma(x)}\,\beta(x) w = 0 \tag{8.3.11}$$

on these characteristics. Finally, to obtain an expression for $u(x,t)$ at P that is independent of the value of $w(x,t)$, we set

$$2\gamma(x) w(x,t)|_P = 2\gamma(\xi) w(\xi,\ \tau) = 1. \tag{8.3.12}$$

The equations (8.3.11)–(8.3.12) represent an initial value problem for $w(x,t)$ on the characteristic curves. The solution is easily found to be

$$w_\pm = \frac{1}{2}\left[\frac{\gamma(\xi)}{\gamma^3(x)}\right]^{1/2} \exp\left[-\frac{1}{2}\int_\xi^x \frac{\alpha(s) \pm \gamma(s)\beta(s)}{\gamma^2(s)}\,ds\right], \tag{8.3.13}$$

with $w_{\pm}$ evaluated on the characteristics $\partial G_{\pm}$. The equation (8.3.10) for w and the conditions (8.3.12)–(8.3.13) comprise a characteristic initial value problem for $w(x,t)$. (The problem is to be solved backward rather than forward in time.) Such a problem is well posed and has been discussed at the end of Section 6.5 where an iteration method for solving the characteristic initial value problem for the canonical form of second order hyperbolic equations was presented. We assume that a solution can be found for the present problem as well and obtain explicit solutions for two simple cases in Example 8.6.

The solution $w(x,t)$ is called the *Riemann function*. Since it depends on the point $P = (\xi, \tau)$, we express it as $w(x,t) = R(x,t;\xi,\tau)$. Then we have from (8.3.9), since $M[u] = F$,

$$u(\xi,\ \tau) = \{\gamma f R|_{A_-} + \gamma f R|_{A_+}\} + \iint_G FR\,dx\,dt - \int_{A_-}^{A_+} [fR_t - (g + \beta f)R]\,dx. \tag{8.3.14}$$

The solution formula (8.3.14) shows that the domain of dependence of the solution $u(x,t)$ of (8.3.1)–(8.3.2) is the interior of the backward characteristic triangle with vertex at the point (x,t). Reversing the argument shows that the domain of influence of a point $(x,0)$ is the interior of the forward characteristic sector formed by the two characteristics issuing from the point $(x,0)$.

To establish a connection between the *Riemann function*, the *Green's function*, and the *fundamental solution*, we set $F(x,t) = \delta(x - \hat{\xi})\delta(t - \hat{\tau})$ in (8.3.1) and (8.3.14) and assume that the initial data $f(x)$ and $g(x)$ vanish. Then (8.3.14) yields, with (ξ,τ) replaced by (x,t),

$$u(x,t) = \iint_G R(\tilde{x},\tilde{t};x,t)\delta(\tilde{x} - \hat{\xi})\delta(\tilde{t} - \hat{\tau})\,d\tilde{x}\,d\tilde{t} = \begin{cases} R(\hat{\xi},\hat{\tau};x,t), & (\hat{\xi},\hat{\tau}) \in G, \\ 0, & (\hat{\xi},\hat{\tau}) \notin G. \end{cases} \tag{8.3.15}$$

This shows that the Riemann function $R(\xi,\tau;x,t)$ represents the effect of a point source located at the point (ξ,τ). As shown in Figure 8.3, if the point (x,t) lies in the forward characteristic sector with vertex at (ξ,τ), the solution is nonzero. Otherwise, $u(x,t)$ vanishes, since disturbances cannot move faster than the characteristic speed $x'(t) = \gamma(x)$. In fact, the Riemann function $R(\xi,\tau;x,t)$ is equal to the causal fundamental solution for the operator Mu [see (8.3.3)] in the forward characteristic sector, that is, in the region where the causal fundamental solution is nonzero.

Similarly, the Riemann function $R(x,t;\xi,\tau)$ (now expressed in terms of the original variables) agrees with the free space Green's function for the operator Mu in the backward characteristic sector that issues from (ξ,τ) (see Figure 8.2). The causal fundamental solution and the free space Green's function both vanish outside the forward and backward characteristic sectors, respectively. Thus, each of these functions is completely determined once the Riemann function together with its domain of definition is known. Prior to the development of the theory of generalized functions, the importance of whose role in the construction of Green's functions and fundamental solutions was demonstrated in the preceding chapters, most efforts to

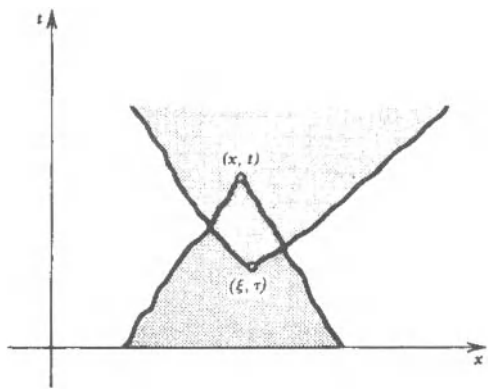

Figure 8.3 Forward and backward characteristic sectors.

solve Cauchy problems for hyperbolic equations were based on generalizations of Riemann's method.

Example 8.6. Riemann's Function for Equations with Constant Coefficients. With γ= constant and $\alpha = \beta = c = 0$ in (8.3.1), we obtain the inhomogeneous wave equation

$$u_{tt}(x,t) - \gamma^2 u_{xx}(x,t) = F(x,t). \tag{8.3.16}$$

In this case the operator M as given in (8.3.3) is self-adjoint. The Riemann function equals $1/2\gamma$ on the characteristics, and we immediately conclude that

$$R(x,t;\xi,\tau) = \frac{1}{2\gamma} \tag{8.3.17}$$

in the characteristic triangle. The *causal fundamental solution* for this problem has the form (6.7.45) and agrees with (8.3.15). It follows easily that (8.3.14) reduces to *d'Alembert's solution* of the initial value problem for (8.3.16).

If γ = constant, $\alpha = \beta = 0$, c = constant, with $c < 0$, we obtain for (8.3.1),

$$u_{tt}(x,t) - \gamma^2 u_{xx}(x,t) + cu(x,t) = F(x,t). \tag{8.3.18}$$

The operator M in (8.3.3) is self-adjoint and we have $R = 1/2\gamma$ on the characteristic curves. By looking for a solution of the equation for R that depends on the hyperbolic distance defined in (6.7.41) with c replaced by γ, we find as in Example 6.14 that the *Riemann function* R is given as

$$R(x,t;\xi,\tau) = \frac{1}{2\gamma} I_0\left[\sqrt{-\frac{c}{\gamma^2}}\sqrt{\gamma^2(t-\tau)^2 - (x-\xi)^2}\right], \tag{8.3.19}$$

since the modified Bessel function $I_0(z)$ has the property that $I_0(0) = 1$. The *causal fundamental solution* (8.3.15) agrees with (6.7.47), and (8.3.14) agrees with (7.4.43), as shown in the exercises.

Exercises 8.3

8.3.1. Determine the conditions under which equation (8.3.1) is self-adjoint and show that the Riemann function is symmetric in that case. That is, $R(x,t;\xi,\tau) = R(\xi,\tau;x,t)$.

8.3.2. Show that (8.3.14) reduces to d'Alembert's solution of the wave equation if $R(x,t;\xi,\tau)$ is given by (8.3.17), $\gamma(x)$ is a constant, and $\beta(x) = 0$.

8.3.3. Verify that (8.3.14) and (7.4.43) are identical if $R(x,t;\xi,\tau)$ is given by (8.3.19) and $c = -\hat{c}^2$.

8.3.4. Obtain the Riemann function for (8.3.18) with $c > 0$ (i.e., the Klein- Gordon equation). Show that the solution formula for this problem agrees with (7.4.41).

8.4 MAXIMUM AND MINIMUM PRINCIPLES

In Chapter 1, the diffusion, telegrapher's, and Laplace's equations were derived as limiting forms of certain random walk problems. It was indicated that some of the properties of the discrete random walk models carry over to the limiting PDEs. For example, a mean value property for the discrete form of Laplace's equation was shown in Section 7.5 to be valid for solutions of Laplace's equation. In this section we show that the maximum and minimum properties valid for discrete forms of the diffusion and Laplace equations carry over to their limiting forms. These properties can be used to prove uniqueness and continuous dependence on data of the solutions of these equations. In addition, we will show for the telegrapher's equation that if the initial data for the Cauchy problem are positive, the solution must be positive. This is consistent with the interpretation of the solution in Section 1.2 as a probability density function that must be nonnegative.

Maximum and Minimum Principles for the Diffusion Equation

We begin by presenting a maximum principle for the *diffusion* or *heat equation*

$$u_t(x,t) - c^2 u_{xx}(x,t) = 0 \tag{8.4.1}$$

in the closed rectangular region R given as $0 \le x \le l$ and $0 \le t \le T$. We assume that c^2 is a constant. [A more general equation of parabolic type with constant coefficients in the form

$$v_t(x,t) - c^2 v_{xx}(x,t) = a v_x(x,t) + b v(x,t) \tag{8.4.2}$$

can be reduced to (8.4.1) by way of the transformation $v(x,t) = \exp(\alpha x + \beta t)u(x,t)$ with $\alpha = -a/2c^2$ and $\beta = b - a^2/4c^2$.]

The *maximum and minimum principle* is given as follows. Let $u(x,t)$ be a solution of (8.4.1) in $0 < x < l$ and $0 < t < T$ that is continuous in the closed region R. Then the maximum and minimum values of $u(x,t)$ are assumed on the initial line $t = 0$ or at points on the boundary $x = 0$ and $x = l$.

In proving this principle we make use of a corresponding *maximum and minimum principle* for the *inhomogeneous diffusion equation*

$$u_t(x,t) - c^2 u_{xx}(x,t) = F(x,t) \tag{8.4.3}$$

given over the closed region R. Let $u(x,t)$ be a solution of (8.4.3) that is continuous in the closed region R. Then if $F(x,t) < 0$ in that region, $u(x,t)$ attains its maximum values on $t = 0$, $x = 0$, or $x = l$ and not in the interior of the region or at $t = T$. If $F(x,t) > 0$ in R, $u(x,t)$ attains its minimum values on $t = 0$, $x = 0$, or $x = l$ and not in the interior R or at $t = T$.

The proof of this principle for the inhomogeneous equation (8.4.3) is based on showing that if a maximum or minimum occurs at an (interior) point (x_0, t_0) with $0 < x_0 < l$ and $0 < t_0 \leq T$, we are led to a contradiction. We consider the case with $F(x,t) < 0$ first. Since $u(x,t)$ is continuous in the closed and bounded region R, it must assume its maximum there. At the assumed interior maximum point (x_0, t_0), we must have

$$u_t(x_0,t_0) \geq 0, \qquad u_{xx}(x_0,t_0) \leq 0, \tag{8.4.4}$$

as is known from calculus. If $t_0 < T$, then $u_t(x_0,t_0) = 0$, since both $u_x(x,t)$ and $u_t(x,t)$ vanish at an interior maximum point. However, it may happen that $t_0 = T$, in which case $(x_0,t_0) = (x_0,T)$ is on the boundary of the closed region and we can merely assert that $u_t(x_0,T) \geq 0$, since $u(x,t)$ may be increasing at that point.

If we insert the inequalities (8.4.4) into (8.4.3), we obtain a contradiction. The left side of the equation is nonnegative, whereas the right side is strictly negative. Consequently, the maximum must be assumed on the initial line or on the boundary. If $F(x,t) > 0$, we assume there is an interior minimum point (x_0,t_0) and obtain the inequalities (8.4.4) with the signs reversed. Again this leads to a contradiction, so that the minimum must be assumed on the initial line or on the boundary.

The inequalities (8.4.4) at a maximum or their reversed form at a minimum, do not yield a contradiction when they are inserted into (8.4.1), as $u_{xx}(x,t)$ and $u_t(x,t)$ *may both vanish at* (x_0,t_0). *To prove the maximum principle for (8.4.1) based on the* foregoing argument, we introduce the auxiliary function

$$w(x,t) = u(x,t) + \epsilon x^2, \tag{8.4.5}$$

where $\epsilon > 0$ is a constant and $u(x,t)$ satisfies (8.4.1). Now $w(x,t)$ is continuous in R, so that it has a maximum at some point (x_1,t_1) in the region. If we assume that $0 < x_1 < l$ and $0 < t_1 \leq T$, we again conclude that

$$w_t(x_1,t_1) \geq 0, \qquad w_{xx}(x_1,t_1) \leq 0. \tag{8.4.6}$$

But

$$w_t(x,t) - c^2 w_{xx}(x,t) = u_t(x,t) - c^2 u_{xx}(x,t) - 2c^2\epsilon = -2c^2\epsilon < 0, \quad (8.4.7)$$

since $u(x,t)$ satisfies (8.4.1). Inserting (8.4.6) into (8.4.7) now leads to a contradiction, since the left side is nonnegative and the right side is strictly negative. Therefore, $w(x,t)$ assumes its maximum on the initial line or on the boundary since $w(x,t)$ satisfies (8.4.3) with $F(x,t) < 0$.

Let M be the maximum value of $u(x,t)$ on $t = 0$, $x = 0$, and $x = l$ (i.e., on the initial and boundary lines). Then $w(x,t) = u(x,t) + \epsilon x^2 \leq M + \epsilon l^2$ in R since $w(x,t)$ has its maximum on $t = 0$, $x = 0$ or $x = l$. Consequently,

$$u(x,t) = w(x,t) - \epsilon x^2 \leq w(x,t) \leq M + \epsilon l^2. \quad (8.4.8)$$

Since ϵ is arbitrary, we can let $\epsilon \to 0$ in (8.4.8) and conclude that $u(x,t) \leq M$ throughout the closed region R, and the proof is complete.

To obtain a *minimum principle*, we consider $-u(x,t)$, where $u(x,t)$ is a solution of (8.4.1). Clearly, $-u(x,t)$ is also a solution of (8.4.1) and the maximum values of $-u(x,t)$ correspond to the minimum values of $u(x,t)$. Since $-u(x,t)$ satisfies the maximum principle, we conclude that $u(x,t)$ assumes its minimum values on the initial line or on the boundary lines. This implies that if the initial and boundary data for the problem are nonnegative, the solution must be nonnegative, a result consistent with the interpretation of the solution as a probability density, a concentration or a temperature.

For the *Cauchy problem* for (8.4.1) over the initial line $t = 0$, the solution $u(x,t)$ is nonnegative if $u(x,0) = f(x)$ is nonnegative. This follows directly from the representation of the solution given in Example 5.2. We have

$$u(x,t) = \frac{1}{\sqrt{4\pi c^2 t}} \int_{-\infty}^{\infty} \exp\left[-\frac{(x-s)^2}{4c^2 t}\right] f(s)\, ds. \quad (8.4.9)$$

The exponential term and the algebraic coefficient in (8.4.9) are positive, so that $u(x,t)$ is nonnegative as long as the initial value $f(x)$ is nonnegative.

On examining the proofs of the maximum and minimum principles for (8.4.1) and (8.4.3), we observe the following basic difference between the results. For the inhomogeneous equation (8.4.3) it was shown that the maximum or minimum values must be attained either on the initial line or the boundary lines and that they cannot be assumed in the interior. This result is known as a *strong maximum and minimum principle*. For the homogeneous equation (8.4.1) we determined that the maximum and minimum values are attained on the initial line or on the boundary lines. However, the possibility that the solution can also assume these maximum and minimum values at interior points was not excluded. (This would be the case, for example, if the solution is identically constant.) This result is referred to as a *weak maximum and minimum principle*. It can be proven (but we do not carry this out) that a strong maximum and minimum principle is valid for (8.4.1). It states that if the solution

does assume the maximum and minimum values in the interior, it must be identically constant.

We now show that if the solution $u(x,t)$ of (8.4.1) has a maximum value at a point on the boundary $x = 0$ or $x = l$, the normal derivative $\partial u(x,t)/\partial n$ is positive at that point unless $u(x,t)$ is identically constant. This means that $u_x(l,t) > 0$ and $u_x(0,t) < 0$. At a minimum point the normal derivative is negative. [We note that at $x = l$, say, we must have $u_x(x,t) \geq 0$ at a maximum point since the point lies on the boundary, but the theorem states that $u_x(x,t)$ is strictly positive there.]

A separate discussion is required at $x = 0$ and at $x = l$. We consider the boundary line $x = l$ and assume that $(x,t) = (l,t_0)$ is a maximum point for the solution $u(x,t)$ of (8.4.1) with $u(l,t_0) = M$. Let $0 < \alpha < t_0 < \beta < T$ and $0 < a < l$, and construct the rectangle $\hat{R}$ bounded by $x = a$, $x = l$, $t = \alpha$, and $t = \beta$. We assume that $u(l,t) < M$ for $t \neq t_0$ on the side of $\hat{R}$ that coincides with the boundary and we also have $u(x,t) < M$ on the three other sides of the rectangle because of the strong maximum principle. Given the auxiliary function

$$v(x,t) = u(x,t) + \epsilon(x^2 - 4lx + 3l^2), \tag{8.4.10}$$

where $\epsilon > 0$, we have $v(l,t) = u(l,t)$ and we choose ϵ to be sufficiently small so that $v(x,t) < M$ on the sides $x = a$ and $t = \alpha$ of the rectangle. [This is possible because $u(x,t)$ is strictly less than M on these sides.] Further,

$$v_t(x,t) - c^2 v_{xx}(x,t) = u_t(x,t) - c^2 u_{xx}(x,t) - 2c^2\epsilon = -2c^2\epsilon < 0. \tag{8.4.11}$$

Thus, the maximum principle for the inhomogeneous diffusion equation applied to the rectangle $\hat{R}$ implies that $v(x,t)$ has its maximum at the point (l,t_0). As a result, $v_x(l,t_0) \geq 0$, and this means that $u_x(l,t_0) \geq 2l\epsilon > 0$, as was to be shown. (This result is valid even if the maximum point is not isolated.) If $u(x,t)$ is a constant, the proof is clearly not valid. The proof that $u_x(x,t) < 0$ at a maximum point on $x = 0$ is left as an exercise.

The maximum and minimum principles may be used to prove *uniqueness* and *continuous dependence on the data* for solutions of initial and boundary value problems for the diffusion equation. They can also be extended to higher space dimensions and to equations with variable coefficients. These matters are considered in the exercises.

Maximum and Minimum Principle for Poisson's and Laplace's Equations

We begin with a *maximum and minimum principle* for *Poisson's equation*,

$$\nabla^2 u(x,y) = -F(x,y) \tag{8.4.12}$$

in two dimensions. Consider the bounded region G and its boundary ∂G. Then, the maximum values of a solution $u(x,y)$ of (8.4.12) are attained on ∂G if $F(x,y) < 0$ in G, and the minimum values of $u(x,y)$ are assumed on ∂G if $F(x,y) > 0$ in G.

To prove this result we note that since $u(x, y)$ is continuous in a closed and bounded region by assumption, it must assume its maximum in G or in ∂G. Let us suppose that the maximum is assumed at a point (x_0, y_0) in G and consider the case where $F(x, y) < 0$ in G. Then at the interior maximum point (x_0, y_0), we must have

$$u_{xx}(x_0, y_0) \leq 0, \qquad u_{yy}(x_0, y_0) \leq 0. \tag{8.4.13}$$

But since $F(x, y) < 0$, (8.4.12) yields

$$u_{xx}(x_0, y_0) + u_{yy}(x_0, y_0) > 0, \tag{8.4.14}$$

which contradicts (8.4.13). Thus, the maximum of $u(x, y)$ must occur on ∂G.

To show that the minimum of $u(x, y)$ is attained on ∂G if $F(x, y) > 0$ in G, we replace $u(x, y)$ by $-u(x, y)$ in the preceding argument. This is equivalent to replacing $F(x, y)$ by $-F(x, y)$ in (8.4.12). Since $F(x, y) > 0$, we obtain $-F(x, y) < 0$ and conclude that $-u(x, y)$ assumes its maximum on ∂G. Therefore, $u(x, y)$ assumes its minimum on ∂G and the minimum principle is proven.

Now the *Green's function* $K(x, y; \xi, \eta)$ for *Laplace's equation* with Dirichlet boundary conditions satisfies [see (1.3.12)–(1.3.13)]

$$K_{xx}(x, y; \xi, \eta) + K_{yy}(x, y; \xi, \eta) = -\delta(x-\xi)\delta(y-\eta), \ (x, y), (\xi, \eta) \in G, \tag{8.4.15}$$

and the boundary condition $K(x, y; \xi, \eta) = 0$ on ∂G. We would like to conclude that $K(x, y; \xi, \eta) \geq 0$ in G in view of its probabilistic interpretation in Section 1.3. Since the delta function is not an ordinary function, it is not obvious that the preceding theorem can be applied. However, we may obtain the delta function in two dimensions as a limit of sequences of ordinary nonnegative functions in the manner of Example 1.2 and Exercise 7.2.1, as is easily shown. For each of these functions we obtain a nonnegative solution of (8.4.15), where the delta function is replaced by a member of the sequence of functions. This follows since $K(x, y; \xi, \eta)$ and each of its approximations vanishes on ∂G and the right side of (8.4.15) is nonpositive. Technically, the minimum principle proved in the foregoing requires the right side of (8.4.12) to be strictly negative in G, and the members of the delta sequence are such that $-F(x, y)$ in (8.4.12) can only be said to be nonpositive. However, the minimum principle is valid in that case also, as easily follows from the proof of that principle given for Laplace's equation below. In the limit, as the sequence tends to the delta function, the sequence of approximate Green's functions tends to the Green's function $K(x, y; \xi, \eta)$, and this, in turn, must also be nonnegative.

The *maximum principle* for the *Laplace's equation* in two dimensions is proven by following the procedure used above for the diffusion equation. We consider a solution $u(x, y)$ of Laplace's equation,

$$u_{xx}(x, y) + u_{yy}(x, y) = 0, \tag{8.4.16}$$

in the region G with the boundary ∂G. Let M be the maximum value of $u(x, y)$ on ∂G and suppose that the (bounded) region G can be enclosed in a square of side $2l$

with center at the origin. Then the auxiliary function $w(x, y) = u(x, y) + \epsilon x^2$, with $\epsilon > 0$, satisfies the equation

$$\nabla^2 w(x, y) = \nabla^2[u(x, y) + \epsilon x^2] = 2\epsilon > 0. \tag{8.4.17}$$

Based on our considerations for Poisson's equation with $F(x, y) < 0$, we conclude that $w(x, y)$ attains its maximum on ∂G. Then for $(x, y) \in G$, we have $w(x, y) \leq M + \epsilon l^2$, since $l^2 \geq x^2$ for all x on ∂G by our assumption. Finally, $u(x, y) = w(x, y) - \epsilon x^2 \leq w(x, y) \leq M + \epsilon l^2$. Since ϵ is arbitrary, we conclude that

$$u(x, y) \leq M \tag{8.4.18}$$

on letting $\epsilon \to 0$, so that the maximum is assumed on the boundary.

By considering $-u(x, y)$ in place of $u(x, y)$, we conclude, as was done earlier, that the minimum is also assumed on the boundary. As we have not excluded the possibility that $u(x, y)$ can assume the maximum and minimum values in the region G, our result is a *weak maximum and minimum principle*. It can also be shown for Laplace's equation that unless $u(x, y)$ is identically constant, the normal derivative $\partial u(x, y)/\partial n$ on ∂G is positive at a point where $u(x, y)$ assumes its maximum and negative at a point where $u(x, y)$ assumes its minimum.

These results can be used to prove *uniqueness* and *continuous dependence on data* for boundary value problems for Laplace's and Poisson's equations. They can be extended to three dimensions and to certain problems with variable coefficients of the type considered in this text. This is done in the exercises.

The *minimum principle* for Laplace's equation again shows that nonnegative boundary data imply nonnegative solutions as required by our random walk formulation in Section 1.3. We have already derived the *mean value property* for solutions of Laplace's equation (see Exercise 7.5.27). This property can be used to prove the maximum and minimum principles for solutions for Laplace's equation. It can also be used to show that whenever the maximum or minimum is attained in the interior of the region (as well as on the boundary), the solution $u(x, y)$ must be identically constant. This is the *strong maximum and minimum principle* for Laplace's equation. It is discussed in the exercises.

Positivity Principle for the Telegrapher's Equation

To conclude this section, we consider the *telegrapher's equation* of Section 1.2, that is,

$$v_{tt}(x, t) - \gamma^2 v_{xx}(x, t) + 2\lambda v_t(x, t) = 0, \tag{8.4.19}$$

and express it in the form of a system

$$\alpha_t(x, t) + \gamma \alpha_x(x, t) = -\lambda \alpha(x, t) + \lambda \beta(x, t), \tag{8.4.20}$$

$$\beta_t(x, t) - \gamma \beta_x(x, t) = \lambda \alpha(x, t) - \lambda \beta(x, t), \tag{8.4.21}$$

as in (1.2.23)–(1.2.24), with $v(x,t) = \alpha(x,t) + \beta(x,t)$. Since $\alpha(x,t)$ and $\beta(x,t)$ are interpreted as probability densities, we require that $\alpha(x,t)$ and $\beta(x,t)$ be nonnegative. In particular, if the initial data

$$\alpha(x,0) = f(x), \qquad \beta(x,0) = g(x), \qquad -\infty < x < \infty \tag{8.4.22}$$

are nonnegative, we must have $\alpha(x,t) \geq 0$ and $\beta(x,t) \geq 0$ for $t > 0$. In terms of $v(x,t)$ we should have $v(x,t) \geq 0$ if $v(x,0) \geq 0$ when solving a Cauchy problem for $v(x,t)$. We will show that if the data $f(x)$ and $g(x)$ in (8.4.22) are strictly positive, the solutions $\alpha(x,t)$ and $\beta(x,t)$ must also be positive. The case of nonnegative data is considered in the exercises.

To verify these results we consider the characteristic curves for (8.4.19)–(8.4.21), which are given as $x \pm \gamma t =$ constant. On the curves $x \pm \gamma t =$ constant we have for any function $w(x,t)$

$$\frac{dw}{dt} = w_x \frac{dx}{dt} + w_t = \mp \gamma w_x + w_t. \tag{8.4.23}$$

This yields for (8.4.20),

$$\alpha_t + \gamma \alpha_x = \frac{d\alpha}{dt} = -\lambda\alpha + \lambda\beta \qquad \text{on} \quad x - \gamma t = \text{constant} \tag{8.4.24}$$

and for (8.4.21),

$$\beta_t - \gamma \beta_x = \frac{d\beta}{dt} = \lambda\alpha - \lambda\beta \qquad \text{on} \quad x + \gamma t = \text{constant}. \tag{8.4.25}$$

These equations can be written as

$$\frac{d}{dt}[e^{\lambda t}\alpha] = \lambda e^{\lambda t}\beta \qquad \text{on} \quad x - \gamma t = \text{constant}, \tag{8.4.26}$$

$$\frac{d}{dt}[e^{\lambda t}\beta] = \lambda e^{\lambda t}\alpha \qquad \text{on} \quad x + \gamma t = \text{constant}. \tag{8.4.27}$$

Now if $\alpha(x,0) > 0$ and $\beta(x,0) > 0$, it follows from (8.4.26)–(8.4.27) that $(d/dt)[e^{\lambda t}\alpha]|_{t=0} > 0$ and $(d/dt)[e^{\lambda t}\beta]|_{t=0} > 0$ since $\lambda > 0$. Initially, therefore, $e^{\lambda t}\alpha(x,t)$ and $e^{\lambda t}\beta(x,t)$ must be increasing functions of time. Since $\alpha(x,t)$ and $\beta(x,t)$ are initially positive, it is not possible that they become negative at any later time t, for (8.4.26)–(8.4.27) are valid for all time. Consequently, $v(x,t) = \alpha(x,t) + \beta(x,t)$ is also positive, as was to be shown.

Exercises 8.4

8.4.1. Use the maximum principle to prove that the initial and boundary value problem for the heat equation, $u_t(x,t) - c^2 u_{xx}(x,t) = F(x,t)$, $0 < x < l$, $t > 0$, $u(x,0) = f(x)$, $u(0,t) = g_1(t)$, $u(l,t) = g_2(t)$, has a unique continuous solution. *Hint*:

Assume there are two solutions $u_1(x,t)$ and $u_2(x,t)$ and consider the difference $\hat{u}(x,t) = u_1(x,t) - u_2(x,t)$.

8.4.2. Show that if $u_1(x,t)$ and $u_2(x,t)$ are solutions of the heat equation $u_t(x,t) = c^2 u_{xx}(x,t)$ whose initial and boundary data at $x = 0$, $t = 0$, and $x = l$ satisfy the inequality $|u_1(x,t) - u_2(x,t)| < \epsilon$, $x = 0$, $t = 0$, $x = l$, $|u_1(x,t) - u_2(x,t)| < \epsilon$ for all (x,t) in $0 \leq x \leq l$, $t \geq 0$. This proves that solutions of the first boundary value problem for the heat equation depend continuously on the data. *Hint*: Consider the solutions $v_1(x,t) = -\epsilon$, $u(x,t) = u_1(x,t) - u_2(x,t)$ and $v_2(x,t) = \epsilon$ of the heat equation and use the maximum principle to conclude that $v_1(x,t) \leq u(x,t) \leq v_2(x,t)$ by examining these functions two at time.

8.4.3. Prove that the Cauchy problem for the heat equation $u_t(x,t) - c^2 u_{xx}(x,t) = F(x,t)$, $-\infty < x < \infty$, $t > 0$, $u(x,0) = f(x)$, $-\infty < x < \infty$, $|u(x,t)| < M$, $-\infty < x < \infty$, $t \geq 0$, has a unique continuous solution. Assume that the problem has two solutions $u_1(x,t)$ and $u_2(x,t)$ and consider the difference $\hat{u}(x,t) = u_1(x,t) - u_2(x,t)$. Consider the interval $-l \leq x \leq l$ and the function $v(x,t) = (4M/l^2)(c^2 t + x^2/2)$ which is a solution of the (homogeneous) heat equation. Show that $|\hat{u}(x,t)| \leq v(x,t)$ on the boundary $x = -l$, $t = 0$, $x = l$ and conclude from the maximum principle in the bounded interval $|x| \leq l$ that $|\hat{u}(x,t)| \leq v(x,t)$ in $|x| \leq l$ and $t \geq 0$. By letting $l \to \infty$, conclude that $\hat{u}(x,t) = 0$.

8.4.4. Consider the initial and boundary value problem given in Exercise 8.4.1. Let R be the region $0 \leq x \leq l$ and $0 \leq t \leq T$, and introduce the auxiliary function $w(x,t) = u(x,t) + \frac{1}{2}(x^2/c^2) \max_R |F(x,t)|$, as well as the function $\hat{w}(x,t)$, which has the form of $w(x,t)$ except that $u(x,t)$ is replaced by $-u(x,t)$. Use the maximum principle for the inhomogeneous heat equation to show that $|u(x,t)| \leq \max_S |u(x,t)| + \frac{1}{2}(l^2/c^2) \max_R |F(x,t)|$, where S represents the part of ∂R that coincides with $t = 0$, $x = 0$, and $x = l$. Use this inequality to prove continuous dependence on the data.

8.4.5. Show that the problem of Exercise 8.4.1 with Neumann boundary conditions at $x = 0$ and $x = l$, has a unique solution. *Hint*: Use the fact that $u_x(x,t) \neq 0$ at a maximum or minimum point on $x = 0$ and $x = l$ to conclude that the homogeneous version of the problem has its extreme values at $t = 0$.

8.4.6. Show that the problem of Exercise 8.4.1 with Robin or mixed boundary conditions has a unique solution. *Hint*: Consider what happens at a maximum or minimum point on the boundary and use the hint in Exercise 8.4.5.

8.4.7. Determine a maximum and minimum principle for the heat equation $u_t(\mathbf{x},t) - c^2 \nabla^2 u(\mathbf{x},t) = 0$ in a bounded region in two or three dimensions. *Hint*: Proceed as in the one-dimensional case and replace $w(x,t)$ in (8.4.5) by $w(\mathbf{x},t) = u(\mathbf{x},t) + \epsilon r^2$, where $r^2 = x^2 + y^2$ and $r^2 = x^2 + y^2 + z^2$ in two and three dimensions, respectively. Note that the bounded region may be assumed to be contained in a circle or sphere of sufficiently large radius.

8.4.8. Obtain a maximum principle for the parabolic equation $\rho(\mathbf{x}) u_t(\mathbf{x},t) - \nabla \cdot (p(\mathbf{x}) \nabla u(\mathbf{x},t)) = 0$, in a bounded region, with bounded $\rho(\mathbf{x}) > 0$ and $p(\mathbf{x}) > 0$. Similarly, obtain a maximum principle for the equation in Exercise 1.1.17. *Hint*:

Let $w(\mathbf{x},t) = u(\mathbf{x},t) + \epsilon \exp[ax]$ and then choose a such that $\rho(\mathbf{x})w_t(\mathbf{x},t) - \nabla \cdot (p(\mathbf{x})\nabla w(\mathbf{x},t)) < 0$. Conclude that $w(\mathbf{x},t)$ must have its maximum at $t = 0$ or at the boundary of the given region for $t > 0$. Proceed as in the one-dimensional case in the text.

8.4.9. Use the maximum principle of Exercise 8.4.8 to prove that the first boundary and initial value problem for $\rho(\mathbf{x})u_t(\mathbf{x},t) - \nabla \cdot (p(\mathbf{x})\nabla u(\mathbf{x},t)) = \rho(\mathbf{x})F(\mathbf{x},t)$ has a unique solution.

8.4.10. Consider the elliptic equation $\nabla \cdot (p(\mathbf{x})\nabla u(\mathbf{x})) = -F(\mathbf{x})$, $p(\mathbf{x}) > 0$ in two or three dimensions in a bounded region G with the boundary ∂G. Show that if $F(\mathbf{x}) < 0$ in G, the solution $u(\mathbf{x})$ assumes its maximum on ∂G, and if $F(\mathbf{x}) > 0$ in G, the solution $u(\mathbf{x})$ assumes its minimum on ∂G. *Hint*: The first partial derivatives of $u(\mathbf{x})$ vanish at an interior maximum point.

8.4.11. Obtain a maximum principle for the elliptic equation $\nabla \cdot (p(\mathbf{x})\nabla u(\mathbf{x})) = 0$, $p(\mathbf{x}) > 0$, in the two- and three-dimensional cases. Similarly, obtain a maximum principle for equation (1.3.21). *Hint*: Proceed as in Exercise 8.4.8 and use Exercise 8.4.10

8.4.12. Use the maximum principle to prove continuous dependence on the data for the Dirichlet problem for the elliptic equation $\nabla \cdot (p(\mathbf{x})\nabla u(\mathbf{x})) = 0$, $p(\mathbf{x}) > 0$, in a bounded region. *Hint*: Adapt the method of Exercise 8.4.2.

8.4.13. Use the fact that the normal derivative $\partial u(\mathbf{x})/\partial n$ must be nonzero at a maximum or minimum point on the boundary ∂G to prove that the second, third, and mixed boundary value problems for the elliptic equation of Exercise 8.4.10 have a unique solution. In the case of the second boundary value problem the solution is unique up to an arbitrary constant.

8.4.14. Consider the elliptic equation $-\nabla \cdot (p(\mathbf{x})\nabla u(\mathbf{x})) + q(\mathbf{x})u(\mathbf{x}) = 0$, $p(\mathbf{x}) > 0$, $q(\mathbf{x}) > 0$ in a bounded region G. Show that $u(\mathbf{x}) \leq 0$ at an interior maximum point and that $u(\mathbf{x}) \geq 0$ at an interior minimum point. Conclude from this that the Dirichlet problem for the equation with $u(\mathbf{x}) = f(\mathbf{x})$ on ∂G must have a unique solution.

8.4.15. Let $u(x,y,z)$ satisfy $\nabla^2 u(x,y,z) = 0$ in G, and let P_0 and the sphere S_0 centered at P_0 with radius a lie completely within G. Then the mean value theorem for harmonic functions states [see (7.5.35)] $u(P_0) = \frac{1}{4\pi a^2} \iint_{S_0} u(x,y,z)\, ds$, with the integration taken over the sphere S_0. Assuming that $u(x,y,z)$ is continuous up to ∂G so that it has a maximum in the closed region, use the mean value theorem to show that the maximum is achieved on the boundary ∂G. *Hint*: Assume that the maximum point M_0 is interior to G. Apply the mean value principle to all spheres centered at M_0 extending up to the boundary ∂G. Since $u(M_0) \geq u(P)$ for each P in G or ∂G by assumption, conclude that if $u(P) < u(M_0)$, we would be led to the contradiction that $u(M_0) < u(M_0)$. This argument can be extended to conclude that if $u(x,y,z)$ does attain its maximum in G, it must be a constant function.

8.4.16. Show that the solution $u(x,t)$ of the Cauchy problem for the wave equation, $u_{tt}(x,t) - c^2 u_{xx}(x,t) = 0$, $-\infty < x < \infty$, $t > 0$, $u(x,0) = f(x)$, $u_t(x,0) =$

$g(x)$, $-\infty < x < \infty$, is nonnegative if $f(x) \geq 0$ and $g(x) \geq 0$. *Hint*: Use d'Alembert's solution.

8.4.17. Use the solution formula (7.4.43) for the Cauchy problem for the modified telegrapher's equation (7.4.42) and the properties of the modified Bessel functions $I_0(z)$ and $I_1(z)$ to show that if the data $f(x)$, $g(x)$, and $F(x,t)$ are nonnegative, so is $u(x,t)$. Show how this relates to the result for the telegrapher's equation given in the text.

8.5 SOLUTION METHODS FOR HIGHER-ORDER PDEs AND SYSTEMS OF PDEs

With the exception of Chapter 2, which deals with first order partial differential equations, we have mostly studied second order partial differential equations in this text. In this section we consider a number of important physical processes that are governed by higher-order PDEs and systems of PDEs. In a number of cases the techniques used previously for second order equations can be applied, whereas in other cases special methods need to be introduced. We begin by considering the higher order equations that govern the *lateral vibrations of rods and plates*. Then we discuss the *equations of fluid dynamics, Maxwell's equations of electromagnetic theory*, and the *equations of elasticity theory*. Both linear and nonlinear problems are considered.

Although the problems to be considered are not of the general form studied previously in the text, it will be shown in some cases that they can be reduced to a consideration of PDEs studied previously. However, we only touch upon some of the methods used for these problems, and each of the theories of fluid dynamics, electromagnetics, and elasticity has associated with it a rich body of mathematical methods. Additionally, in Chapters 9 to 11 we consider approximate methods for solving PDEs, and these techniques are useful not only for the PDEs studied previously but also for the types of problems considered in this section.

Lateral Vibration of a Rod of Infinite Length

The *lateral vibration of a thin homogeneous rod* is governed by the PDE

$$u_{tt}(x,t) + c^2\, u_{xxxx}(x,t) = 0, \tag{8.5.1}$$

where $u(x,t)$ is the deflection of the rod at the time t. The constant c^2 depends on the physical properties of the rod. (The *longitudinal vibration of a thin homogeneous rod* is governed by the wave equation.) We do not derive (8.5.1) but note that in contrast to the wave equation for the vibrating string, which contains the term $u_{xx}(x,t)$ proportional to the curvature of the string, (8.5.1) contains a fourth derivative term, $u_{xxxx}(x,t)$. Because of the extreme rigidity of the rod, the internal force on a portion of the rod is proportional to a fourth rather than a second derivative of the displacement $u(x,t)$, as for the stretched string. Invoking Section 3.3, we find that (8.5.1) is an equation of *parabolic type* and the lines t = constant are the characteristics.

The *Cauchy problem* for (8.5.1) with prescribed data $u(x,0)$ and $u_t(x,0)$ is *well posed.* To see this, we look for *normal mode solutions* as in Section 3.5, $u(x,t) = a(k)\exp[ikx + \lambda(k)t]$. This yields, for all real k, $\lambda^2(k) + c^2k^4 = 0$, so that $\lambda(k) = \pm ick^2$. The stability index $\Omega = 0$, so that the Cauchy problem is well posed. Furthermore, $\omega(k) = ck^2$, $\lambda(k) = \pm i\omega(k)$, implies that (8.5.1) is *conservative* as well as being of *dispersive type.* The dispersion relation is $\omega(k) = ck^2$, and the group velocity $d\omega(k)/dk = 2ck$ exceeds the phase velocity $\omega(k)/k = ck$ in magnitude. It is worth noting that even though the PDE (8.5.1) is not of hyperbolic type, it is, nevertheless, an equation of dispersive type.

We use *Fourier transform methods* to solve the Cauchy problem for (8.5.1). On proceeding as in Section 5.2, with the initial data for (8.5.1) at $t = 0$ given as

$$u(x,0) = f(x), \qquad u_t(x,0) = g(x), \qquad -\infty < x < \infty, \tag{8.5.2}$$

we denote the Fourier transform of $u(x,t)$ as

$$U(\lambda,t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{i\lambda x}u(x,t)\,dx, \tag{8.5.3}$$

and obtain the Fourier-transformed problem

$$\frac{\partial^2 U(\lambda,t)}{\partial t^2} + c^2\lambda^4 U(\lambda,t) = 0,\ t > 0, \quad U(\lambda,0) = F(\lambda),\ \frac{\partial U(\lambda,0)}{\partial t} = G(\lambda), \tag{8.5.4}$$

where $F(\lambda)$ and $G(\lambda)$ are the Fourier transforms of $f(x)$ and $g(x)$, respectively. It has the solution

$$U(\lambda,t) = \left[\frac{1}{2}F(\lambda) + \frac{1}{2ic\lambda^2}G(\lambda)\right]e^{ic\lambda^2 t} + \left[\frac{1}{2}F(\lambda) - \frac{1}{2ic\lambda^2}G(\lambda)\right]e^{-ic\lambda^2 t}. \tag{8.5.5}$$

Inverting the transform yields

$$u(x,t) = \frac{1}{\sqrt{2\pi}}\left\{\int_{-\infty}^{\infty} H_+(\lambda)e^{i(\omega(\lambda)t - \lambda x)}\,d\lambda + \int_{-\infty}^{\infty} H_-(\lambda)e^{i(-\omega(\lambda)t - \lambda x)}\,d\lambda\right\}, \tag{8.5.6}$$

where $\omega(\lambda) = c\lambda^2$ and $H_\pm(\lambda)$ are the coefficients of $e^{\pm ic\lambda^2 t}$ in (8.5.5).

Rather than evaluate (8.5.6) for specific choices of $f(x)$ and $g(x)$, we use the *method of stationary phase* of Section 5.7 to analyze the behavior of the solution $u(x,t)$ as t gets large. The integrals in (8.5.6) may be written as

$$I_\pm(x,t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} H_\pm(\lambda)\exp[i\phi_\pm(x,t;\lambda)]\,d\lambda, \tag{8.5.7}$$

with $\phi_\pm(x,t;\lambda) = \pm\omega(\lambda)t - \lambda x = \pm c\lambda^2 t - \lambda x$. We assume that x and t are large. Then the stationary points are determined from $d\phi_\pm(x,t;\lambda)/d\lambda = \pm 2c\lambda t - x = 0$, so that $\lambda = \lambda_\pm = \pm x/2ct$. Thus, say, if $x > 0$, we have $\lambda_+ = x/2ct$ and

$\lambda_- = -x/2ct$ as positive and negative stationary points for the integrals I_+ and I_-, respectively. Also, $\phi_\pm(x,t;\lambda_\pm) = \pm c\left(x^2/4c^2t^2\right)t \mp x^2/2ct = \mp x^2/4ct$ and $d^2\phi_\pm(x,t;\lambda_\pm)/d\lambda^2 = \pm 2ct$, so that $d^2\phi_\pm/d\lambda^2/|d^2\phi_\pm/d\lambda^2| = \pm 1$. Furthermore, $d^2\omega(\lambda_\pm)/d\lambda^2 = 2c$, so that using an appropriate modification of (5.7.4) gives

$$u(x,t) \approx \frac{H_+(\lambda_+)}{\sqrt{2ct}} \exp\left(-\frac{ix^2}{4ct} + \frac{i\pi}{4}\right) + \frac{H_-(\lambda_-)}{\sqrt{2ct}} \exp\left(\frac{ix^2}{4ct} - \frac{i\pi}{4}\right), \quad |x|, t \to \infty. \tag{8.5.8}$$

If the initial data $f(x)$ and $g(x)$ are real valued, it is easy to see, on using the definition of the Fourier transform, that $H_+(\lambda)$ and $H_-(-\lambda)$ are complex conjugates. Since $\lambda_+ = -\lambda_-$, we conclude that (8.5.8) is real valued.

Thus near the (group) lines $x/2ct$ = constant, the solution of the initial value problem for (8.5.1) for large t is approximated by (8.5.8) and the solution decays like $1/\sqrt{t}$ as $t \to \infty$.

Lateral Vibration of a Rod of Finite Length

The *lateral vibration of a thin homogeneous rod* of finite extent, say, of length l, is again governed by (8.5.1), with $u(x,t)$ as the deflection of the rod. We place the rod in the interval $0 < x < l$ and prescribe boundary conditions at $x = 0$ and $x = l$. The boundary conditions

$$u(0,t) = u_x(0,t) = 0, \; u(l,t) = u_x(l,t) = 0, \tag{8.5.9}$$

imply that the rod is *clamped* at both edges so that u and the slope u_x both vanish. The resulting initial and boundary value problem for $u(x,t)$ in $0 < x < l$, where $u(x,t)$ satisfies (8.5.1), the initial conditions (8.5.2) in the finite interval, and the boundary conditions (8.5.9), can be solved by *separation of variables*.

We set $u(x,t) = F(x)G(t)$ in (8.5.1), and this yields the separated equation $G''(t)/c^2G(t) = -F''''(x)/F(x) = -\lambda^2$, with $-\lambda^2$ as the separation constant. Then we obtain

$$G''(t) + (\lambda c)^2 G(t) = 0, \quad F''''(x) - \lambda^2 F(x) = 0. \tag{8.5.10}$$

The conditions (8.5.9) imply that $F(0) = F'(0) = F(l) = F'(l) = 0$, so that we have an eigenvalue problem for $F(x)$. We do not discuss general properties of higher-order eigenvalue problems but solve the preceding problem directly.

The general solution of the ODE for $F(x)$ in (8.5.10) is (for $\lambda > 0$) $F(x) = a\cos(\sqrt{\lambda}\,x) + b\cosh(\sqrt{\lambda}\,x) + c\sin(\sqrt{\lambda}\,x) + d\sinh(\sqrt{\lambda}\,x)$. The conditions $F(0) = F'(0) = 0$ imply that $b = -a$ and $d = -c$. Thus $F(x) = a[\cos(\sqrt{\lambda}\,x) - \cosh(\sqrt{\lambda}\,x)] + c[\sin(\sqrt{\lambda}\,x) - \sinh(\sqrt{\lambda}\,x)]$. At $x = l$ we have $\{a[\cos(\sqrt{\lambda}\,l) - \cosh(\sqrt{\lambda}\,l)] + c[\sin(\sqrt{\lambda}\,l) - \sinh(\sqrt{\lambda}\,l)] = 0,\; a[-\sin(\sqrt{\lambda}\,l) - \sinh(\sqrt{\lambda}\,l)] + c[\cos(\sqrt{\lambda}\,l) - \cosh(\sqrt{\lambda}\,l)] = 0\}$. For a and c to be nonzero, the determinant of the coefficients of this system must vanish. This gives the eigenvalue equation $2 - 2\cos(\sqrt{\lambda}\,l)\cosh(\sqrt{\lambda}\,l) = 0$. This equation be written as $\cosh(\sqrt{\lambda}\,l) = \sec(\sqrt{\lambda}\,l)$.

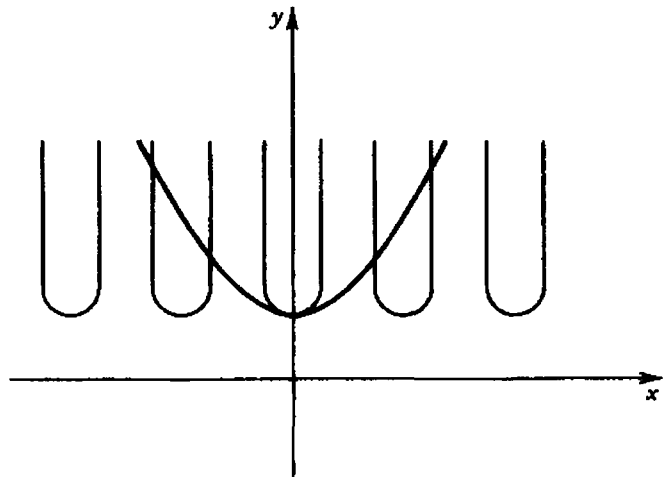

Figure 8.4 The graphs of $\cosh(x)$ and $\sec(x)$.

As the graph in Figure 8.4 shows, the curves for $\cosh(\sqrt{\lambda}\,l)$ and $\sec(\sqrt{\lambda}\,l)$ have infinitely many intersections. [Zero is not an eigenvalue since $F(x)$ vanishes for $\lambda = 0$.] Clearly, the eigenvalues tend to infinity. Since $\cosh(\sqrt{\lambda}l)$ grows exponentially with $\sqrt{\lambda}$, the large eigenvalues are given approximately by the zeros of $\cos(\sqrt{\lambda})$ since $\sec(\sqrt{\lambda}) = 1/\cos(\sqrt{\lambda})$. Let λ_k $(k = 1, 2, \ldots)$ denote the eigenvalues. Then a set of eigenfunctions is given by

$$F_k(x) = [\sin(\sqrt{\lambda_k}\,l) - \sinh(\sqrt{\lambda_k}\,l)][\cos(\sqrt{\lambda_k}\,x) - \cosh(\sqrt{\lambda_k}\,x)]$$
$$- [\cos(\sqrt{\lambda_k}\,l) - \cosh(\sqrt{\lambda_k}\,l)][\sin(\sqrt{\lambda_k}\,x) - \sinh(\sqrt{\lambda_k}\,x)]. \quad (8.5.11)$$

The $F_k(x)$ are not normalized. It can be shown that $\int_0^l F_k(x)F_j(x)\,dx = 0$ if $\lambda_k \neq \lambda_j$ by direct verification.

Furthermore, it can be shown that the set $\{F_k(x)\}$ is complete and any smooth function in $0 < x < l$ can be expanded in a series of the $F_k(x)$.

Solving for $G_k(t)$ from (8.5.10) gives $G_k(t) = \alpha_k \cos(\lambda_k ct) + \beta_k \sin(\lambda_k ct)$, with α_k and β_k as arbitrary constants. The formal solution of (8.5.1) is obtained by superposition as

$$u(x,t) = \sum_{k=1}^{\infty} F_k(x)G_k(t). \quad (8.5.12)$$

Formally applying the initial conditions to the series (8.5.12) leads to the specification of α_k and β_k in terms of the (Fourier) coefficients of the functions $f(x)$ and $g(x)$ when expanded in a series of eigenfunctions $F_k(x)$.

Other types of boundary conditions for (8.5.1) and properties of the corresponding eigenvalue problems for $F(x)$ are given in the exercises.

Vibration of a Plate

The *vibration of a thin plate* is governed by the equation

$$u_{tt}(x,y,t) + c^2\nabla^2\nabla^2 u(x,y,t) = 0, \tag{8.5.13}$$

where $u(x,y,t)$ is the transverse displacement of the point (x,y) on the plate at the time t, and ∇^2 is the two-dimensional Laplacian operator. The constant c^2 depends on the physical properties of the plate. The wave equation that governs the *vibration of a membrane* is given as $u_{tt}(x,y,t) - c^2\nabla^2 u(x,y,t) = 0$ and contains only the Laplacian operator, while (8.5.13) contains the iterated Laplacian because of the greater rigidity of the plate. If we assume that the plate occupies the region G in the (x,y)-plane and is *clamped* at the boundary or edge ∂G, we have the boundary conditions

$$u(x,y,t)\Big|_{\partial G} = 0, \qquad \frac{\partial u(x,y,t)}{\partial n}\Big|_{\partial G} = 0. \tag{8.5.14}$$

Also, $u(x,y,t)$ must satisfy the initial conditions

$$u(x,y,0) = f(x,y), \qquad u_t(x,y,0) = g(x,y). \tag{8.5.15}$$

Because (8.5.13) is of fourth order in the spatial variables, we require two boundary conditions. However, since there are only two time derivatives in (8.5.13) we have two initial conditions. The problem (8.5.13)–(8.5.15) may be shown to be well posed.

We attempt to solve this initial and boundary value problem by using separation of variables. Let

$$u(x,y,t) = F(x,y)G(t). \tag{8.5.16}$$

This yields, on inserting (8.5.16) into (8.5.13),

$$\frac{G''(t)}{c^2 G(t)} = -\frac{\nabla^2\nabla^2 F(x,y)}{F(x,y)} = -\lambda, \tag{8.5.17}$$

where $-\lambda$ is the separation constant. The equation for $G(t)$ is

$$G''(t) + \lambda c^2 G(t) = 0, \tag{8.5.18}$$

and $F(x,y)$ must satisfy

$$\nabla^2\nabla^2 F(x,y) - \lambda F(x,y) = 0 \tag{8.5.19}$$

for $(x,y) \in G$ with the boundary conditions

$$F(x,y)\Big|_{\partial G} = 0, \qquad \frac{\partial F(x,y)}{\partial n}\Big|_{\partial G} = 0. \tag{8.5.20}$$

That is, we obtain an eigenvalue problem for $F(x,y)$.

To make any further progress with this problem we must be able to determine the eigenvalues λ and the eigenfunctions $F(x, y)$ for specific regions. It turns out, however, that (8.5.19) is only separable in polar coordinates, so that more complicated or, possibly, approximation methods are needed if the problem does not involve a circular clamped plate. Indeed, in rectangular coordinates, (8.5.19) takes the form $F_{xxxx}(x, y) + 2F_{xxyy}(x, y) + F_{yyyy}(x, y) - \lambda F(x, y) = 0$. Then, if we set $F(x, y) = A(x)B(y)$, we have

$$A''''(x)B(y) + 2A''(x)B''(y) + B''''(y)A(x) - \lambda A(x)B(y) = 0, \qquad (8.5.21)$$

and it is not possible to separate the variables x and y in (8.5.21).

For the *circular plate* the eigenfunctions $F(x, y)$ can be obtained as follows. Let $\lambda = k^4$ for simplicity of notation and factor (8.5.19) as

$$(\nabla^2\nabla^2 - k^4)F(x, y) = (\nabla^2 + k^2)(\nabla^2 - k^2)F(x, y) = 0. \qquad (8.5.22)$$

This factorization can be carried out irrespective of the choice of coordinates, but only in the case of polar coordinate problem does it lead to a solution of the eigenvalue problem (8.5.19)–(8.5.20). We express F in polar coordinates as $F = F(r, \theta)$, and assume that the region G is the disk $r < R$ with the circle $r = R$ as the boundary ∂G. Then $F(r, \theta)$ must be bounded at $r = 0$ and must be periodic of period 2π in θ—that is, $F(r, \theta + 2\pi) = F(r, \theta)$—since $F(r, \theta)$ must be single valued in G. Thus, if $F(r, \theta)$ satisfies either of the equations

$$(\nabla^2 \pm k^2)F(r, \theta) = F_{rr}(r, \theta) + \frac{1}{r}F_r(r, \theta) + \frac{1}{r^2}F_{\theta\theta}(r, \theta) \pm k^2 F(r, \theta) = 0, \quad (8.5.23)$$

it is a solution of (8.5.19) since the order of the operators in (8.5.22) can be interchanged.

The equations (8.5.23) can be solved using separation of variables. Let $F(r, \theta) = V(r)W(\theta)$, and (8.5.23) can be expressed as

$$V''(r) + \frac{1}{r}V'(r) \pm k^2 V(r) - \frac{n^2}{r^2}V(r) = 0, \; W''(\theta) + n^2 W(\theta) = 0 \qquad (8.5.24)$$

after separating the variables, with n^2 as the separation constant. The periodicity of $F(r, \theta)$ implies that $W(\theta + 2\pi) = W(\theta)$, so that n must be a (positive) integer or zero (i.e., $n = 0, 1, 2, \ldots$). Consequently, we have $W_n(\theta) = \cos(n\theta)$ and $W_n(\theta) = \sin(n\theta)$ as the appropriate solutions $W(\theta)$ in (8.5.24). For each value of n, (8.5.24) is a Bessel equation for $V(r)$ of order n if the plus sign (i.e., $+k^2$) is chosen. When the minus sign (i.e., $-k^2$) is selected in (8.5.24), we have the modified Bessel equation of order n.

The product functions $F_n(r, \theta) = V_n(r)W_n(\theta)$ must be bounded at $r = 0$, so that the appropriate solution of (8.5.24) with the term $+k^2$ is $V_n = J_n(kr)$, the Bessel function of nth order. For the case $-k^2$ we obtain $V_n = I_n(kr)$, the modified Bessel

function of nth order. Consequently, we find that for each integer n, we have the solutions

$$F_n(r,\theta) = [\alpha_n J_n(kr) + \gamma_n I_n(kr)]\cos(n\theta) + [\beta_n J_n(kr) + \delta_n I_n(kr)]\sin(n\theta) \tag{8.5.25}$$

of (8.5.22), where α_n, β_n, γ_n, and δ_n are constants. On applying the boundary conditions (8.5.20) to each of the $F_n(r,\theta)$, we determine the eigenvalues $\lambda = k^4$. For the circular region G, the normal derivative $\partial F(r,\theta)/\partial n$ becomes $\partial F(R,\theta)/\partial r$ on $r = R$. Since the boundary conditions must be valid for all θ, we see that the coefficients of $\cos(n\theta)$ and $\sin(n\theta)$ must vanish separately once the boundary conditions are applied. The coefficient of $\cos(n\theta)$ yields the system

$$\alpha_n J_n(kR) + \gamma_n I_n(kR) = 0, \quad \alpha_n J_n'(kR) + \gamma_n I_n'(kR) = 0. \tag{8.5.26}$$

A second system with α_n and γ_n replaced by β_n and δ_n, respectively, results from the coefficient of $\sin(n\theta)$. Since the coefficients α_n and γ_n must not vanish, we conclude that the determinant of the coefficients in (8.5.26) vanishes and this yields the transcendental equation

$$J_n(kR)I_n'(kR) - I_n(kR)J_n'(kR) = 0 \tag{8.5.27}$$

for the determination of the eigenvalues. (The second system yields the same result.) The appropriate values of k must be specified approximately and we do not consider this matter here. Once the full set of eigenvalues has been found we solve (8.5.18) for $G(t)$. Then, on summing over all the product solutions (8.5.16), we may specify the two sets of arbitrary constants in the solution of (8.5.18) so as to satisfy the initial conditions (8.5.15).

Since the general problem is quite complicated, we consider a simplified case where the initial data (8.5.15) are functions of the radial variable r only, that is,

$$u(r,\theta,0) = f(r), \qquad u_t(r,\theta,0) = g(r). \tag{8.5.28}$$

Then the eigenfunctions $F(r,\theta)$ may be chosen to be independent of θ, and we obtain instead of (8.5.25), $F(r) = \alpha_0 J_0(kr) + \gamma_0 I_0(kr)$. The equations (8.5.26)–(8.5.27) remain valid, but we must put $n = 0$. Let k_m $(m = 1, 2, 3, \ldots)$ represent the (real) roots of (8.5.27) with $n = 0$. The (unnormalized) eigenfunctions $F_m(r)$ then have the form

$$F_m(r) = I_0(k_m R)J_0(k_m r) - J_0(k_m R)I_0(k_m r). \tag{8.5.29}$$

The $\{F_m(r)\}$ are an orthogonal set with weight function r over the interval $0 \le r \le R$. With $\lambda_m = k_m^4$, the solutions of (8.5.18) are

$$G_m(t) = a_m \cos(k_m^2 ct) + b_m \sin(k_m^2 ct) \tag{8.5.30}$$

with arbitrary constants a_m and b_m.

The formal series $u = \sum_{m=1}^{\infty} F_m(r)G_m(t)$ satisfies the equation (8.5.13) in the disk $0 \leq r < R$, the boundary conditions (8.5.14) on $r = R$, and the initial conditions (8.5.28) once the constants a_m and b_m in (8.5.30) are specified in terms of $f(r)$ and $g(r)$. Using the orthogonality properties of the eigenfunctions $\{F_m(r)\}$, this is easy to carry out. The eigenfunctions for the eigenvalue problem (8.5.19)–(8.5.20) are known to form a complete set, so that the expansions discussed are valid if the initial data satisfy certain smoothness conditions.

Static Deflection of a Plate: The Biharmonic Equation

It is of considerable interest to study the stationary or static version of the vibrating plate equation (8.5.13). Then, the displacements $u = u(x, y)$ are independent of time, and we find that $u(x, y)$ satisfies the *biharmonic equation*

$$\nabla^2\nabla^2 u(x, y) = 0. \tag{8.5.31}$$

Clearly, any harmonic function $u(x, y)$ [i.e., a solution of Laplace's equation $\nabla^2 u(x, y) = 0$] will also satisfy (8.5.31). This can occasionally be used to advantage in solving boundary value problems for the biharmonic equation.

We consider a *clamped circular plate* of radius R and let $u(r, \theta)$ represent the deflection at a point (r, θ). Using direct verification it may be shown that the function

$$u(r, \theta) = (r^2 - R^2)v(r, \theta) + w(r, \theta) \tag{8.5.32}$$

is a solution of the biharmonic equation if $v(r, \theta)$ and $w(r, \theta)$ are both harmonic functions. We want to determine $u(r, \theta)$ to be a solution of (8.5.31) in the disk $0 \leq r < R$ with the clamped plate boundary conditions

$$u(r, \theta)\Big|_{r=R} = f(\theta), \qquad \frac{\partial u(r, \theta)}{\partial r}\Big|_{r=R} = g(\theta). \tag{8.5.33}$$

The problem can be solved by choosing $v(r, \theta)$ and $w(r, \theta)$ in (8.5.32) appropriately. The boundary condition $u(R, \theta) = f(\theta)$ implies that $u(R, \theta) = w(R, \theta) = f(\theta)$. Thus $w(r, \theta)$ must be a solution of Laplace's equation in the disk $0 \leq r < R$, which satisfies the Dirichlet condition $w(r, \theta) = f(\theta)$ on the boundary $r = R$. This problem has been solved previously and completely specifies $w(r, \theta)$. The condition $\partial u(r, \theta)/\partial r = g(\theta)$ on $r = R$ implies that $\partial u(R, \theta)/\partial r = 2Rv(R, \theta) + \partial w(R, \theta)/\partial r = g(\theta)$. Since $w(r, \theta)$ has already been specified, $\partial w(R, \theta)/\partial r$ is a known function and we obtain $v(R, \theta) = (1/2R)g(\theta) - (1/2R)\partial w(R, \theta)/\partial r$ as the Dirichlet boundary condition for the harmonic function $v(r, \theta)$. Again, $v(r, \theta)$ is thereby uniquely specified and we have obtained a solution of the boundary value problem (8.5.31) and (8.5.33) in terms of the solution of two boundary value problems for Laplace's equation in a circle.

Next, we briefly consider the *fundamental solution* and *Green's function* methods for the biharmonic equation. We proceed as in Section 6.7 and consider the inhomogeneous biharmonic equation

$$\nabla^2\nabla^2 u(x,y) = -F(x,y). \tag{8.5.34}$$

With P_0 as the point (x_0, y_0), we integrate over a region G with boundary ∂G that contains P_0 in its interior and obtain

$$\iint_G \nabla \cdot [\nabla\nabla^2 u(x,y)]\,dx\,dy = \int_{\partial G} \frac{\partial \nabla^2 u(x,y)}{\partial n} ds = -\iint_G F(x,y)\,dx\,dy = -1, \tag{8.5.35}$$

where we have applied the divergence theorem and chosen $F(x,y)$ to represent a point source with singularity at P_0 so that the integral of $F(x,y)$ equals unity. Introducing polar coordinates with the pole located at P_0 and taking the limit in (8.5.35), we obtain

$$\lim_{\partial G \to P_0} \int_{\partial G} \frac{\partial}{\partial n}(\nabla^2 u)\,ds = \lim_{r\to 0} \int_0^{2\pi} \frac{\partial}{\partial r}(\nabla^2 u)\Big|_{\substack{x=x_0+r\cos(\theta)\\ y=y_0+r\sin(\theta)}} r\,d\theta$$

$$= \lim_{r\to 0}\left[2\pi r \frac{\partial}{\partial r}(\nabla^2 u)\right] = -1 \tag{8.5.36}$$

if we let G represent a disk of radius r centered at P_0. Continuing as in Section 6.7, we assume that $u = u(r)$ and we have $\nabla^2 u(r) = (1/r)\,d/dr\,(r\,du(r)/dr)$. This yields the approximate equation $2\pi r d/dr\,[(1/r)\,d/dr\,(r\,du(r)/dr)] \approx -1$. We integrate and retain only the most singular terms and obtain the fundamental solution

$$u(r) = -\frac{1}{8\pi} r^2 \log r, \tag{8.5.37}$$

where $r^2 = (x-x_0)^2 + (y-y_0)^2$. It can be verified that (8.5.37) solves the biharmonic equation everywhere except at $r = 0$. In fact, apart from the constant factor, (8.5.37) may be derived by looking for a solution of (8.5.31) that depends only on r.

We obtain the *Green's theorem* for an arbitrary region G,

$$\iint_G [u\nabla\nabla^2 v - v\nabla\nabla^2 u]dxdy$$

$$= \int_{\partial G}\left[u\frac{\partial(\nabla^2 v)}{\partial n} - (\nabla^2 v)\frac{\partial u}{\partial n} + (\nabla^2 u)\frac{\partial v}{\partial n} - v\frac{\partial(\nabla^2 u)}{\partial n}\right] ds \tag{8.5.38}$$

on the basis of the identity $u\nabla^2\nabla^2 v - v\nabla^2\nabla^2 u = \nabla \cdot \mathbf{p}$, where $\mathbf{p} = u\nabla\nabla^2 v - v\nabla\nabla^2 u + (\nabla^2 u)\nabla v - (\nabla^2 v)\nabla u$, and the divergence theorem.

The Green's function $K(x, y; \xi, \eta)$ for the biharmonic equation with *clamped plate* boundary conditions is a solution of $\nabla^2\nabla^2 K(x, y; \xi, \eta) = -\delta(x - \xi)\delta(y - \eta)$ that satisfies the boundary conditions $K(x, y; \xi, \eta)\big|_{\partial G} = 0$, $\partial K(x, y; \xi, \eta)/\partial n\big|_{\partial G} = 0$. Now if $u(x, y)$ satisfies the inhomogeneous equation (8.5.34) in G and assumes prescribed boundary conditions $u = s(x, y)$ and $\partial u/\partial n = r(x, y)$ on ∂G, the Green's theorem (8.5.38) [where we set $v = K(x, y; \xi, \eta)$] yields the result

$$u(\xi, \eta) = \iint_G K(x, y; \xi, \eta) F(x, y)\, dx\, dy$$

$$+ \int_{\partial G} \left[(\nabla^2 K(x, y; \xi, \eta)) r(x, y) - s(x, y) \frac{\partial (\nabla^2 K(x, y; \xi, \eta))}{\partial n} \right] ds. \quad (8.5.39)$$

Using special techniques, it is possible to construct the Green's function $K(x, y; \xi, \eta)$ for the case of circular clamped plate. We do not derive or exhibit this Green's function but note that in view of (8.5.37), $K(x, y; \xi, \eta)$ can be expressed as $K(x, y; \xi, \eta) = -(1/8\pi) r^2 \log r + V(x, y)$, where $V(x, y)$ is a regular solution of the homogeneous biharmonic equation. Applying the boundary conditions (8.5.50) to $K(x, y; \xi, \eta)$ yields a set of boundary conditions for $V(x, y)$ and $\partial V(x, y)/\partial n$ on ∂G. Since the boundary value problem for the biharmonic equation with clamped plate boundary conditions was solved previously, it is possible to determine $V(x, y)$ and, thereby, $K(x, y; \xi, \eta)$ for the circular clamped plate.

Euler's Equations of Inviscid Fluid Dynamics

We begin by deriving the *Euler equations* of motion for a nonviscous or *inviscid fluid*. Let $\mathbf{u} = \mathbf{u}(x, y, z, t)$ be the velocity vector of a fluid particle situated at the point (x, y, z) at the time t. Also let $p = p(x, y, z, t)$ and $\rho = \rho(x, y, z, t)$ represent the pressure and the density, respectively, of the fluid. It is assumed that no external forces are acting on the fluid. In the general case, with variable pressure and density, we are dealing with a compressible fluid or gas, and the resulting equations characterize *gas dynamics*.

We consider a volume of the fluid occupying a region R with boundary ∂R. Assuming that there are no sources or sinks for the fluid in R, the rate of increase of the fluid in R in unit time must be balanced by the flux or inflow of fluid through the boundary ∂R. Analytically, this may be expressed as

$$\frac{\partial}{\partial t} \iint_R \rho(x, y, z, t)\, dv = -\int_{\partial R} \rho(x, y, z, t)\, \mathbf{u}(x, y, z, t) \cdot \mathbf{n}\, ds, \quad (8.5.40)$$

where $\mathbf{n}$ is the exterior unit normal vector and the minus sign occurs since we are concerned with the inflow. [We note that $\rho(x, y, z, t)\, dv$ is an element of mass and that $\rho(x, y, z, t)\mathbf{u}(x, y, z, t) \cdot \mathbf{n}\, ds$ represents the flow of mass through the boundary element ds.] Applying the divergence theorem to the integral over ∂R and bringing the time derivative inside the integral over R in (8.5.40)—since R is fixed in time—we obtain

$$\iint_R [\rho_t(x, y, z, t) + \nabla \cdot (\rho(x, y, z, t)\mathbf{u}(x, y, z, t))]\, dv = 0. \quad (8.5.41)$$

Since (8.5.41) must be valid for an arbitrary region R, we conclude that the integrand vanishes identically and obtain

$$\rho_t(x,y,z,t) + \nabla \cdot (\rho(x,y,z,t)\mathbf{u}(x,y,z,t)) = 0, \tag{8.5.42}$$

which is known as the *equation of continuity.*

Next we consider the law of *conservation of momentum* for the fluid. Since we are neglecting frictional forces due to the effects of viscosity, as well as all external forces, the only force acting on the volume of fluid in the region R is due to the pressure along the boundary ∂R. The total pressure is given as $\mathbf{P}(x,y,z,t) = -\int_{\partial R} p(x,y,z,t)\mathbf{n}\, ds$, which is effectively a resultant of all the internally directed normal pressures at the boundary. According to Newton's second law of motion, the pressure force equals the mass times the acceleration of the fluid element or particle. The acceleration of a particle is given by $d\mathbf{u}/dt$, but since the particle is moving along a path $x = x(t),\ y = y(t),\ z = z(t)$, we obtain from the chain rule

$$\frac{d\mathbf{u}}{dt} = \frac{\partial \mathbf{u}}{dt} + \frac{\partial \mathbf{u}}{dx}\frac{dx}{dt} + \frac{\partial \mathbf{u}}{dy}\frac{dy}{dt} + \frac{\partial \mathbf{u}}{dz}\frac{dz}{dt} = \mathbf{u}_t + (\mathbf{u}\cdot\nabla)\mathbf{u}. \tag{8.5.43}$$

Here we have used the fact that

$$\mathbf{u} = [x'(t),\ y'(t),\ z'(t)], \quad \mathbf{u}\cdot\nabla = x'\frac{\partial}{dx} + y'\frac{\partial}{dy} + z'\frac{\partial}{dz}. \tag{8.5.44}$$

The derivative $d\mathbf{u}/dt$ in (8.5.43) is often referred to as a *material derivative.* The total momentum of the fluid particles in R must equal the resultant pressure forces, and we obtain

$$\iiint_R \rho(x,y,z,t)\frac{d\mathbf{u}(x,y,z,t)}{dt}\, dv = -\int_{\partial R} p(x,y,z,t)\mathbf{n}\, ds, \tag{8.5.45}$$

since $\rho(x,y,z,t)\, dv$ is an element of mass. Using the *gradient theorem,* which is a simple consequence of the *divergence theorem,* the integral over ∂R can be converted into an integral over R and (8.5.45) becomes

$$\iiint_R \left[\rho(x,y,z,t)\frac{d\mathbf{u}(x,y,z,t)}{dt} + \nabla p(x,y,z,t)\right] dv = 0. \tag{8.5.46}$$

The arbitrariness of R implies the vanishing of the integrand and we obtain the *equation of conservation of momentum,*

$$\mathbf{u}_t(x,y,z,t) + (\mathbf{u}\cdot\nabla)\mathbf{u}(x,y,z,t) + \frac{1}{\rho(x,y,z,t)}\nabla p(x,y,z,t) = \mathbf{0}. \tag{8.5.47}$$

Further we assume that we are dealing with *adiabatic flow.* This means that the entropy $s(x,y,z,t)$, a thermodynamic quantity, remains constant for each fluid particle as it moves along the particle path $(x(t), y(t), z(t))$. (A thermodynamic relation between the entropy and other physical properties of the fluid is given below

in our discussion of the Navier-Stokes equations.) Thus the material derivative of the entropy $ds(x, y, z, t)/dt$ must equal zero. On using the chain rule and (8.5.44), we obtain

$$s_t(x, y, z, t) + (u \cdot \nabla)s(x, y, z, t) = 0. \tag{8.5.48}$$

The equations (8.5.42), (8.5.47), and (8.5.48) are incomplete as they stand. We have only five equations and six unknowns [i.e., $p(x, y, z, t)$, $\rho(x, y, z, t)$, $s(x, y, z, t)$, and the three components of $\mathbf{u}(x, y, z, t)$]. An *equation of state* that relates the three quantities p, ρ, s completes the system. We give it as

$$p = f(\rho, s), \tag{8.5.49}$$

where the function $f(\rho, s)$ is determined from thermodynamic considerations and differs for different fluids.

The equations (8.5.42), (8.5.47), (8.5.48), and (8.5.49) constitute *Euler's equations of fluid dynamics* for *inviscid adiabatic fluid flow*. If the entropy $s(x, y, z, t)$ is initially uniform, say $s(x, y, z, t) = s_0$, then $ds(x, y, z, t)/dt = 0$ implies that it remains uniform for all time, as each particle carries the same constant value $s = s_0$ for the entropy. (The flow is assumed to remain smooth.) A fluid flow with uniform entropy is said to be *isentropic*. Then we can drop (8.5.48) and replace the equation of state (8.5.49) by

$$p = f(\rho), \tag{8.5.50}$$

since $s = s_0$ for all t. The reduced set of equations (8.5.42), (8.5.47), and (8.5.50) are *Euler's equations for inviscid, isentropic flow*. Each of the systems is a quasilinear first order system that is not easy to solve in general. Consequently, certain simplifying assumptions are introduced to deal with specific types of fluid flow.

Incompressible and Irrotational Fluid Flow

In dealing with the *flow of liquids* rather than gases, we assume that the flow is *isentropic* and that, in addition, $\rho(x, y, z, t) =$ constant. This means that we are dealing with an *incompressible fluid* and this equation replaces the equation of state (8.5.62). (Otherwise, the fluid is compressible, as is the case for a gas.) Noting the above, the equation of continuity (8.5.42) reduces to

$$\nabla \cdot \mathbf{u}(x, y, z, t) = 0. \tag{8.5.51}$$

We further assume that the fluid flow is *irrotational*, which means that the velocity field $\mathbf{u}(x, y, z, t)$ has the property that its curl vanishes; that is,

$$\nabla \times \mathbf{u}(x, y, z, t) = \mathbf{0}. \tag{8.5.52}$$

On the basis of Stokes' theorem from vector analysis, we conclude that if (8.5.52) is satisfied everywhere, the integral of the velocity around any closed curve vanishes; that is,

$$\oint \mathbf{u}(x, y, z, t) \cdot d\mathbf{r} = 0. \tag{8.5.53}$$

The integral in (8.5.53) defines the *circulation* of the fluid around the closed curve and irrotationality means that the flow is circulation free. Since the Euler equations describe the fluid motion as it develops in time, it cannot be assumed that if the flow is irrotational at one time, it does not become rotational at a later time. It can be shown, however, that the fluid remains irrotational for all t (based on Euler's equations) if it was so at a given time.

Rather than consider this question, we examine the problem of *steady fluid flow* for which the pressure $p = p(x, y, z)$ and the velocity $\mathbf{u} = \mathbf{u}(x, y, z)$ are independent of time. Now it is easy to show that if (8.5.53) is satisfied for all closed curves, the velocity vector $\mathbf{u}(x, y, z)$ can be derived from a (velocity) *potential function* $\phi(x, y, z)$ and can be expressed as $\mathbf{u}(x, y, z) = \nabla\phi(x, y, z)$. Inserting this into (8.5.51) gives $\nabla \cdot \mathbf{u}(x, y, z) = \nabla \cdot \nabla\phi(x, y, z) = \nabla^2\phi(x, y, z) = 0$, so that $\phi(x, y, z)$ satisfies *Laplace's equation*. Once Laplace's equation for the velocity potential $\phi(x, y, z)$ is solved, we may determine the pressure p from (8.5.47), where we set $\mathbf{u}_t(x, y, z) = \mathbf{0}$. We see that for *steady, incompressible, and irrotational flow*, that the study of the fluid motion is reduced to the solution of Laplace's equation, subject, of course, to appropriate boundary conditions.

Linearization of Euler's Equations: Acoustics

A linearization procedure for Euler's equations for isentropic flow that results in a study of the wave equation, leads to the *theory of acoustics*. We perturb $\mathbf{u}(x, y, z, t)$, $p(x, y, z, t)$, and $\rho(x, y, z, t)$ around a (constant) equilibrium state $\mathbf{u} = \mathbf{0}$, $p = p_0$, and $\rho = \rho_0$ [with $p_0 = f(\rho_0)$] and write, with ϵ = constant,

$$\begin{cases} \mathbf{u}(x, y, z, t) = \mathbf{0} + \epsilon\mathbf{u}_1(x, y, z, t) + \cdots, \\ p(x, y, z, t) = p_0 + \epsilon p_1(x, y, z, t) + \cdots, \\ \rho(x, y, z, t) = \rho_0 + \epsilon\rho_1(x, y, z, t) + \cdots, \end{cases} \tag{8.5.54}$$

and $0 < \epsilon \ll 1$, where the dots refer to higher-order terms in ϵ. We are considering *small disturbances* around the equilibrium state. (The perturbation around the constant density is often referred to as the *condensation*.) Inserting (8.5.54) into Euler's equations and retaining only $O(\epsilon)$ terms gives

$$\frac{\partial\rho_1(x, y, z, t)}{\partial t} + \rho_0\nabla \cdot \mathbf{u}_1(x, y, z, t) = 0, \tag{8.5.55}$$

$$\frac{\partial\mathbf{u}_1(x, y, z, t)}{\partial t} + \frac{1}{\rho_0}\nabla p_1(x, y, z, t) = \mathbf{0}, \tag{8.5.56}$$

$$p_1(x, y, z, t) = f'(\rho_0)\rho_1(x, y, z, t), \tag{8.5.57}$$

as the *linearized versions* of the Euler equations.

From (8.5.56), we find, on integrating, that $\mathbf{u}_1(x, y, z, t) = \mathbf{u}_1(x, y, z, 0) - (1/\rho_0)\nabla\int_0^t p_1(x, y, z, t)\, dt$. We assume that initially, $\mathbf{u}_1(x, y, z, t) = \mathbf{0}$ or that $\mathbf{u}_1(x, y, z, t)$ is irrotational so that $\mathbf{u}_1(x, y, z, 0) = -\nabla\psi(x, y, z)$ [i.e., it can be

derived from a potential function $\psi(x, y, z)$]. (We recall the vector identity $\nabla \times \nabla\psi = 0$.) Then

$$\mathbf{u}_1(x, y, z, t) = -\nabla\left[\psi(x, y, z) - \frac{1}{\rho_0}\int_0^t p_1(x, y, z, t)\, dt\right] \equiv -\nabla\phi(x, y, z, t), \tag{8.5.58}$$

where $\phi(x, y, z, t)$ equals the term in brackets in (8.5.58). We have shown that if the flow is initially irrotational, it remains so for all time.

Putting (8.5.58) into (8.5.56) gives $\nabla\left[-\phi_t(x, y, z, t) + (1/\rho_0)p_1(x, y, z, t)\right] = \mathbf{0}$, and we may choose

$$p_1(x, y, z, t) = \rho_0\phi_t(x, y, z, t). \tag{8.5.59}$$

Then (8.5.57) implies that

$$\rho_1(x, y, z, t) = \left[\frac{\rho_0}{f'(\rho_0)}\right]\phi_t(x, y, z, t). \tag{8.5.60}$$

Inserting (8.5.58) and (8.5.60) into (8.5.55) yields

$$\phi_{tt}(x, y, z, t) - \left[f'(\rho_0)\right]\nabla^2\phi(x, y, z, t) = 0, \tag{8.5.61}$$

so that $\phi(x, y, z, t)$ satisfies the wave equation if $f'(\rho_0) > 0$.

In fact, for a *polytropic gas* we have

$$f(\rho) = A\rho^\gamma, \tag{8.5.62}$$

where A and γ are positive constants with $\gamma > 1$ and given as $\gamma = c_p/c_v$, with c_p and c_v equal to the specific heats of the gas at constant pressure and constant volume, respectively. We set

$$c^2(\rho) = f'(\rho), \tag{8.5.63}$$

so that $c(\rho_0) = \sqrt{f'(\rho_0)}$ in (8.5.61) represents the *propagation speed* for disturbances. We note that not only $\phi(x, y, z, t)$ but $p_1(x, y, z, t)$, $\rho_1(x, y, z, t)$, and each component of $\mathbf{u}_1(x, y, z, t)$ also satisfy the *wave equation* (8.5.61).

For the one-dimensional case we replace $\mathbf{u}(x, y, z, t)$ by $u(x, t)$ and obtain the one-dimensional wave equation for $\phi(x, t)$ with the general solution $\phi(x, t) = F(x - c_0t) + G(x + c_0t)$, with $c_0 = c(\rho_0)$. The function $\phi(x, t)$ must satisfy appropriate initial and boundary conditions and then $u_1(x, t)$, $p_1(x, t)$, and $\rho_1(x, t)$ can be expressed in terms of $\phi(x, t)$ by means of (8.5.58), (8.5.59), and (8.5.60). For example, if $\phi(x, t) = F(x - c_0t)$, so that only right-moving waves occur, we find that

$$\begin{cases} u_1(x, t) = -F'(x - c_0t), \\ p_1(x, t) = -c_0\rho_0F'(x - c_0t), \\ \rho_1(x, t) = -(\rho_0/c_0)F'(x - c_0t), \end{cases} \tag{8.5.64}$$

so that both $p_1(x, t)$ and $\rho_1(x, t)$ can be expressed in terms of $u_1(x, t)$ in the form

$$p_1(x, t) = c_0\rho_0u_1(x, t), \qquad \rho_1(x, t) = (\rho_0/c_0)u_1(x, t). \tag{8.5.65}$$

Euler's Equations for One-Dimensional Fluid Flow

The result (8.5.65) suggests that we look for special solutions of the full nonlinear system of Euler's equations of isentropic flow in one dimension in the form $p = p(u)$, $\rho = \rho(u)$, where $u(x,t)$ is the velocity. Using (8.5.63), the one-dimensional forms of Euler's equations are

$$\rho_t + \rho u_x + u\rho_x = 0, \tag{8.5.66}$$

$$u_t + uu_x + \frac{c^2(\rho)}{\rho}\rho_x = 0. \tag{8.5.67}$$

Since p has been eliminated from (8.5.67), we need only use $\rho = \rho(u)$. Inserting this into (8.5.66) and (8.5.84) gives

$$u_t + uu_x = -\frac{\rho}{\rho'}u_x, \quad u_t + uu_x = -\frac{c^2(\rho)\rho'}{\rho}u_x, \tag{8.5.68}$$

where $\rho' = d\rho/du$. Equating the right sides of these equations yields

$$(c(\rho)\rho')^2 = \rho^2, \quad \rho' = \pm\frac{\rho}{c(\rho)}. \tag{8.5.69}$$

Inserting (8.5.69) in the first equation of (8.5.68) gives, with $\hat{c}(u) = c[\rho(u)]$,

$$u_t + [u \pm \hat{c}(u)]u_x = 0, \tag{8.5.70}$$

a first order *quasilinear wave equation* studied in Chapter 2. Once $u(x,t)$ is found from (8.5.70), we may use the solution of (8.5.69) to specify $\rho(x,t)$.

Rather than deal with the general case, we consider a *polytropic gas* for which (8.5.62) is valid and choose the plus sign in (8.5.69)–(8.5.70). As a result $c^2(\rho) = dA\rho^\gamma/d\rho = \gamma A\rho^{\gamma-1}$ and (8.5.69) becomes $\sqrt{\gamma A}\,\rho^{(\gamma-3)/2}\rho' = 1$. This ODE is separable. Assuming that at $u = 0$ we have $\rho = \rho_0$ and $c(\rho)|_{u=0} = c(\rho_0) = c_0$, the solution given in terms of $\hat{c}(u)$ is

$$\hat{c}(u) = c_0 + \frac{\gamma - 1}{2}u, \tag{8.5.71}$$

from which ρ can be determined. Then (8.5.70) takes the form (with the plus sign chosen)

$$u_t + \left(c_0 + \frac{\gamma+1}{2}u\right)u_x = 0. \tag{8.5.72}$$

If the initial waveform of $u(x,t)$ at $t = 0$ is given as $u(x,0) = h(x)$, the result of Exercise 2.3.7 implies that

$$u = h\left[x - \left(c_0 + \frac{\gamma+1}{2}u\right)t\right] \tag{8.5.73}$$

is the implicit form of the solution of (8.5.71). This represents a wave traveling to the right for positive u, and according to the results of Section 2.3, each point x on the waveform $u(x, t)$ travels with speed

$$\frac{dx}{dt} = c_0 + \frac{\gamma + 1}{2} u = u + \left(c_0 + \frac{\gamma - 1}{2} u \right). \tag{8.5.74}$$

In fact, for the general case of (8.5.70) we have

$$\frac{dx}{dt} = u \pm \hat{c}(u). \tag{8.5.75}$$

We see from (8.5.74)–(8.5.75) that $\pm\hat{c}(u)$ represent an addition to the actual velocity u of the fluid at which disturbances propagate in the fluid. Therefore, $c = \sqrt{f'(\rho)}$ is called the *local speed of sound* of the fluid or gas. For the linearized equations of acoustics where a perturbation around a zero velocity was carried out, $c_0 = \sqrt{f'(\rho_0)}$ represents the (constant) speed of wave propagation or the *speed of sound* for the gas. By choosing the minus sign in (8.5.69), we may construct solutions that represent waves traveling to the left. If we assume that the velocity u is small and approximate $\hat{c}(u)$ by $\hat{c}(u_0) = c_0$, our results reduce to those for the equations of linear acoustics.

In general, solutions of (8.5.70) in the form $u = h[x - (u \pm \hat{c}(u))t]$ are known as *simple waves* and represent *unidirectional wave motion.* An alternative method for deriving simple waves is based on the use of the *Riemann invariants* for the one-dimensional Euler equations in either the adiabatic or isentropic case (see Example 3.5 and the exercises for this section). Simple waves play an important role in a number of problems in fluid dynamics, a particular case of which is the *piston problem.* This concerns the (one-dimensional) flow that results when a piston is pushed into or withdrawn from a long tube filled with gas at rest in a uniform state. (That is, the velocity $u = 0$ and the density, pressure, and entropy are all constant.) As a result, the flow can be assumed to be isentropic and can be described in terms of a simple wave under circumstances to be considered in the exercises.

However, as was found for solutions of quasilinear wave equations in Section 2.3 the simple wave solutions can break down after a finite time because they become multivalued or singular. In that case the solution can be continued beyond the breaking time by introducing *shock waves* as was done in Section 2.3. A set of shock conditions for the one-dimensional case known as the *Rankine-Hugoniot shock conditions* is considered in the exercises.

Navier-Stokes Equations for One-Dimensional Viscous Fluid Flow

An alternative approach to dealing with the breakdown of solutions of the Euler equations is to include the effects of viscosity and dependence on temperature variations that were neglected in our derivation of the Euler fluid flow equations. These effects introduce higher derivative terms into the system of fluid dynamics equations, and as we have seen on a number of occasions, the inclusion of higher derivative terms may have a smoothing effect on the solutions so that no breakdown occurs at all.

Rather than discuss the general case, we restrict ourselves to one dimension. Without presenting a derivation we write down the *Navier-Stokes equations* for *viscous fluid flow* in one dimension. We choose as the unknowns the velocity u, the density ρ, the pressure p, and the temperature T. Each of them is a function of x and t. We have

$$\rho_t + \rho u_x + u\rho_x = 0, \tag{8.5.76}$$

$$u_t + uu_x + \frac{1}{\rho} p_x = \frac{4}{3}\frac{\mu}{\rho} u_{xx}, \tag{8.5.77}$$

$$\rho c_v[T_t + uT_x] + pu_x = kT_{xx} + \frac{4}{3}\mu(u_x)^2, \tag{8.5.78}$$

$$p = R\rho T. \tag{8.5.79}$$

In these equations μ is the (constant) coefficient of viscosity, k is the coefficient of heat conduction, c_p and c_v are the specific heats at constant pressure and volume, respectively, and $R = c_p - c_v$ is the gas constant. We have assumed that there are no external forces or heat sources. (We note that in the literature the Navier-Stokes equations are often written in different forms.)

The first two equations, (8.5.76) and (8.5.77) are the *equations of continuity* and *conservation of momentum*. The new term in (8.5.77) represents forces due to viscous effects and vanishes when $\mu = 0$. The equation (8.5.78) characterizes the *conservation of energy*. If we neglect conductivity and viscosity (i.e., we put $k = \mu = 0$), we obtain $\rho c_v(T_t + uT_x) + pu_x = 0$, From (8.5.76) we can replace u_x by $-(1/\rho)(\rho_t + u\rho_x)$ and obtain $\rho(e_t + ue_x) - (p/\rho)(\rho_t + u\rho_x)$, where the *internal energy* e is given as $e = c_v T$. Using the material derivative, we can write the above as $de/dt - (p/\rho^2)\, d\rho/dt = de/dt + p\, d/dt\,(1/\rho) = 0$. Then the thermodynamic relation $Tds = de + p\, d\,(1/\rho)$, where s is the *entropy*, reduces the internal energy equation to $T\, ds/dt = 0$. We see that the entropy for individual fluid particles remains constant, so that this represents *adiabatic flow*. If at the initial time we have uniform entropy, it remains so for all time, and this corresponds to the isentropic flow considered.

The final equation (8.5.79) is an equation of state for a *perfect gas* that takes the place of (8.5.49). We do not discuss the relationship between these equations, say, if $\mu = k = 0$. In any case, the introduction of viscosity and heat conduction increases the order of the system of equations from first to second order by means of the addition of derivatives in x. It is shown in the exercises that for isentropic flow the *Euler equations* are a *hyperbolic system* whose characteristic curves for a given solution are (8.5.75), with $\hat{c}$ replaced by $c = c(\rho)$. The equations of adiabatic flow are also hyperbolic, as shown in Exercise 3.3.22. The full *Navier-Stokes equations* are of *parabolic type*. We do not discuss the solution of these equations here, but in Section 10.3 we present an approximation method that enables us to draw some conclusions about the effect of additional higher derivative terms in the Navier-Stokes equations on certain solutions of the fluid flow equations.

Steady Two-Dimensional Isentropic Flow

We conclude our discussion of fluid flow by examining the Euler equations in the case of *steady* (i.e., time independent) *two-dimensional isentropic flow*. With $u(x, y)$ and $v(x, y)$ as the velocity components, $\rho(x, y)$ as the density, and $c^2 = f'(\rho)$, we have from (8.5.42) and (8.5.47),

$$\rho u u_x + \rho v u_y + c^2 \rho_x = 0, \ \rho u v_x + \rho v v_y + c^2 \rho_y = 0, \ \rho u_x + \rho v_y + u\rho_x + v\rho_y = 0. \tag{8.5.80}$$

The system (8.5.80) can be written in matrix form as

$$A\mathbf{w}_x + B\mathbf{w}_y = \mathbf{0}, \tag{8.5.81}$$

where the matrices A and B and the vector $\mathbf{w}$ are given as

$$A = \begin{bmatrix} \rho u & 0 & c^2 \\ 0 & \rho u & 0 \\ \rho & 0 & u \end{bmatrix}, \quad B = \begin{bmatrix} \rho v & 0 & 0 \\ 0 & \rho v & c^2 \\ 0 & \rho & v \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} u \\ v \\ \rho \end{bmatrix}. \tag{8.5.82}$$

This first order quasilinear system may be classified according to the method of Section 3.3 for a fixed solution $\mathbf{w}$. The characteristic curves $y = y(x)$ are determined from [see (3.3.38)]

$$\det(B - y'A) = \begin{vmatrix} \rho(v - y'u) & 0 & -c^2 y' \\ 0 & \rho(v - y'u) & c^2 \\ -\rho y' & \rho & (v - y'u) \end{vmatrix} = 0, \tag{8.5.83}$$

and this gives us

$$[v - uy'][(v - uy')^2 - (1 + y'^2)c^2] = 0. \tag{8.5.84}$$

The first factor of (8.5.84) yields the characteristic curves

$$y' = \frac{dy}{dx} = \frac{v}{u}, \tag{8.5.85}$$

and these represent the *streamlines* of the flow, that is, the curves whose tangents are the velocity vectors $[u, v]$. The second factor in (8.5.84) can be written as

$$(c^2 - u^2)y'^2 + (2uv)y' + c^2 - v^2 = 0. \tag{8.5.86}$$

The roots y' of this quadratic equation may be characterized by means of the discriminant, which is $D = 4c^2(u^2 + v^2 - c^2)$, as is easy to show.

If $u^2 + v^2 > c^2$, both roots y' are real and distinct and neither is equal to the third root (8.5.85). Thus there are three real and distinct characteristics, and the system is strictly hyperbolic. Since c is the sound speed and $\sqrt{u^2 + v^2}$ is the speed of the fluid, this case corresponds to *supersonic flow*. If $u^2 + v^2 < c^2$, the corresponding

roots y' are complex. This case corresponds to *subsonic flow*. (According to our classification scheme, since not all the y' are complex, this is not an elliptic system.) If $u^2+v^2 = c^2$, there is a real double root, and the situation is complicated in that case. The flow may be termed *sonic flow* since the fluid speed and sound speed coincide.

The distinction between *subsonic* and *supersonic flow* can be brought out in another, simpler manner by linearizing the system (8.5.80). We set

$$u = u_0 + \epsilon u_1 + \cdots, \quad v = \epsilon v_1 + \cdots, \quad \rho = \rho_0 + \epsilon \rho_1 + \cdots, \tag{8.5.87}$$

where u_0 and ρ_0 are constants and $0 < \epsilon \ll 1$. The dots correspond to higher-order terms in ϵ. In (8.5.87) we are perturbing around a constant state in which there is a uniform flow with velocity $u_0 > 0$ parallel to the x-axis.

Inserting (8.5.87) in (8.5.80) and retaining only the terms of order ϵ yields

$$\rho_0 u_0 \frac{\partial u_1}{\partial x} + c_0^2 \frac{\partial \rho_1}{\partial x} = 0, \ \rho_0 u_0 \frac{\partial v_1}{\partial x} + c_0^2 \frac{\partial \rho_1}{\partial y} = 0, \ \rho_0 \frac{\partial u_1}{\partial x} + \rho_0 \frac{\partial v_1}{\partial y} + u_0 \frac{\partial \rho_1}{\partial x} = 0, \tag{8.5.88}$$

where $c_0 = c(\rho_0)$. From the third equation in (8.5.88) we have $\rho_0 \partial v_1/\partial y = -u_0 \partial \rho_1/\partial x - \rho_0 \partial u_1/\partial x = -u_0 \partial \rho_1/\partial x + (c_0^2/u_0)\partial \rho_1/\partial x$, in view of the first equation in (8.5.88). Since $\partial^2 v_1/\partial x \partial y = \partial^2 v_1/\partial y \partial x$, we have $\rho_0 \partial^2 v_1/\partial y \partial x = -(c_0^2/u_0)\partial^2 \rho_1/\partial y^2$, $\rho_0 \partial^2 v_1/\partial x \partial y = \left(c_0^2/u_0 - u_0\right) \partial^2 \rho_1/\partial x^2$, which follow from the above, and imply that

$$(1 - M^2) \frac{\partial^2 \rho_1(x,y)}{\partial x^2} + \frac{\partial^2 \rho_1(x,y)}{\partial y^2} = 0, \tag{8.5.89}$$

where the *Mach number* M is defined as $M = u_0/c_0$; that is, it is the ratio of the underlying constant velocity and the uniform sound speed $c_0 = c(\rho_0)$.

If $M > 1$, (8.5.89) is clearly of *hyperbolic type*. This corresponds to *supersonic flow* since $u_0 > c_0$ in the unperturbed state. If $M < 1$, (8.5.89) is of *elliptic type* and we have *subsonic flow* since $u_0 < c_0$. If $M = 1$, the equation is *parabolic* and we have *sonic flow* since $u_0 = c_0$.

We note that the characteristics corresponding to the streamlines for the nonlinear system [i.e., (8.5.85)] play no role in the linearized problem. In fact, if we insert (8.5.87) into (8.5.85), we obtain to leading order $y' = \epsilon v_1/u_0$, and from (8.5.86) we have $y'^2 = M^2 - 1$. Thus with $\epsilon = 0$ the characteristic equations are $y' = 0$ and $y' = \pm\sqrt{M^2 - 1}$. The solutions of $y' = 0$ (i.e., y = constant) are the streamlines of the steady uniform flow $u = u_0$ and $v = 0$. The characteristics of (8.5.89) (i.e., $y = \pm\sqrt{M^2 - 1}\, x$ = constant) are solutions of the equation $y' = \pm\sqrt{M^2 - 1}$.

Maxwell's Equations of Electromagnetic Theory

The *electric* and *magnetic field vectors* $\mathbf{E}$ and $\mathbf{H}$, respectively, which are studied in electromagnetic theory, are solutions of the fundamental system of PDEs known as *Maxwell's equations*. For the most part, each equation in the system characterizes an experimentally observed property of electric and magnetic fields. We do not present

a derivation of these equations but write out the equations in mks units and discuss some of their properties.

Maxwell's equations are given as follows:

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \tag{8.5.90}$$

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} + \mathbf{J}, \tag{8.5.91}$$

$$\nabla \cdot \mathbf{B} = 0, \tag{8.5.92}$$

$$\nabla \cdot \mathbf{D} = \rho. \tag{8.5.93}$$

In addition there are the *constitutive relations* that express $\mathbf{B}$, $\mathbf{D}$, and $\mathbf{J}$ in terms of $\mathbf{E}$ and $\mathbf{H}$. These equations are

$$\mathbf{D} = \epsilon \mathbf{E}, \quad \mathbf{B} = \mu \mathbf{H}, \quad \mathbf{J} = \sigma \mathbf{E}. \tag{8.5.94}$$

The vectors $\mathbf{D}$, $\mathbf{B}$, and $\mathbf{J}$ are the *electric displacement*, the *magnetic induction*, and the *conduction current density*, respectively. Further, ρ is the *density of electric charges*, ϵ is the *dielectric constant*, μ is the *magnetic permeability*, and σ is the *conductivity*.

We now introduce some simplifying assumptions that render this system more tractable. In problems of *electrostatics* we set all time derivatives equal to zero. From (8.5.90) we then find that $\nabla \times \mathbf{E} = \mathbf{0}$, so that $\mathbf{E}$ is an *irrotational field*. As was shown above, this implies that $\mathbf{E}$ can be derived from a potential, so that $\mathbf{E} = -\nabla\phi$, with ϕ given as the *electrostatic potential*. Further, if in (8.5.94) ϵ is a constant,we have from (8.5.93) $\nabla \cdot \mathbf{E} = \rho/\epsilon = -\nabla \cdot \nabla\phi$, so that $\nabla^2\phi = -\rho/\epsilon$, and ϕ is a solution of *Poisson's equation*. If the charge density ρ vanishes, ϕ satisfies *Laplace's equation*. If the medium is nonconducting, so that $\sigma = 0$ and $\mathbf{J} = 0$, and, in addition, μ is a constant, the same result is obtained for the magnetic field $\mathbf{H}$. However, since there are no magnetic charges, the *magnetostatic potential* satisfies *Laplace's equation*.

We return to the consideration of *electrodynamics*. If the medium is homogeneous (i.e., ϵ, μ, and σ are constant) and we also have $\rho = 0$, the following simplification of Maxwell's equations results. We apply the curl operator to (8.5.90) and obtain $\nabla \times \nabla \times \mathbf{E} = -\nabla \times \partial(\mu\mathbf{H})/\partial t = -\mu\, \partial(\nabla \times \mathbf{H})/\partial t$. Using the vector identity

$$\nabla \times \nabla \times \mathbf{E} = \nabla(\nabla \cdot \mathbf{E}) - \nabla^2\mathbf{E}, \tag{8.5.95}$$

and (8.5.91) for $\mathbf{H}$ gives

$$\nabla(\nabla \cdot \mathbf{E}) - \nabla^2\mathbf{E} = -(\mu\epsilon)\frac{\partial^2 \mathbf{E}}{\partial t^2} - (\mu\sigma)\frac{\partial \mathbf{E}}{\partial t}. \tag{8.5.96}$$

But $\nabla \cdot \mathbf{D} = \epsilon \nabla \cdot \mathbf{E} = 0$, in view of (8.5.93), so that

$$\nabla^2\mathbf{E} - (\epsilon\mu)\frac{\partial^2 \mathbf{E}}{\partial t^2} - (\sigma\mu)\frac{\partial \mathbf{E}}{\partial t} = \mathbf{0}, \tag{8.5.97}$$

with an identical equation satisfied by $\mathbf{H}$. Thus each Cartesian component of $\mathbf{E}$ and $\mathbf{H}$ satisfies a *telegrapher's equation* of the form

$$u_{tt} - c^2\nabla^2 u + \lambda u_t = 0, \tag{8.5.98}$$

where $c^2 = 1/\epsilon\mu$ and $\lambda = \sigma/\epsilon$. If the medium is nonconducting so that $\sigma = 0$, each component satisfies the *wave equation*. In either case, disturbances are propagated with the speed c. We note that if $\rho \neq 0$, we obtain an inhomogeneous form of (8.5.97).

It is also possible to introduce *scalar* and *vector potential functions* to treat the *time-dependent Maxwell's equations*. We assume that the medium is homogeneous as well as nonconducting, so that $\sigma = 0$ and $\mathbf{J} = \mathbf{0}$. The equation (8.5.92), that is, $\nabla \cdot \mathbf{B} = 0$, implies that $\mathbf{B}$ is a *solenoidal field* and it is shown in vector analysis that there exists a *vector potential* $\mathbf{A}$ such that

$$\mathbf{B} = \nabla \times \mathbf{A}. \tag{8.5.99}$$

Then (8.5.90) yields $\nabla \times \mathbf{E} = -\partial\mathbf{B}/\partial t = -\nabla \times \partial\mathbf{A}/\partial t$, and we obtain $\nabla \times [\mathbf{E} + \partial\mathbf{A}/\partial t] = \mathbf{0}$, so that $\mathbf{E} + \partial\mathbf{A}/\partial t$ is an irrotational vector. Consequently, it can be derived from a potential ϕ and we have

$$\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} = -\nabla\phi. \tag{8.5.100}$$

From the identity $\nabla \times \nabla\psi = \mathbf{0}$, we see that if $\mathbf{A}$ is replaced by $\mathbf{A} = \mathbf{A}_0 + \nabla\psi$, (8.5.99) will still be satisfied by $\mathbf{A}_0$. Further, in terms of ϕ we have from (8.5.100) $\mathbf{E} + \partial\mathbf{A}_0/\partial t = -\nabla\psi_t - \nabla\phi = -\nabla(\phi + \psi_t)$, so that a new scalar potential ϕ_0 can be defined as $\phi = \phi_0 - \psi_t$. To remove this ambiguity we impose the *Lorentz condition* or the *Lorentz gauge*,

$$\nabla \cdot \mathbf{A} = -(\epsilon\mu)\frac{\partial \phi}{\partial t}. \tag{8.5.101}$$

With this choice, $\mathbf{A}$ and ϕ both satisfy wave equations, as we now show.

From (8.5.100), (8.5.93), and (8.5.101) we have $\nabla \cdot [\mathbf{E} + \partial\mathbf{A}/\partial t] = \nabla \cdot \mathbf{E} + \partial(\nabla \cdot \mathbf{A})/\partial t = \rho/\epsilon - (\epsilon\mu)\partial^2\phi/\partial t^2 = -\nabla^2\phi$, so that, with $c^2 = 1/\epsilon\mu$, we have

$$c^2\nabla^2\phi - \phi_{tt} = -\frac{\rho}{\epsilon^2\mu}, \tag{8.5.102}$$

which is an *inhomogeneous wave equation*. To obtain an equation for $\mathbf{A}$ we use (8.5.91) in the form $\nabla \times (\mu\mathbf{H}) = \nabla \times \mathbf{B} = (\epsilon\mu)\,\partial\mathbf{E}/\partial t$, and insert the expression (8.5.99) to obtain $\nabla \times \mathbf{B} = \nabla \times \nabla \times \mathbf{A} = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2\mathbf{A} = -(\epsilon\mu)\left[\partial^2\mathbf{A}/\partial t^2 + \nabla(\partial\phi/\partial t)\right]$, where (8.5.95) and (8.5.100) were used. Rearranging terms and using the Lorentz condition (8.5.101) yields

$$(\epsilon\mu)\frac{\partial^2 \mathbf{A}}{\partial t^2} - \nabla^2\mathbf{A} = -\nabla\left[\nabla \cdot \mathbf{A} + (\epsilon\mu)\frac{\partial\phi}{\partial t}\right] = \mathbf{0}, \tag{8.5.103}$$

so that $\mathbf{A}$ satisfies the *homogeneous wave equation*. Once $\mathbf{A}$ and ϕ are specified, the electric and magnetic fields are given as

$$\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t} - \nabla\phi, \quad \mathbf{H} = \frac{1}{\mu}\nabla \times \mathbf{A}. \tag{8.5.104}$$

Maxwell's Equations in a Vacuum

We have shown the relationship between *Maxwell's equations* and *Laplace's* and *Poisson's equations* in the time-independent case and the *wave* and *telegrapher's equations* in the time-varying case. Many problems for Maxwell's equations can be reduced to a study of these simpler equations subject to appropriate initial and boundary conditions.

As a simple example we consider the *initial value problem* for *Maxwell's equations in a vacuum*. We have $\epsilon = \mu =$ constant, $\sigma = \rho = 0$, and $\mathbf{J} = \mathbf{0}$ in (8.5.90)–(8.5.94). Maxwell equations can then be written as

$$\nabla \times \mathbf{E}(\mathbf{x}, t) = -\mu \frac{\partial \mathbf{H}(\mathbf{x}, t)}{\partial t}, \quad \nabla \times \mathbf{H}(\mathbf{x}, t) = \epsilon \frac{\partial \mathbf{E}(\mathbf{x}, t)}{\partial t}, \tag{8.5.105}$$

$$\nabla \cdot \mathbf{E}(\mathbf{x}, t) = \nabla \cdot \mathbf{H}(\mathbf{x}, t) = 0. \tag{8.5.106}$$

These equations are valid throughout space, and initially we have

$$\mathbf{E}(\mathbf{x}, 0) = \mathbf{F}(\mathbf{x}), \qquad \mathbf{H}(\mathbf{x}, 0) = \mathbf{G}(\mathbf{x}). \tag{8.5.107}$$

The procedure that led to (8.5.97) can be applied to (8.5.105) to yield the equations

$$\nabla^2 \mathbf{E}(\mathbf{x}, t) - (\epsilon\mu)\mathbf{E}_{tt}(\mathbf{x}, t) = \mathbf{0}, \quad \nabla^2 \mathbf{H}(\mathbf{x}, t) - (\epsilon\mu)\mathbf{H}_{tt}(\mathbf{x}, t) = \mathbf{0}. \tag{8.5.108}$$

Each Cartesian component of $\mathbf{E}(\mathbf{x}, t)$ and $\mathbf{H}(\mathbf{x}, t)$ satisfies a *scalar wave equation*. However, we have only one initial condition for $\mathbf{E}(\mathbf{x}, t)$ and $\mathbf{H}(\mathbf{x}, t)$ in (8.5.106) and we must also prescribe $\mathbf{E}_t(\mathbf{x}, t)$ and $\mathbf{H}_t(\mathbf{x}, t)$ initially to obtain a unique solution. From (8.5.105) we find that

$$\begin{cases} \partial \mathbf{E}(\mathbf{x}, 0)/\partial t = (1/\epsilon)\nabla \times \mathbf{H}(\mathbf{x}, 0) = (1/\epsilon)\nabla \times \mathbf{G}(\mathbf{x}), \\ \partial \mathbf{H}(\mathbf{x}, 0)/\partial t = -(1/\mu)\nabla \times \mathbf{E}(\mathbf{x}, 0) = -(1/\mu)\nabla \times \mathbf{F}(\mathbf{x}). \end{cases} \tag{8.5.109}$$

The initial value problem for the scalar wave equation has been solved previously in the text in one, two, and three space dimensions. Thus $\mathbf{E}(\mathbf{x}, t)$ and $\mathbf{H}(\mathbf{x}, t)$ can be determined. To complete the verification of the solution we must show that (8.5.106) is satisfied for $t > 0$ provided that it was true at the time $t = 0$. Thus we must assume that the initial fields $\mathbf{F}(\mathbf{x})$ and $\mathbf{G}(\mathbf{x})$ are such that

$$\nabla \cdot \mathbf{F}(\mathbf{x}) = \nabla \cdot \mathbf{G}(\mathbf{x}) = 0. \tag{8.5.110}$$

Since $\nabla \cdot \nabla \times \mathbf{A}(\mathbf{x}, t) = \mathbf{0}$ for any vector $\mathbf{A}(\mathbf{x}, t)$, we have $0 = \nabla \cdot \nabla \times \mathbf{E}(\mathbf{x}, t) = -\mu \nabla \cdot \mathbf{H}_t(\mathbf{x}, t) = -\mu \partial(\nabla \cdot \mathbf{H}(\mathbf{x}, t))/\partial t$, $0 = \nabla \cdot \nabla \times \mathbf{H}(\mathbf{x}, t) = \epsilon \nabla \cdot \mathbf{E}_t(\mathbf{x}, t) = \epsilon \partial(\nabla \cdot \mathbf{E}(\mathbf{x}, t))/\partial t$, on using (8.5.105). Thus $\nabla \cdot \mathbf{E}(\mathbf{x}, t)$ and $\nabla \cdot \mathbf{H}(\mathbf{x}, t)$ are constant in time and since they vanish at $t = 0$, they do so for all time.

For the sake of concreteness we consider a specific problem in one dimension. Let

$$\mathbf{E}(x, 0) = \mathbf{F}(x) = \sin(x)\mathbf{j}, \qquad \mathbf{H}(x, 0) = \mathbf{G}(x) = \cos(x)\mathbf{k}. \tag{8.5.111}$$

Then (8.5.110) is satisfied and $\nabla \times \mathbf{F}(x) = \cos(x)\mathbf{k}$, $\nabla \times \mathbf{G}(x) = \sin(x)\mathbf{j}$, so that $\mathbf{E}_t(x,0) = (1/\epsilon)\sin(x)\mathbf{j}$, $\mathbf{H}_t(x,0) = -(1/\mu)\cos(x)\mathbf{k}$.

Each component of $\mathbf{E}(x,t)$ and $\mathbf{H}(x,t)$ satisfies the one-dimensional wave equation $u_{xx}(x,t) - (\epsilon\mu)u_{tt}(x,t) = 0$. The y-component of $\mathbf{E}(x,t)$ has the initial conditions $u(x,0) = \sin(x)$, $u_t(x,0) = (1/\epsilon)\sin(x)$, while the z-component of $\mathbf{H}(x,t)$ has the initial conditions $u(x,0) = \cos(x)$, $u_t(x,0) = -(1/\mu)\cos(x)$. All other components have zero initial conditions and, therefore, vanish for all time. Using d'Alembert's solution of the wave equation, we easily obtain (with $c^2 = 1/\epsilon\mu$) the standing wave representations

$$\mathbf{E}(x,t) = \sin(x)\left(\cos(ct) + \frac{\sin(ct)}{c\epsilon}\right)\mathbf{j},\ \mathbf{H}(x,t) = \cos(x)\left(\cos(ct) - \frac{\sin(ct)}{c\mu}\right)\mathbf{k}. \tag{8.5.112}$$

Clearly, $\mathbf{E}(x,t)$ and $\mathbf{H}(x,t)$ satisfy the initial conditions (8.5.111). Further, $\nabla \times \mathbf{E}(x,t) = \cos(x)\left[\cos(ct) + (1/c\epsilon)\sin(ct)\right]\mathbf{k}$, and $-\mu\ \partial\mathbf{H}(x,t)/\partial t = \cos(x)\left[\mu c\sin(ct) + \cos(ct)\right]\mathbf{k}$, so that (8.5.105) is satisfied since $\mu c = 1/c\epsilon$. It can easily be verified that (8.5.106) are also satisfied. Finally, it may be noted that although the fields $\mathbf{E}(x,t)$ and $\mathbf{H}(x,t)$ propagate along the x-direction, the nonzero field components lie in the transverse y- and z-directions.

Navier's Equation of Elasticity Theory

Let the vector $\mathbf{u}(\mathbf{x},t)$ represent the displacement of a point $\mathbf{x}$ in an *elastic medium* at the time t. Assuming that the medium is *isotropic*, the equation of motion satisfied by $\mathbf{u}(\mathbf{x},t)$, which is known as *Navier's equation*, is

$$\rho\frac{\partial^2\mathbf{u}(\mathbf{x},t)}{\partial t^2} = (\lambda+2\mu)\nabla(\nabla\cdot\mathbf{u}(\mathbf{x},t)) - \mu\nabla\times\nabla\times\mathbf{u}(\mathbf{x},t) + \mathbf{F}(\mathbf{x},t). \tag{8.5.113}$$

Here ρ is the density of the medium, λ and μ are the *Lamé constants*, which depend on the properties of the medium, and $F(\mathbf{x},t)$ represents the external forces. Using the vector identity (8.5.95), we can put (8.5.113) into the equivalent form

$$\rho\frac{\partial^2\mathbf{u}(\mathbf{x},t)}{\partial t^2} = (\lambda+\mu)\nabla(\nabla\cdot\mathbf{u}(\mathbf{x},t)) + \mu\nabla^2\mathbf{u}(\mathbf{x},t) + \mathbf{F}(\mathbf{x},t). \tag{8.5.114}$$

We assume in our discussion that ρ, λ, and μ are constants.

We begin our analysis of the equations of elasticity by assuming that $\mathbf{u} = \mathbf{u}(x,t)$, so that the problem is one-dimensional. Let $\mathbf{u}(x,t) = u(x,t)\mathbf{i} + v(x,t)\mathbf{j} + w(x,t)\mathbf{k}$. The divergence $\nabla\cdot\mathbf{u}(x,t)$ has the form $\nabla\cdot\mathbf{u}(x,t) = \partial u(x,t)/\partial x$, since $\mathbf{u}$ is independent of y and z. Thus (8.5.114) becomes

$$\rho\frac{\partial^2\mathbf{u}(x,t)}{\partial t^2} = (\lambda+\mu)\frac{\partial^2 u(x,t)}{\partial x^2}\mathbf{i} + \mu\nabla^2\mathbf{u}(x,t) + \mathbf{F}(x,t). \tag{8.5.115}$$

With $\mathbf{F}(x,t) = F_1(x,t)\mathbf{i} + F_2(x,t)\mathbf{j} + F_3(x,t)\mathbf{k}$, we can uncouple the three equations (8.5.115) for the Cartesian components of $\mathbf{u}(x,t)$ and obtain

$$\rho u_{tt}(x,t) = (\lambda + 2\mu)u_{xx}(x,t) + F_1(x,t), \tag{8.5.116}$$

$$\rho v_{tt}(x,t) = \mu v_{xx}(x,t) + F_2(x,t), \quad \rho w_{tt}(x,t) = \mu w_{xx}(x,t) + F_3(x,t). \tag{8.5.117}$$

Thus each component of $\mathbf{u}(x,t)$ satisfies a wave equation in one dimension. However, the speed of wave propagation for $u(x,t)$ equals $[(\lambda + 2\mu)/\rho]^{1/2}$, while the corresponding speed for $v(x,t)$ and $w(x,t)$ is $(\mu/\rho)^{1/2}$.

These results show that there are two modes of wave propagation for elastic waves in one dimension. Since the waves associated with $u(x,t)$ have the property that the displacement is in the direction of wave propagation, they are called *longitudinal waves* and travel with the speed $[(\lambda + 2\mu)/\rho]^{1/2}$. The second wave motion is associated with $v(x,t)$ and $w(x,t)$ and thus corresponds to displacements in a plane perpendicular to the direction of wave propagation. Therefore, they are called *transverse waves* and travel with the speed $(\mu/\rho)^{1/2}$. Since λ, μ, and ρ are all positive, we see that the longitudinal waves have greater speed than the transverse waves. (The longitudinal waves are often called *compression waves* and the transverse waves are called *shear waves*.)

It may be noted that if we put $\mathbf{u}(x,t) = u(x,t)\mathbf{i} + [v(x,t)\mathbf{j} + w(x,t)\mathbf{k}] \equiv u(x,t)\mathbf{i} + \mathbf{V}(x,t)$, we have (since u and $\mathbf{V}$ are functions of x and t only) $\nabla \times \mathbf{u}(x,t) = \nabla \times \mathbf{V}(x,t)$, $\nabla \times (u(x,t)\mathbf{i}) = \mathbf{0}$, $\nabla \cdot \mathbf{u}(x,t) = \nabla \cdot (u(x,t)\mathbf{i})$, $\nabla \cdot \mathbf{V}(x,t) = 0$, as is easily verified. This suggests that in the general case, we try to decompose the displacement vector $\mathbf{u}(\mathbf{x},t)$ into two parts, one involving $\nabla \cdot \mathbf{u}(\mathbf{x},t)$ and the other $\nabla \times \mathbf{u}(\mathbf{x},t)$. (It is shown in vector analysis that such a decomposition is always possible for a vector field. That is, it can be decomposed into a sum of two vectors, one of which has zero divergence whereas the other has zero curl.)

To achieve this decomposition, we multiply (8.5.113) by $\nabla\cdot$ and then by $\nabla\times$; that is, we take the divergence and curl of these equations. We have

$$\rho \frac{\partial^2(\nabla \cdot \mathbf{u})}{\partial t^2} = (\lambda + 2\mu)\nabla^2(\nabla \cdot \mathbf{u}) + \nabla \cdot \mathbf{F} \tag{8.5.118}$$

since the divergence of a curl vanishes. Also,

$$\rho \frac{\partial^2(\nabla \times \mathbf{u})}{\partial t^2} = \mu\nabla^2(\nabla \times \mathbf{u}) + \nabla \times \mathbf{F} \tag{8.5.119}$$

because of the vector identities $\nabla \times [\nabla \times \nabla \times \mathbf{u}] = -\nabla^2(\nabla \times \mathbf{u}) + \nabla(\nabla \cdot \nabla \times \mathbf{u}) = -\nabla^2(\nabla \times \mathbf{u})$, $\nabla \cdot \nabla \times \mathbf{u} = 0$, and $\nabla \times \phi = \mathbf{0}$.

We have shown that in the general case, the elasticity equations can be reduced to the study of two wave equations for $\nabla \cdot \mathbf{u}(\mathbf{x},t)$ and $\nabla \times \mathbf{u}(\mathbf{x},t)$. The term $\nabla \cdot \mathbf{u}(\mathbf{x},t)$ characterizes the compression and expansion of the body and its waves travel with the speed $[(\lambda + 2\mu)/\rho]^{1/2}$. The term $\nabla \times \mathbf{u}(\mathbf{x},t)$ characterizes the distortion of the body and its waves travel with the speed $(\mu/\rho)^{1/2}$. In dealing with initial value problems for elastic wave propagation with no boundaries present, this decomposition can be used to great advantage since we know how to solve the initial value problem for the wave equation. However, if there are boundaries, the boundary conditions couple the components of $\mathbf{u}(\mathbf{x},t)$ and the problems are more complicated.

Exercises 8.5

8.5.1. Show that if $f(x) = \delta(x)$ (the Dirac delta function) and $g(x) = 0$ in (8.5.2), the asymptotic solution (8.5.8) of (8.5.1) is $u(x,t) \approx (1/\sqrt{4\pi ct})\cos(x^2/4ct - \pi/4)$ and this is the exact solution of the initial value problem.

8.5.2. Use the result of Exercise 8.5.1 to show that the solution of (8.5.1) with $u(x,0) = f(x)$, $u_t(x,0) = 0$ is $u(x,t) = \frac{1}{\sqrt{4\pi ct}} \int_{-\infty}^{\infty} f(s)\cos\left(\frac{(x-s)^2}{4ct} - \frac{\pi}{4}\right) ds$.

8.5.3. A rod is said to be simply supported at $x = 0$ if $u(0,t) = u_{xx}(0,t) = 0$. Use the Fourier sine transform to solve the following problem for the vibration of a semi-infinite rod: $u_{tt}(x,t) + c^2 u_{xxxx}(x,t) = 0$, $0 < x < \infty$, $t > 0$, $u(x,0) = f(x)$, $u_t(x,0) = g(x)$, $0 < x < \infty$, $u(0,t) = u_{xx}(0,t) = 0$, $t > 0$.

8.5.4. Apply the Fourier cosine transform to solve the following problem: $u_{tt}(x,t) + c^2 u_{xxxx}(x,t) = 0$, $0 < x < \infty$, $t > 0$, $u(x,0) = f(x)$, $u_t(x,0) = g(x)$, $0 < x < \infty$, $u_x(0,t) = u_{xxx}(0,t) = 0$, $t > 0$.

8.5.5. The vibration of a finite rod simply supported at both ends is determined by the solution of the following problem: $u_{tt}(x,t) + c^2 u_{xxxx}(x,t) = 0$, $0 < x < l$, $t > 0$, $u(x,0) = f(x)$, $u_t(x,0) = g(x)$, $0 < x < l$, $u(0,t) = u_{xx}(0,t) = u(l,t) = u_{xx}(l,t) = 0$, $t > 0$. Obtain the solution by using separation of variables.

8.5.6. A rod is free or unsupported at a point if $u_{xx}(x,t)$ and $u_{xxx}(x,t)$ vanish there. Use separation of variables to solve the following problem for the vibration of a rod clamped at one end and free at the other end: $u_{tt}(x,t) + c^2 u_{xxxx}(x,t) = 0$, $0 < x < l$, $t > 0$, $u(x,0) = f(x)$, $u_t(x,0) = g(x)$, $0 < x < l$, $u(0,t) = u_x(0,t) = u_{xx}(l,t) = u_{xxx}(l,t) = 0$, $t > 0$.

8.5.7. Consider the eigenvalue problem $v''''(x) - \lambda^2 v(x) = 0$, $0 < x < l$, with the following boundary conditions: (a) $v(0) = v'(0) = v(l) = v'(l) = 0$; (b) $v(0) = v''(0) = v(l) = v''(l) = 0$; (c) $v'(0) = v'''(0) = v'(l) = v'''(l) = 0$; (d) $v(0) = v'(0) = v''(l) = v'''(l) = 0$. Let λ_n and λ_m be distinct eigenvalues for each of these problems and $v_n(x)$ and $v_m(x)$ be the corresponding eigenfunctions. Use the method of Section 4.3 to show that the eigenfunctions are orthogonal with respect to the following inner product $(v_n(x), v_m(x)) = \int_0^l v_n(x)v_m(x)\,dx = 0$, $n \neq m$.

8.5.8. Construct the Green's function $K(x;\xi)$ for the problems: $\partial^4 K(x;\xi)/\partial x^4 = -\delta(x-\xi)$, $0 < x, \xi < l$, where $K(x;\xi)$ satisfies the boundary conditions (a), (b), (c), or (d) given in Exercise 8.5.7. *Hint*: $K(x;\xi)$ satisfies a homogeneous equation at $x \neq \xi$, and the jump in $\partial^3 K(x;\xi)/\partial x^3$ at $x = \xi$ is -1. In case (c), a modified Green's function must be constructed.

8.5.9. Show how the finite sine transform method can be used to solve the problem in Exercise 8.5.5 if the equation and the boundary data are inhomogeneous.

8.5.10. A simply supported plate satisfies the boundary conditions $u(x,y,t) = 0$ and $\partial^2 u(x,y,t)/\partial n^2 = 0$ at the edge. [Here $\partial^2 u(x,y,t)/\partial n^2$ is the second exterior normal derivative.] Use separation of variables to solve the problem of the vibrating simply supported rectangular plate. That is, solve the following problem for $u(x,y,t)$:

$u_{tt}(x,y,t) + c^2\nabla^2\nabla^2 u(x,y,t) = 0,\ 0 < x < l,\ 0 < y < \hat{l},\ t > 0, u(x,y,0) = f(x,y),\ u_t(x,y,0) = g(x,y),\ 0 < x < l,\ 0 < y < \hat{l}, u(0,y,t) = u_{xx}(0,y,t) = u(l,y,t) = u_{xx}(l,y,t) = 0,\ 0 < y < \hat{l},\ t > 0,\ u(x,0,t) = u_{yy}(x,0,t) = u(x,\hat{l},t) = u_{yy}(x,\hat{l},t) = 0,\ 0 < x < l,\ t > 0$. *Hint*: Obtain (8.5.21) and set $A_n(x) = \sin[(\pi n/l)x]$ to satisfy the conditions at $x = 0$ and $x = l$. Conclude that $B_m(y) = \sin[(\pi m/\hat{l})y]$, and obtain the solution in terms of a double Fourier sine series.

8.5.11. Use the Fourier transform to solve the Cauchy problem for the vibrating plate equation $u_{tt}(x,y,t) + c^2\nabla^2\nabla^2 u(x,y,t) = 0,\ -\infty < x,y < \infty,\ t > 0,\ u(x,y,0) = f(x,y),\ u_t(x,y,0) = g(x,y),\ -\infty < x,y < \infty$ with appropriate conditions at infinity.

8.5.12. Use Green's theorem (8.5.38) to show that eigenfunctions corresponding to different eigenvalues of the problem (8.5.19)–(8.5.20) are orthogonal.

8.5.13. Verify that (8.5.32) satisfies the biharmonic equation.

8.5.14. Solve the boundary value problem $\nabla^2\nabla^2 u(r,\theta) = 0,\ x^2 + y^2 < R^2,\ u(R,\theta) = f(\theta),\ \partial u(R,\theta)/\partial r = g(\theta)$ by using a Fourier series in the angular variable θ.

8.5.15. Solve the problem of Exercise 8.5.14 using the method presented in the text if the boundary conditions are $f(\theta) = 1,\ g(\theta) = \sin^3\theta$.

8.5.16. Show that (8.5.37) satisfies the biharmonic equation if $r \neq 0$.

8.5.17. Solve the following problem by looking for a solution of the form $u = u(r)$: $\nabla^2\nabla^2 u(r,\theta) = F_0,\ x^2 + y^2 < R^2, u(R,\theta) = 0,\ \partial u(R,\theta)/\partial r = 0$, where F_0 is a constant.

8.5.18. Obtain the solution of the problem $\nabla^2\nabla^2 u(x,y) = -\delta(x)\delta(y),\ x^2 + y^2 < R^2, u(x,y) = \partial u(x,y)/\partial n = 0,\ x^2 + y^2 = R^2$ by looking for a solution in the form $u(x,y) = -(1/8\pi)r^2 \log r + v(r)$, where $r^2 = x^2 + y^2$.

8.5.19. Take the total or material derivative of the equation of state (8.5.49) with respect to t and obtain $dp/dt = c^2\, d\rho/dt$ on using (8.5.48) and defining $c^2 = \partial f(\rho,s)/\partial\rho$. Use this result to replace (8.5.48) by $p_t(x,y,z,t) + (\mathbf{u}(x,y,z,t) \cdot \nabla)p(x,y,z,t) - c^2[\rho_t(x,y,z,t) + (\mathbf{u}(x,y,z,t) \cdot \nabla)\rho(x,y,z,t)] = 0$. Show that the one-dimensional version of the modified form of Euler's equations for adiabatic flow is given as in Exercise 3.3.22.

8.5.20. Use the equation $T\, ds/dt = de/dt + p\, d/dt\,(1/\rho) = 0$, together with the equations (8.5.42) and (8.5.47), to replace the (entropy) equation (8.5.48) by the equation

$$\rho(x,y,z,t)\frac{d}{dt}\left(e(x,y,z,t) + \frac{1}{2}|\mathbf{u}(x,y,z,t)|^2\right) + \nabla\cdot(p(x,y,z,t)\mathbf{u}(x,y,z,t)) = \mathbf{0}.$$

This the equation of *conservation of energy*.

8.5.21. Use the result of Exercise 8.5.20 to express the one-dimensional form of Euler's equations for adiabatic flow in conservation form (see Section 2.3), $\rho_t(x,t) +$

$(\rho(x,t)u(x,t))_x = 0, (\rho(x,t)u(x,t))_t + (\rho(x,t)u^2(x,t) + p(x,t))_x = 0, [\frac{1}{2}\rho(x,t)u^2(x,t) + \rho(x,t)e(x,t)]_t + [\left(\frac{1}{2}\rho(x,t)u^2(x,t) + \rho(x,t)e(x,t)\right)u(x,t) + p(x,t)u(x,t)]_x = 0$. It is from this system of conservation laws that the *Rankine-Hugoniot shock conditions* are derived, in the manner of Section 2.3. We do not write them down.

8.5.22. Multiply (8.5.66) by $c(\rho)$, add (8.5.67) to it, and then subtract (8.5.67) from it to obtain $c\,[\rho_t + (u \pm c)\rho_x] \pm \rho\,[u_t + (u \pm c)u_x] = 0$. Conclude from these equations that $c\,d\rho/dt + \rho\,du/dt = 0$ on $dx/dt = u + c$, and $c\,d\rho/dt - \rho\,du/dt = 0$ on $dx/dt = u - c$, where $d/dt = \partial/\partial t + (dx/dt)\partial/\partial x$. Since $c = c(\rho)$, show that these equations can be written as $\int^\rho c(\rho)/\rho\,d\rho + u =$ constant on $dx/dt = u + c$, and $\int^\rho c(\rho)/\rho\,d\rho - u =$ constant on $dx/dt = u - c$. They are the *Riemann invariants*.

8.5.23. Show that if one of the Riemann invariants in Exercise 8.5.22 is set identically constant for all x and t (i.e., it is not just constant on individual characteristics), we must have $\rho = \rho(u)$. Conclude that the other Riemann invariant has the form of a simple wave.

8.5.24. In the case of a polytropic gas [see (8.5.62)], show that the Riemann invariants of Exercise 8.5.23 take the form $2c/(\gamma - 1) \pm u =$ constant on $dx/dt = u \pm c$.

8.5.25. Consider the problem of the withdrawal of a piston from a polytropic gas initially at rest. The problem is one-dimensional, and at the time $t = 0$ the gas is located at $x \geq 0$ with velocity $u = 0$, a constant propagation speed $c = c_0$, and constant entropy $s = s_0$. The path of the withdrawing piston is $x = g(t)$, with $g(0) = 0$ and $g'(t) < 0$. Use the Riemann invariants of Exercise 8.5.24 and conclude that $2c/(\gamma - 1) - u = 2c_0/(\gamma - 1)$ since each of the characteristics associated with this invariant issues from the positive x-axis. Show that $u = 0$ and $c = c_0$ in the sector $c_0 t < x < \infty$. In the region between the piston curve and the line $x = c_0 t$, show that $u =$ constant on $dx/dt = c_0 + [(\gamma + 1)/2]u$, that is, the solution is a simple wave there. Since the fluid velocity on the piston equals the piston velocity, we have the boundary condition $u = g'(t)$ on $x = g(t)$. Obtain the equation of the characteristic $x'(t) = c_0 + [(\gamma + 1)/2]u$ in the parametric form $x = g(\tau) + [c_0 + (\gamma + 1/2)\,g'(\tau)]\,(t - \tau)$. The solution is given as $u = g'(\tau)$, $c = c_0 + [(\gamma - 1)/2]\,g'(\tau)$, $s = s_0$, with τ determined as a function of x and t from the equation for the characteristic. We do not discuss the formation of shocks.

8.5.26. Consider the *Hertz vector* $\boldsymbol{\pi}(\mathbf{x}, t)$ and express the scalar and vector potentials $\phi(\mathbf{x}, t)$ and $\mathbf{A}(\mathbf{x}, t)$ for Maxwell's equations as $\phi(\mathbf{x}, t) = -\nabla \cdot \boldsymbol{\pi}(\mathbf{x}, t)$, $\mathbf{A}(\mathbf{x}, t) = (1/c^2)\partial\boldsymbol{\pi}(\mathbf{x}, t)/\partial t$, where $\epsilon\mu = 1/c^2$. Show that the Lorentz condition (8.5.101) is thereby satisfied. Express $\mathbf{E}(\mathbf{x}, t)$ and $\mathbf{H}$ in terms of $\boldsymbol{\pi}(\mathbf{x}, t)$, and if $\rho(\mathbf{x}, t) = \sigma(\mathbf{x}, t) = 0$, show that $\boldsymbol{\pi}(\mathbf{x}, t)$ satisfies the equation $\boldsymbol{\pi}_{tt}(\mathbf{x}, t) - c^2\nabla^2\boldsymbol{\pi}(\mathbf{x}, t) = \mathbf{0}$.

8.5.27. Determine the form taken by Maxwell's equations if $\rho(\mathbf{x}, t) = \sigma(\mathbf{x}, t) = 0$, $\epsilon(\mathbf{x}, t)$, $\mu(\mathbf{x}, t)$ are constant, and $\mathbf{E}(\mathbf{x}, t)$ and $\mathbf{H}(\mathbf{x}, t)$ are given as $\mathbf{E}(\mathbf{x}, t) = \mathbf{e}(\mathbf{x})e^{-i\omega t}$, $\mathbf{H}(\mathbf{x}, t) = \mathbf{h}(\mathbf{x})e^{-i\omega t}$, where ω is a constant.

8.5.28. Verify that $\nabla \cdot \mathbf{F}(x) = \nabla \cdot \mathbf{G}(x) = 0$ for the data (8.5.111) and show that the solution (8.5.112) also satisfies $\nabla \cdot \mathbf{E}(x, t) = \nabla \cdot \mathbf{H}(x, t) = 0$.

8.5.29. Formulate Exercise 8.5.26 for the Hertz vector if the problem depends only on x and t. Assume that the Hertz vector $\boldsymbol{\pi}(x,t)$ is given initially as $\boldsymbol{\pi}(x,0) = (\cos x)\mathbf{i} + \mathbf{j}$, $\boldsymbol{\pi}_t(x,0) = \mathbf{k}$, $-\infty < x < \infty$. Determine $\boldsymbol{\pi}(x,t)$, $\mathbf{E}(x,t)$, and $\mathbf{H}(x,t)$ for $t > 0$ by solving the Cauchy problem for $\boldsymbol{\pi}(x,t)$.

8.5.30. Let the Hertz vector $\boldsymbol{\pi}(\mathbf{x},t)$ be given as $\boldsymbol{\pi}(\mathbf{x},t) = \boldsymbol{\alpha}(\mathbf{x})e^{-i\omega t}$ (see Exercise 8.5.26). Assuming that the conditions given in Exercise 8.5.27 apply, show that $\nabla^2\boldsymbol{\alpha}(\mathbf{x}) + k^2\boldsymbol{\alpha}(\mathbf{x}) = \mathbf{0}$, where $k = \omega/c$ and $\mathbf{e}(\mathbf{x}) = \nabla(\nabla\cdot\boldsymbol{\alpha}(\mathbf{x})) + k^2\boldsymbol{\alpha}(\mathbf{x})$, $\mathbf{h}(\mathbf{x}) = -(ik/\mu c)\,\nabla\times\boldsymbol{\alpha}(\mathbf{x})$.

8.5.31. To study the possibility that electromagnetic waves can propagate in a waveguide, we consider Maxwell's equations in a cylindrical region whose walls are perfectly conducting. The interior of the guide is such that ϵ and μ are constant and $\sigma = 0$. In addition, there are no sources so that $\rho = 0$ and $\mathbf{E}(x,y,z,t)$ and $\mathbf{H}(x,y,z,t)$ are assumed to have harmonic time dependence, as in Exercise 8.5.27. The tangential component of $\mathbf{E}(x,y,z,t)$ vanishes on the waveguide wall. Assume that the cylinder has its generators parallel to the z-axis and show that if we set the Hertz vector $\boldsymbol{\pi}(x,y,z,t)$ equal to $\boldsymbol{\pi}(x,y,z,t) = \boldsymbol{\alpha}(x,y,z)e^{-i\omega t} = u(x,y,z)e^{-i\omega t}\mathbf{k}$, the tangential component of $\mathbf{E}(x,y,z,t)$ vanishes if $u(x,y,z) = 0$ on the cylinder wall (use Exercise 8.5.30). Also, conclude that $u(x,y,z)$ satisfies the reduced wave equation $\nabla^2u(x,y,z) + k^2u(x,y,z) = 0$, $k = \omega/c$. Let $u(x,y,z) = M(x,y)v(z)$ and determine $M(x,y)$ from the eigenvalue problem $\nabla^2M(x,y) + \lambda M(x,y) = 0$, $(x,y) \in G$, $M(x,y) = 0$, $(x,y) \in \partial G$, where G is a perpendicular cross section of the cylinder. Let λ_n and $M_n(x,y)$ be the eigenvalues and eigenfunctions for this problem with $\lambda_1 \le \lambda_2 \le \cdots$. Then $v_n(z)$ must satisfy $v_n''(z) + (k^2 - \lambda_n)v_n(z) = 0$. Considering only waves traveling in the positive z-direction, show that the Hertz vectors $\boldsymbol{\pi}_n(x,y,z,t)$ must take the form $\boldsymbol{\pi}_n(x,y,z,t) = a_nM_n(x,y)\exp[i(\sigma_nz - \omega t)]\mathbf{k}$, where $\sigma_n = \sqrt{k^2 - \lambda_n}$ and a_n is a constant. Determine that no traveling waves exist in the waveguide if $k^2 < \lambda_1$ and that only a finite number of such waves exist if $\lambda_{n-1} < k^2 < \lambda_n$. [With $k^2 < \lambda_n$, the field $\boldsymbol{\pi}_n(x,y,z,t)$ decays as $z \to \infty$.]

8.5.32. Since $k = \omega/c$, show that if $\omega^2 < \lambda_1c^2$ for a particular waveguide where λ_1 is determined as in Exercise 8.5.31, no traveling waves can exist in the waveguide. Thus $\omega = \sqrt{\lambda_1}\,c$ is termed a *cutoff frequency*. Determine the cutoff frequencies for the waveguides with the following cross sections: (a) $0 \le x \le l$, $\quad 0 \le y \le \hat{l}$; (b) $x^2 + y^2 \le a^2$.

8.5.33. Solve the initial value problem for the one-dimensional Navier equation (8.5.114) with $\mathbf{F} = \mathbf{0}$ and the initial data $\mathbf{u}(x,0) = (\cos x)\mathbf{i} + \mathbf{j} + (\sin x)\mathbf{k}$, $\mathbf{u}_t(x,0) = x\mathbf{i} + e^{-x}\mathbf{j} + \mathbf{k}$.

8.5.34. Show that if the equation (8.5.118) is solved for $\nabla\cdot\mathbf{u}(\mathbf{x},t)$, we may obtain $\mathbf{u}(\mathbf{x},t)$ from (8.5.114) by treating the term involving $\nabla\cdot\mathbf{u}(\mathbf{x},t)$ as an additional inhomogeneous term. Indicate how we may express the solution of the initial value problem for (8.5.114) with $\mathbf{u}(\mathbf{x},0)$ and $\mathbf{u}_t(\mathbf{x},0)$ specified, using the general solution formula for the wave equation.

8.6 MAPLE METHODS

We have constructed a number of Maple procedures that carry out the *Rayleigh-Ritz method* for the approximate determination of *eigenvalues and eigenfunctions*, as presented in Section 8.2, for one, two, and three dimensions. One set of procedures finds only approximate eigenvalues, while a corresponding set of procedures obtains approximate eigenvalues and eigenfunctions. Only the first set of procedures can handle problems that contain parameters. Additional procedures that determine approximate solutions of boundary value problems based on the Rayleigh-Ritz and the Galerkin methods (as presented in the exercises in Sections 8.1 and 8.2) have also been constructed but are not presented here.

Rayleigh-Ritz Method: One Dimension

In the one-dimensional case we consider regular or singular Sturm-Liouville eigenvalue problems for the second order ODE

$$-\frac{d}{dx}\left(p(x)\frac{du(x)}{dx}\right)+q(x)u(x)=\lambda\rho(x)u(x), \qquad a<x<b \tag{8.6.1}$$

with boundary conditions of the type presented in Section 4.3. The procedure $RayleighRitz1d$ finds the approximate eigenvalues for the Sturm-Liouville problem using the n admissible functions $\{\phi_k(x)\}$, $k=1,2,\ldots,n$ in the appropriate *Rayleigh quotient*. It is given as $RayleighRitz1d([p(x),q(x),\rho(x)],x=a..b,[\phi_1(x),\phi_2(x),\ldots,\phi_n(x)],[LBC,RBC])$. The boundary conditions at $x=a$ and $x=b$ are specified in the last argument. If $LBC=Dir$ or $LBC=Neu$, the boundary conditions are $u(a)=0$ or $u'(a)=0$, respectively. If $LBC=h$, the boundary condition is $u'(a)-hu(a)=0$. If the point $x=a$ is a singular point for the differential equation or we have $a=-\infty$, we put $LBC=Bound$. In that case, the admissible functions must be bounded at the left endpoint. The same values are given for RBC, and the corresponding boundary conditions are $u(b)=0$, $u'(b)=0$, and $u'(b)+hu(b)=0$, or the boundedness condition. Only in the case of a Dirichlet boundary condition at an endpoint must the $\{\phi_k(x)\}$ satisfy the admissibility condition that they vanish at that endpoint. If one of the $\phi_k(x)\neq 0$ at a Dirichlet endpoint, an error message is printed out and the function $\phi_k(x)$ must be replaced. We recall that no admissibility condition is required in the case of Neumann or Robin boundary conditions. However, as noted previously, if the $\{\phi_k(x)\}$ satisfy the relevant boundary condition, improved results are obtained.

As an example, we reconsider the eigenvalue problem of Example 8.4 but increase the number of admissible functions from 2 to 10. They are given as $\{\phi_k(x)=\pi^k-x^k\}$, $k=2,3,\ldots,11$. Each function vanishes at $x=\pi$ and has a vanishing derivative at $x=0$. The procedure is $RayleighRitz1d([1,0,1],x=0..\pi,[seq(\pi^k-x^k,k=2..11)],[Neu,Dir])$. The output list the first ten approximate eigenvalues $\hat{\lambda}_k$, $[0.25,2.26,6.26,12.3,20.2,30.4,43.6,63.4,111,222]$. The exact eigenvalues are $\lambda_k=(k+1/2)^2$, and the first ten eigenvalues are $[0.25,2.25,6.25,12.2,20.2,30.2,42.2,56.2,72.2,90.2]$. We find that the first seven approximate eigenvalues are

excellent approximations to the exact eigenvalues. In Example 8.4, where only two admissible functions were used in the Rayleigh-Ritz approximation, the second approximate eigenvalue was found not to provide too good an approximation to the exact eigenvalue.

Next we use the procedure to approximate the leading eigenvalues for Bessel's equation of order 1, as determined in Section 4.3, where we put $n = 1$ and $l = 1$ in (4.3.56). We invoke $RayleighRitz1d([x, 1/x, x], x = 0..1, [x(1-x), x(1-x)^2, x(1-x)^3, x(1-x)^4], [Bound, Dir])$. The $\phi_k(x)$ are chosen to vanish at $x = 1$ to satisfy the Dirichlet boundary condition. They are also chosen to vanish at $x = 0$ rather than simply to be bounded there, to guarantee that the integrals that occur in the Rayleigh quotient converge. The four leading approximate eigenvalues determined from the procedure are $[14.7, 49.2, 115., 226.]$. The exact eigenvalues are the squares of the zeros of the Bessel function of order 1. They are $[14.7, 49.3, 104., 177.]$. The first two eigenvalues are in good agreement.

The procedure $RayleighRitzEF1d$ finds approximate eigenvalues and eigenfunctions for the Sturm-Liouville problems considered above. The arguments of the procedure are identical with those of $RayleighRitz1d$. We apply it to the singular Sturm-Liouville problem of Section 4.3, which involves the Legendre polynomials. The procedure takes the form $RayleighRitzEF1d([1-x^2, 0, 1], x = -1..1, [1, x, x^2], [Bound, Bound])$. Each monomial is bounded at $x = -1$ and $x = 1$. The output of the procedure is; Approximate Eigenvalues $= [0., 2., 6.]$, Approximate Eigenfunctions $= [1, x, x^2 - 1/3]$. The approximate eigenvalues and eigenfunctions are, in fact, exact. The approximate eigenfunctions are the first three Legendre polynomials if $x^2 - 1/3$ is multiplied by $3/2$.

Rayleigh-Ritz Method: Two and Three Dimensions

We consider the eigenvalue problem

$$-\nabla \cdot (p \nabla M) + qM = \lambda \rho M \tag{8.6.2}$$

in the region G in two or three dimensions. In two dimensions we assume that G can be described as $c(x) < y < d(x)$, $a < x < b$ or $c(y) < x < d(y)$, $a < y < b$. The region G may be a rectangular region or the interior of a circle. In three dimensions G must be given as $r(x, y) < z < s(x, y)$, $c(x) < y < d(x)$, $a < x < b$ or in one of five other related sets of inequalities. The region G may be a rectangular region (say the interior of a cube) or the interior of a sphere.

RayleighRitz2d finds approximate eigenvalues for the two-dimensional eigenvalue problem using the n admissible functions $\{\phi_k(x, y)\}$, $k = 1, 2, \ldots, n$ in the appropriate *Rayleigh quotient*. It takes the form $RayleighRitz2d([p(x, y), q(x, y), \rho(x, y)], y = c(x)..d(x), x = a..b, [\phi_1(x, y), \ldots, \phi_n(x, y)], [BC_1, BC_2, BC_3, BC_4])$. [The alternativee form $c(y) < x < d(y)$, $a < y < b$, can also be used.] The boundary conditions are specified in the last argument. It is a list with four elements that specifies the coefficients in the Robin boundary conditions. The order in which the boundaries of the region are taken in the list, follows that of the curves given in the

range arguments. [For example, BC_1 corresponds to the boundary curve $y = c(x)$.] If one of the elements in the list is zero, there can be a Dirichlet or Neumann condition on that side. If it is a Dirichlet condition, all the $\phi_k(x, y)$ must vanish on that side. If they do not vanish, it corresponds to a Neumann condition. If the trial functions satisfy all the boundary conditions, improved results are results are obtained. As in the one dimensional case, the procedure $RayleighRitzEF2d$ determines approximate eigenvalues and eigenfunctions.

As a concrete example, we determine the leading eigenvalue for the problem

$$\nabla^2 M(x, y) + \lambda M(x, y) = 0, \qquad 0 < x < l,\ 0 < y < L, \tag{8.6.3}$$

with the mixed boundary conditions

$$M(0, y) = M_x(l, y) = M(x, 0) = M(x, L) = 0. \tag{8.6.4}$$

We make use of the exact eigenfunction $\phi_1(x, y) = \sin(\pi x/2l)\sin(\pi y/L)$ in $RayleighRitz2d([1, 0, 1], y = 0..L, x = 0..l, [\sin(\pi x/2l)\sin(\pi y/L)], [0, 0, 0, 0])$. The exact leading eigenvalue is $(\pi^2/4)l^2 + \pi^2/L^2$. If we replace the exact eigenfunction in the procedure by $\phi_1(x, y) = (y - L)xy$ we obtain approximate leading eigenvalue $3/l^2 + 10/L^2$. The corresponding procedure $RayleighRitzEF2d$ does not permit parameters in its arguments. If we put $l = 1$ and $L = 2$ in the procedure the approximate eigenvalue equals 5.5, as expected.

In three dimensions, $RayleighRitz3d$ and $RayleighRitzEF3d$ are the relevant approximation procedures. The region is described as above and there are six boundary conditions. We approximate the leading eigenvalue in the sphere as in Example 8.3. The procedure is $RayleighRitz3d([1, 0, 1], z = -\sqrt{1 - x^2 - y^2}..\sqrt{1 - x^2 - y^2}, y = -\sqrt{1 - x^2}..\sqrt{1 - x^2}, x = -1..1, [1 - x^2 - y^2 - z^2], [0, 0, 0, 0, 0, 0])$ The function $\phi_1(x, y, z) = 1 - x^2 - y^2 - z^2$ vanishes on the unit sphere as required. (We must use a function that results in integrals for the Rayleigh quotient that Maple can evaluate in cartesian coordinates. For this reason we do not use the function given in Example 8.3. While it is possible to transform the integrals to spherical coordinates, the procedure uses cartesian coordinates.) The output of the procedure is the approximate eigenvalue 10.5. As shown in Example 8.3 the exact eigenvalue is $\pi^2 \approx 9.87$. The approximation found in Example 8.3 is 10.

We conclude this discussion by approximating the leading eigenvalue for the Laplacian in the unit cube for the case of Robin boundary conditions on each side. That is, we consider

$$\nabla^2 M(x, y, z) + \lambda M(x, y, z) = 0, \quad 0 < x < 1,\ 0 < y < 1,\ 0 < z < 1, \tag{8.6.5}$$

with Robin boundary conditions at $x = 0$ and $x = 1$,

$$M_x(0, y, z) - M(0, y, z) = 0, \qquad M_x(1, y, z) + M(1, y, z) = 0, \tag{8.6.6}$$

and corresponding boundary conditions at $y = 0$, $y = 1$ and $z = 0$, $z = 1$. We use the procedure $RayleighRitz3d([1, 0, 1], z = 0..1, y = 0..1, x = 0..1, [(x^2 - x - 1)(y^2 - y - 1)(z^2 - z - 1)], [1, 1, 1, 1, 1, 1])$. The function $\phi_1(x, y, z) = (x^2 - x - 1)(y^2 - y - 1)(z^2 - z - 1)$ satisfies the boundary conditions on all six sides of the cube. Even though $\phi_1(x, y, z)$ need not satisfy an admissibility requirement for this problem, we expect an improved result with the foregoing choice. The output is the approximate eigenvalue 5.121951220. The exact leading eigenvalue, λ_1, is given in terms of a solution γ of the transcendental equation $\tan(\gamma) = 2\gamma/(\gamma^2 - 1)$ as $\lambda_1 = 3\gamma^2$, and has the approximate value 5.121158925.

Exercises 8.6

8.6.1. Apply the procedure $RayleighRitz1d$ to reproduce the first 10 approximate eigenvalues given at the beginning of this section.

8.6.2. Apply the procedure $RayleighRitz1d$ to obtain the first four approximate eigenvalues given at the beginning of this section. Determine the corresponding exact eigenvalues using the Maple procedure for determining zeros of Bessel functions.

8.6.3. Use the procedure $RayleighRitzEF1d$ to determine the approximate eigenvalues and eigenfunctions given at the end of the subsection on the Rayleigh-Ritz method in one dimension.

8.6.4. Apply the procedure $RayleighRitz2d$ to reproduce the result obtained following (8.6.3)–(8.6.4). Then use the procedure $RayleighRitzEF2d$ as in the discussion that follows the result to determine the approximate eigenvalue 5.5.

8.6.5. Use the procedure $RayleighRitz3d$ to reproduce the result obtained in the text after the first application of this procedure.

8.6.6. Use the procedure $RayleighRitz3d$ to reproduce the result obtained in the text after the second application of this procedure, and use Maple to solve the transcendental equation whose solution yields the exact eigenvalue.

CHAPTER 9

PERTURBATION METHODS

9.1 INTRODUCTION

Chapters 9 and 10 deal with a collection of methods that yield approximate solutions for a large class of initial and boundary value problems for partial differential equations. Generally speaking, these methods are used when a small parameter (or a large parameter) occurs in the given equation or data for the problem. Then the (assumed) solution is expanded in a series of powers (or inverse powers) of the parameter, and this expansion is inserted into the equation and data for the problem. By equating like powers of the parameter, a collection of problems results whose solution is expected to be simpler than that of the given problem.

If the series expansion of the solution converges, or is expected to converge, the aforementioned technique is often referred to as a *perturbation method.* If the series is divergent but asymptotic, so that the first few terms yield a good approximation when extreme values of the parameter are considered, the technique above is called an *asymptotic method.* The terminology is, however, not uniform in the literature and what are termed perturbation series may, in fact, be asymptotic, and vice versa.

In general, it is difficult to specify all the terms in a perturbation or asymptotic series for problems involving PDEs. Thus only the first few terms in the series are de-

Partial Differential Equations of Appplied Mathematics, Third Edition. By Erich Zauderer

termined and the distinction between a convergent or asymptotic series often becomes irrelevant. Although for many problems precise results are available regarding the convergent or asymptotic nature of the result, we concentrate on the formal aspects of constructing the series solutions and generally obtain only the first few terms of the series.

In this chapter we discuss *regular* and *singular perturbation methods*. In Chapter 10 we discuss three methods that are not generally referred to as perturbation methods. Therefore, we collect them all under the heading of *asymptotic methods*.

The appropriate expansion forms for the solution of a given problem are by no means obvious in all cases. Consequently, it is often useful to study exact solutions of particular problems and expand these solutions in a series involving the relevant small or large parameter. This not only suggests appropriate expansion forms for more general problems but may also yield useful information for specifying undetermined quantities that occur in the general case.

Perturbation techniques can also be used to replace given equations by simpler ones whose solutions contain many of the features of the solutions of the original problem. This is especially important for nonlinear equations where perturbation methods are used to linearize the problem, as has been done for Euler's equation of fluid dynamics in Section 8.5, for instance. Even if the linearization procedure breaks down in certain regions, it may still be possible to replace the given equation or system by a simpler nonlinear equation. The equation of simple wave motion that was obtained as an approximation to Euler's hydrodynamic equations is representative of this approach. This idea is developed further in this chapter as well as in Chapter 10.

In our study of hyperbolic equations we have seen that discontinuities or singularities in the solutions, interpreted in the weak sense, occur across characteristics. In fact, any rapid variation in the data for these equations must be carried along the characteristics. Near the characteristics, the solutions may be described by means of series expansions that may contain a small parameter, as we will show. These results are closely related to the asymptotic expansions that will be given for the Helmholtz or reduced wave equation, and they are both discussed in Chapter 10.

It will be seen that more than one type of expansion may be necessary to completely describe the perturbation or asymptotic solution of a given problem. For example, an expansion may break down in some region or may be insufficient to satisfy the data for the problem. These difficulties signify that the given expansion is not uniformly valid over the entire region of interest. Techniques such as the *boundary layer method* or the *method of multiple scales* need to be used to remedy this problem. They are introduced in this chapter but are used in both chapters.

Even though asymptotic equalities, asymptotic expansions, and order of magnitude symbols have often been defined and used in the preceding chapters, we now redefine these concepts in view of their particular relevance for the material in this and the following chapter.

The order symbol O is defined as

$$F(k) = O[G(k)], \qquad k \to a, \tag{9.1.1}$$

if $|F(k)/G(k)| \to A$ as $k \to a$, where A is a nonzero constant (here a may be $\pm\infty$). If (9.1.1) is valid, we say that $F(k)$ *is of order* $G(k)$ near $k = a$. The order symbol o is defined as

$$F(k) = o[G(k)], \qquad k \to a, \tag{9.1.2}$$

if $|F(k)/G(k)| \to 0$ as $k \to a$. For example, if $F(k) = k^2(10 - \cos(k)/k^3 + e^{-k})$ and $G(k) = k^2$, then $F(k) = O[G(k)]$ as $k \to \infty$ with $A = 10$. With $F(k)$ defined as before and $G(k) = k^3$, we have $F(k) = o[G(k)]$ as $k \to \infty$. Finally, the statement $F(k) = O[1]$ means that $F(k)$ is bounded as $k \to a$.

A function $f(\mathbf{x}, k)$ which depends on the parameter k has the *asymptotic expansion*

$$f(\mathbf{x}, k) \approx \sum_{n=0}^{\infty} f_n(\mathbf{x}) k^{-n}, \tag{9.1.3}$$

valid as $k \to \infty$, if for each N we have

$$f(\mathbf{x}, k) = \sum_{n=0}^{N-1} f_n(\mathbf{x}) k^{-n} + O[k^{-N}], \qquad k \to \infty. \tag{9.1.4}$$

We can also write (9.1.4) as

$$f(\mathbf{x}, k) = \sum_{n=0}^{N-1} f_n(\mathbf{x}) k^{-n} + o[k^{-N+1}], \qquad k \to \infty. \tag{9.1.5}$$

Both (9.1.4) and (9.1.5) are assumed to be uniformly valid for all $\mathbf{x}$ in some region. (Here $\mathbf{x}$ may be a real single variable, a real multivariable, or a complex variable.) Note that with $k = 1/\epsilon$, (9.1.3) represents an asymptotic power series expansion in ϵ, valid as $\epsilon \to 0$ through positive values. We shall emphasize the use of the (large) O order symbol in our discussion. The basic feature of an asymptotic expansion is that the remainder is of lower order (in the expansion variable) than the last term retained. The full expansion may, in fact, converge, but convergence is not necessary for the result to be useful. Often, more general forms of asymptotic expansions are required, such as series in fractional powers of k, and these are also defined as having remainders of lower order than that of the last term retained.

As stated previously, we do not prove that the formal series solutions we obtain are either convergent or asymptotic. Occasionally, therefore, strict equality rather than asymptotic equality signs are used even though the full series may diverge. Since only a few terms in the series are found, this distinction is not always significant. It is a general property of asymptotic expansions that finding additional terms need not improve the approximation since the series is generally divergent. As the value of the parameter k in (9.1.2) increases, the approximation provided by the series gets better. Thus we are effectively assuming in our discussion that the parameter is very large (or very small if $\epsilon = 1/k$) since we obtain only the leading terms. However, it is often the case that the results are very good even for moderate values of the parameter. Finally, we note that all formal operations that we carry out on the asymptotic expansion, such as term-by-term differentiation and integration, are assumed to be valid.

9.2 REGULAR PERTURBATION METHODS

We consider a linear or nonlinear differential equation

$$L(u, \epsilon) = 0 \tag{9.2.1}$$

that depends (smoothly) on the small positive parameter ϵ and a problem for (9.2.1) given over a bounded or unbounded spatial region G. If (9.2.1) is of elliptic type, appropriate boundary conditions are assigned on ∂G or at infinity. If (9.2.1) is of hyperbolic or parabolic type, in addition to the boundary conditions assigned on ∂G or at infinity for all $t > 0$, initial data are given in G at the time $t = 0$. The boundary or initial data may depend on ϵ, but the boundary ∂G is for the present assumed to be specified independently of ϵ. (More general problems that involve other types of data or boundaries that depend on ϵ will also be considered below.)

The *reduced* or *unperturbed problem* associated with the problem for (9.2.1) is obtained on formally setting $\epsilon = 0$ in (9.2.1) and its data. That is, we consider the *reduced equation*

$$L(v, 0) = 0 \tag{9.2.2}$$

with the *reduced data* obtained from the data for the given problem for (9.2.1). If the reduced problem has a unique solution, the given problem is called a *regular perturbation problem*. If this is not the case, we have a *singular perturbation problem*. Problems of the latter type are studied in Section 9.3.

Generally speaking, if the reduced equation is of different type or order than the given equation, we have a singular perturbation problem. It may happen, however, that the reduced problem can be solved even if the order or type of the given equation is changed.

For example, the *signaling problem* for the *hyperbolic equation*

$$\epsilon u_{tt}(x,t) - c^2 u_{xx}(x,t) + u_t(x,t) = 0, \qquad x > 0,\ -\infty < t < \infty, \tag{9.2.3}$$

with the boundary condition

$$u(0,t) = f(t), \qquad -\infty < t < \infty, \tag{9.2.4}$$

reduces to the *parabolic problem*

$$-c^2 v_{xx}(x,t) + v_t(x,t) = 0, \qquad x > 0,\ -\infty < t < \infty, \tag{9.2.5}$$

with the boundary condition

$$v(0,t) = f(t), \qquad -\infty < t < \infty. \tag{9.2.6}$$

Both the given and reduced problem can be solved in this case. However, the signaling problem should be interpreted as a large time limit of an initial and boundary value

problem with zero initial data, and in this case the initial value problem for (9.2.3) has an excess of initial conditions in relation to the reduced problem (9.2.5).

Similarly, the *hyperbolic equation*

$$\epsilon\left(u_{tt}(x,t) - c^2 u_{xx}(x,t)\right) + u_x(x,t) = 0, \qquad x > 0,\ -\infty < t < \infty, \quad (9.2.7)$$

with the boundary condition (9.2.4) reduces to

$$v_x(x,t) = 0, \qquad x > 0,\ -\infty < t < \infty, \quad (9.2.8)$$

with the condition (9.2.6). The (unique) solution of the reduced problem for $v(x,t)$ is $v(x,t) = f(t)$. Since (9.2.8) is a first order equation and (9.2.7) is of second order, the comments relating to the prescription of initial values apply to (9.2.7) and (9.2.8).

If the boundary conditions depend on the parameter ϵ, it may happen that the conditions for the reduced problem render it unsolvable. As an example, we consider the boundary value problem for the *biharmonic equation*

$$\nabla^2\nabla^2 u(x,y) = 0, \qquad (x,y) \in G, \quad (9.2.9)$$

with the boundary conditions

$$u(x,y)\Big|_{\partial G} = f(x,y), \qquad \epsilon\,\frac{\partial u(x,y)}{\partial n} + u(x,y)\Big|_{\partial G} = g(x,y). \quad (9.2.10)$$

Since (9.2.9) does not depend explicitly on ϵ, the reduced equation for $u(x,y)$ is also the biharmonic equation, but the reduced boundary conditions require that $u = f(x,y)$ and $u = g(x,y)$ on ∂G. Thus unless $f(x,y) = g(x,y)$, the reduced problem for $u(x,y)$ has no solution. Further, even if $f(x,y) = g(x,y)$, the solution of the reduced problem is not unique since a boundary condition is lost. Thus (9.2.9)–(9.2.10) is, in fact, a *singular perturbation problem.*

Continuing with our discussion of the *regular perturbation method*, we expand the solution u of (9.2.1) in the perturbation series

$$u = \sum_{n=0}^{\infty} u_n \epsilon^n. \quad (9.2.11)$$

The difference between u and u_0 (i.e., $u - u_0$) is referred to as a *perturbation* on the solution u_0 of the reduced or unperturbed problem. Inserting this expansion into (9.2.1) gives

$$L(u,\epsilon) = L\left(\sum_{n=0}^{\infty} u_n \epsilon^n, \epsilon\right) = 0. \quad (9.2.12)$$

We assume that $L(u,\epsilon)$ can be expanded in a power series in u and ϵ. As a result, (9.2.12) can be expressed in the form of a series

$$L(u,\epsilon) = \sum_{n=0}^{\infty} L_n(u_n, u_{n-1}, \ldots, u_1, u_0)\,\epsilon^n = 0, \quad (9.2.13)$$

where the L_n represent differential operators which may be linear or nonlinear, and which act on the functions $u_0, u_1, \ldots, u_n$. The series (9.2.11) is also inserted into the given initial and/or boundary conditions for the problem.

To solve the given problem by means of the perturbation method, we equate the coefficients of ϵ^n in (9.2.13) to zero and obtain

$$L_n(u_n, u_{n-1}, \ldots, u_1, u_0) = 0, \qquad n = 0, 1, \ldots. \tag{9.2.14}$$

Similarly, we equate coefficients of like powers of ϵ in the initial and/or boundary data. This yields the system of equations (9.2.14) with appropriate data that we solve recursively.

That is, we first solve the reduced equation $L_0(u_0) = 0$ with the relevant data. Once u_0 is specified, the equation for u_1, $L_1(u_1, u_0) = 0$, with its data is solved, and then the equations for $u_2, u_3, \ldots$ with their data are solved successively.

If the given equation (9.2.1) is linear, the equations $L_n(u_n, \ldots, u_0) = 0$ are, in general, nonhomogeneous versions of the equation $L_0(u_0) = 0$. However, $L_0(u_0) = 0$ may itself be a nonhomogeneous equation. Even if the given problem is nonlinear but the reduced problem is linear, all the equations (9.2.14) are homogeneous or nonhomogeneous equations of the same form.

For example, if we consider the nonlinear equation

$$L(u(x,t), \epsilon) = u_t(x,t) + u(x,t)u_x(x,t) - \epsilon u(x,t) = 0, \tag{9.2.15}$$

we obtain the reduced equation

$$L_0(u_0) = \frac{\partial u_0(x,t)}{\partial t} + u_0(x,t)\,\frac{\partial u_0(x,t)}{\partial x} = 0 \tag{9.2.16}$$

and

$$L_1(u_1, u_0) = \frac{\partial u_1(x,t)}{\partial t} + u_0(x,t)\frac{\partial u_1(x,t)}{\partial x} + u_1(x,t)\frac{\partial u_0(x,t)}{\partial x} - u_0(x,t) = 0. \tag{9.2.17}$$

However, the nonlinear equation

$$L(u(x,t), \epsilon) = u_t(x,t) + \epsilon u(x,t)u_x(x,t) - u_{xx}(x,t) = 0 \tag{9.2.18}$$

yields the reduced equation

$$L_0(u_0) = \frac{\partial u_0(x,t)}{\partial t} - \frac{\partial^2 u_0(x,t)}{\partial x^2} = 0 \tag{9.2.19}$$

and

$$L_1(u_1, u_0) = \frac{\partial u_1(x,t)}{\partial t} - \frac{\partial^2 u_1(x,t)}{\partial x^2} + u_0(x,t)\,\frac{\partial u_0(x,t)}{\partial x} = 0. \tag{9.2.20}$$

For (9.2.15) the reduced equation (9.2.16) is nonlinear and equation (9.2.17) for $u_1(x,t)$ is linear, whereas for (9.2.18) the reduced equation (9.2.19) as well as all the higher-order equations for the $u_n(x,t)$ are linear and of the same form.

The perturbation method can find approximate solutions to a large class of problems. However, it is often necessary to introduce modifications or special techniques to construct useful results or to extend the validity of the results to the full region of interest for the given problem. We now consider a number of different aspects of the perturbation method and exhibit each one of them by means of an example.

Perturbation Method in a Bounded Region

We begin by considering a boundary value problem in a bounded region for which the conventional perturbation method yields a result that is valid throughout the entire region. In the following example we discuss a problem for Helmholtz's equation in the unit disk for which a closed-form exact solution is easily obtained and compare the results of perturbation theory with the exact result.

Example 9.1. Helmholtz's Equation with a Small Parameter. We consider the two-dimensional *Helmholtz equation*

$$u_{xx}(x,y) + u_{yy}(x,y) + \epsilon^2 u(x,y) = 0 \tag{9.2.21}$$

in the unit disk $x^2 + y^2 < 1$ with the *Dirichlet boundary condition*

$$u(x,y) = 1, \qquad x^2 + y^2 = 1. \tag{9.2.22}$$

The parameter ϵ^2 is assumed to be small, such that the solution of (9.2.21)–(9.2.22) is unique.

To solve, we introduce a perturbation series in powers of ϵ^2,

$$u(x,y) = \sum_{n=0}^{\infty} u_n(x,y)\epsilon^{2n}. \tag{9.2.23}$$

We insert (9.2.23) into (9.2.21)–(9.2.22) to obtain

$$\nabla^2 u(x,y) + \epsilon^2 u(x,y) = \nabla^2 u_0(x,y) + \sum_{n=1}^{\infty} [\nabla^2 u_n(x,y) + u_{n-1}(x,y)]\epsilon^{2n} = 0, \tag{9.2.24}$$

$$u(x,y) = u_0(x,y) + \sum_{n=1}^{\infty} u_n(x,y)\epsilon^{2n} = 1, \qquad x^2 + y^2 = 1. \tag{9.2.25}$$

We have interchanged summation and differentiation in (9.2.24) and have collected like powers of ϵ^2. On equating like powers of ϵ^2 to zero in (9.2.24) and to unity or zero in (9.2.25), we obtain

$$\nabla^2 u_0(x,y) = 0, \quad \nabla^2 u_n(x,y) = -u_{n-1}(x,y), \; n \geq 1, \tag{9.2.26}$$

and the boundary conditions

$$u_0(x,y) = 1, \qquad u_n(x,y) = 0, \quad n \geq 1, \qquad x^2 + y^2 = 1. \tag{9.2.27}$$

The equations for the $u_n(x, y)$ can be solved recursively starting with that for $u_0(x, y)$ and using the boundary conditions (9.2.27). The perturbation method has replaced the *Helmholtz equation* (9.2.21) by the system of *Laplace* and *Poisson equations* (9.2.26). Introducing *polar coordinates* r and θ, the Helmholtz as well as the Laplace and Poisson equations can be solved by looking for solutions independent of θ since the problem has no angular dependence.

With $u = u(r)$, (9.2.21) takes the form

$$u''(r) + \frac{1}{r}u'(r) + \epsilon^2 u(r) = 0 \tag{9.2.28}$$

on expressing the Laplacian operator in polar coordinate form and dropping the θ-derivative. Now (9.2.28) is just *Bessel's equation of zero order* in the variable ϵr. The solution of (9.2.28) that is bounded at $r = 0$ and that satisfies the boundary condition $u(1) = 1$ is clearly

$$u(r) = \frac{J_0(\epsilon r)}{J_0(\epsilon)}, \tag{9.2.29}$$

where $J_0(z)$ is the Bessel function of order zero.

To solve for the u_n in the perturbation series we again assume that $u_n = u_n(r)$ and obtain for $u_0(r)$,

$$\frac{d^2u_0(r)}{dr^2} + \frac{1}{r}\frac{du_0(r)}{dr} = 0, \qquad u_0(1) = 1. \tag{9.2.30}$$

The bounded solution of (9.2.30) is $u_0(r) = 1$. The problem for $u_1(r)$ is

$$\frac{d^2u_1(r)}{dr^2} + \frac{1}{r}\frac{du_1(r)}{dr} = -u_0(r) = -1, \qquad u_1(1) = 0. \tag{9.2.31}$$

and a simple integration yields the bounded solution $u_1(r) = (1 - r^2)/4$.

Thus to leading orders the solution is

$$u(r) = 1 + \epsilon^2\frac{1 - r^2}{4} + O(\epsilon^4). \tag{9.2.32}$$

Since $J_0(z)$ has the expansion $J_0(z) = 1 - z^2/4 + O(z^4)$, we have

$$\frac{J_0(\epsilon r)}{J_0(\epsilon)} = \frac{1 - (\epsilon r)^2/4 + O(\epsilon^4)}{1 - \epsilon^2/4 + O(\epsilon^4)} = 1 + \epsilon^2\frac{1 - r^2}{4} + O(\epsilon^4). \tag{9.2.33}$$

The series obtained from (9.2.33) converges for sufficiently small ϵ, and it agrees to leading orders with the perturbation result (9.2.32). We observe that ϵ must be smaller than the first zero of $J_0(z)$, which occurs at $z \approx 2.4$, for the solution of (9.2.21)–(9.2.22) to exist. By noting that the order of magnitude of the first perturbation term $u_1(r) = \epsilon^2(1 - r^2)/4$ must be smaller than that of the leading term $u_0(r) = 1$ in (9.2.32) for $0 \leq r \leq 1$, we conclude that $\epsilon < 2$ for the perturbation result to be valid. Thus the perturbation solution yields a good approximation to the exact solution throughout the unit circle, even if we retain only the first few terms, as long as ϵ is small.

Perturbation Method in an Unbounded Region: Methods of Multiple Scales and Renormalization

Next we consider problems over an unbounded region for which the conventional perturbation method yields a result that is not uniformly valid over the entire region of interest. In the following example we examine the process of heat conduction with a slow radiation loss, which is governed by a parabolic equation with a small parameter. The perturbation solution of the relevant problem is not valid for all time, and modifications of the method are presented that extend the validity of the result. The *method of multiple scales* is a widely applicable approach to remedy the problem of nonuniformity. Another approach is the *method of renormalization*. Both methods are presented in the example.

Example 9.2. Heat Conduction with Slow Radiation. We consider the *Cauchy problem* for the *parabolic equation*

$$u_t(x,t) + \epsilon u(x,t) = u_{xx}(x,t), \qquad -\infty < x < \infty,\ t > 0, \tag{9.2.34}$$

with the initial value

$$u(x,0) = f(x), \qquad -\infty < x < \infty. \tag{9.2.35}$$

The change of variables

$$u(x,t) = e^{-\epsilon t} v(x,t) \tag{9.2.36}$$

yields

$$v_t(x,t) = v_{xx}(x,t), \qquad -\infty < x < \infty,\ t > 0, \qquad v(x,0) = f(x), \tag{9.2.37}$$

so that $v(x,t)$ satisfies the *heat equation*. The equation (9.2.34) describes heat conduction in a rod in which there is heat loss due to radiation on the surface. The radiative effect gives rise to the term $\epsilon u(x,t)$ in (9.2.34), and (9.2.36) shows that this yields a slow heat loss since $0 < \epsilon \ll 1$.

The solution of (9.2.34)–(9.2.35) is immediately given in terms of the solution of the Cauchy problem for the heat equation (9.2.37) studied previously. However, we want to apply the perturbation method to solve (9.2.34)–(9.2.35) in order to study the effect of the unboundedness of the (x,t) domain for the problem on the perturbation result. Let

$$u(x,t) = \sum_{n=0}^{\infty} u_n(x,t)\epsilon^n, \tag{9.2.38}$$

and insert (9.2.38) into (9.2.34)–(9.2.35). Equating like powers of ϵ gives

$$\frac{\partial u_0(x,t)}{\partial t} - \frac{\partial^2 u_0(x,t)}{\partial x^2} = 0, \quad \frac{\partial u_n(x,t)}{\partial t} - \frac{\partial^2 u_n(x,t)}{\partial x^2} = -u_{n-1}(x,t),\ n \geq 1. \tag{9.2.39}$$

The initial conditions are

$$u_0(x,0) = f(x), \quad u_n(x,0) = 0, \; n \geq 1. \tag{9.2.40}$$

Putting $u_0(x,t) = v(x,t)$, where $v(x,t)$ is the solution of (9.2.37) with the initial condition $v(x,0) = f(x)$, we easily conclude that

$$u_n(x,t) = \frac{(-t)^n}{n!}\, v(x,t), \qquad n \geq 0, \tag{9.2.41}$$

on using mathematical induction. We verify that $\partial u_n(x,t)/\partial t - \partial^2 u_n(x,t)/\partial x^2 = (-t)^n/n!\,[v_t(x,t) - v_{xx}(x,t)] - [(-t)^{n-1}/(n-1)!]\, v(x,t) = -u_{n-1}(x,t)$ since $v(x,t)$ satisfies (9.2.37). The full perturbation solution is

$$u(x,t) = \sum_{n=0}^{\infty} \left[\frac{(-t)^n}{n!}\, v(x,t)\right] \epsilon^n = \sum_{n=0}^{\infty} \frac{(-\epsilon t)^n}{n!}\, v(x,t) = e^{-\epsilon t} v(x,t), \tag{9.2.42}$$

and this is identical with the exact solution (9.2.36). However, the purpose of the perturbation approach is to approximate the exact solution by determining and retaining only the first few terms in the perturbation series. If we retain only the first two terms in the series (9.2.42) and write $u(x,t) = v(x,t) - \epsilon t v(x,t) + O(\epsilon^2)$, we conclude on comparing with the exact solution that for $\epsilon t \ll 1$, the above yields a good approximation.

But if $\epsilon t = O(1)$ or, equivalently, $t = O(1/\epsilon)$, we find that the first perturbation $\epsilon u_1(x,t) = -\epsilon t v(x,t)$ is of the same order of magnitude in ϵ as the leading term $u_0(x,t)$. In fact, (9.2.42) shows that every term $\epsilon^n u_n(x,t)$ in the series is of the same order in ϵ as the leading term $u_0(x,t)$ if $t = O(1/\epsilon)$. Consequently, no matter how small ϵ is, there is a time t at which all terms in the perturbation series are of the same order in ϵ and cannot be neglected on the basis that they constitute small corrections for small ϵ. Thus even though the leading terms of the perturbation series yield a good approximation for $\epsilon t \ll 1$, the result is not uniformly valid for all time. Terms of the form $\epsilon t v(x,t)$ are called *secular terms* and the difficulty caused by the occurrence of such terms in a perturbation series is referred to as *secular behavior.*

There are several methods for remedying the difficulties caused by secular behavior in a perturbation series. One approach involves the summation of all the secular terms in the series, thereby assigning equal importance to all of them. This process, called *renormalization*, can be carried out in one form or another for a number of problems. In our example each term in the perturbation series is a secular term, so that the summation of the series yields the exact solution (9.2.36).

Another approach, known as the *method of multiple scales*, takes note of the fact that the perturbation solution appears to exhibit two time scales for our problem. There is a *slow* time scale ϵt and a (comparatively) *rapid* time scale t, as shown by (9.2.42). Therefore, we look for a solution of the form $u(x,t) = \hat{u}(x,t,\tau)$, $\tau = \epsilon t$, that depends on x, t, and $\tau = \epsilon t$. That is, we treat the fast and slow time scales as independent variables.

Since

$$\frac{\partial u(x,t)}{\partial t} = \frac{\partial \hat{u}(x,t,\tau)}{\partial t} + \frac{\partial \hat{u}(x,t,\tau)}{\partial \tau}\frac{\partial \tau}{\partial t} = \frac{\partial \hat{u}(x,t,\tau)}{\partial t} + \epsilon \frac{\partial \hat{u}(x,t,\tau)}{\partial \tau}, \quad (9.2.43)$$

we have, instead of (9.2.34),

$$\hat{u}_t(x,t,\tau) + \epsilon[\hat{u}_\tau(x,t,\tau) + \hat{u}(x,t,\tau)] = \hat{u}_{xx}(x,t,\tau). \quad (9.2.44)$$

Expanding $\hat{u}(x,t,\tau)$ in a perturbation series

$$\hat{u}(x,t,\tau) = \sum_{n=0}^{\infty} \hat{u}_n(x,t,\tau)\epsilon^n \quad (9.2.45)$$

yields the recursive system

$$\frac{\partial \hat{u}_0(x,t,\tau)}{\partial t} - \frac{\partial^2 \hat{u}_0(x,t,\tau)}{\partial x^2} = 0, \quad (9.2.46)$$

$$\frac{\partial \hat{u}_n(x,t,\tau)}{\partial t} - \frac{\partial^2 \hat{u}_n(x,t,\tau)}{\partial x^2} = -\left[\frac{\partial \hat{u}_{n-1}(x,t,\tau)}{\partial \tau} + \hat{u}_{n-1}(x,t,\tau)\right], \; n \geq 1, \quad (9.2.47)$$

on equating like powers of ϵ.

We solve (9.2.46) using the initial condition $\hat{u}_0(x,t,\tau) = f(x)$ at $t = 0$. Thus $\hat{u}_0(x,t,\tau) = c(\tau)\, v(x,t)$, where $v(x,t)$ is defined as before and $c(\tau)$ is an arbitrary function of τ that need only satisfy $c(0) = 1$, since $t = 0$ implies that $\tau = \epsilon t = 0$. The equation (9.2.47) for $\hat{u}_1(x,t,\tau)$ becomes

$$\frac{\partial \hat{u}_1(x,t,\tau)}{\partial t} - \frac{\partial^2 \hat{u}_1(x,t,\tau)}{\partial x^2} = -[c'(\tau) + c(\tau)]\, v(x,t), \quad (9.2.48)$$

with the initial condition $\hat{u}_1(x,t,\tau) = 0$ at $t = 0$. The terms $c'(\tau)$ and $c(\tau)$ are constants as far the operator on the left side of (9.2.48) is concerned, so we obtain the solution

$$\hat{u}_1(x,t,\tau) = -t\,[c'(\tau) + c(\tau)]v(x,t) + d(\tau)\hat{v}(x,t), \quad (9.2.49)$$

where $d(\tau)$ vanishes at $\tau = 0$ and $\hat{v}(x,t)$ is a solution of the heat equation (9.2.37), but they are both otherwise arbitrary. Consequently, we still have a secular term that grows with t in the solution (9.2.49). However, we are now in a position to remove the secularity by specifying $c(\tau)$ such that $c'(\tau) + c(\tau) = 0$, $c(0) = 1$, which yields $c(\tau) = e^{-\tau} = e^{-\epsilon t}$. We also set $d(\tau) = 0$, for otherwise a further secular term would arise at the next level of approximation. Consequently, $\hat{u}_1(x,t,\tau) = 0$ and, as a result, $\hat{u}_n(x,t,\tau) = 0$ for all $n > 1$. The series (9.2.45) thus terminates with the first term and yields the exact solution (9.2.36).

In the preceding example an exactly solvable problem was solved by the perturbation method. Next we consider a nonlinear problem that cannot be solved exactly. The perturbation method reduces it to a class of solvable linear problems. The problem exhibits *secular behavior*, and the difficulty is (partially) resolved by using a *renormalization method* rather than the method of multiple scales.

Example 9.3. A Nonlinear Klein-Gordon Equation. The nonlinear hyperbolic equation

$$w_{tt}(x,t) - \gamma^2 w_{xx}(x,t) + c^2 w(x,t) - \sigma w^3(x,t) = 0 \tag{9.2.50}$$

reduces to the *Klein-Gordon equation* when the parameter σ is equated to zero and arises in a number of physical contexts. The coefficients in (9.2.50) are assumed to be constants. We consider the initial value problem for (9.2.50) over the infinite line $-\infty < x < \infty$ with the data

$$w(x,0) = \epsilon \cos(kx), \qquad w_t(x,0) = 0, \tag{9.2.51}$$

where $0 < \epsilon \ll 1$ and k is a prescribed constant.

Since the initial data (9.2.51) are uniformly small in magnitude, we look for a solution of (9.2.50)–(9.2.51) in the form

$$w(x,t) = \epsilon\, u(x,t), \tag{9.2.52}$$

and apply the perturbation method to the problem for $u(x,t)$. We must solve

$$u_{tt}(x,t) - \gamma^2 u_{xx}(x,t) + c^2 u(x,t) - \epsilon^2 \sigma u^3(x,t) = 0, \tag{9.2.53}$$

with the initial data

$$u(x,0) = \cos(kx), \qquad u_t(x,0) = 0. \tag{9.2.54}$$

Expanding $u(x,t)$ as

$$u(x,t) = \sum_{n=0}^{\infty} u_n(x,t)\, \epsilon^n, \tag{9.2.55}$$

and inserting (9.2.55) into (9.2.53)–(9.2.54) yields the equations for the leading terms $u_0(x,t), \ldots, u_2(x,t)$ as

$$L_n(u_n) = \frac{\partial^2 u_n}{\partial t^2} - \gamma^2 \frac{\partial^2 u_n}{\partial x^2} + c^2 u_n = 0, \qquad n = 0, 1, \tag{9.2.56}$$

$$L_2(u_2,\, u_0) = \frac{\partial^2 u_2}{\partial t^2} - \gamma^2 \frac{\partial^2 u_2}{\partial x^2} + c^2 u_2 - \sigma u_0^3 = 0. \tag{9.2.57}$$

The initial conditions are

$$u_0(x,0) = \cos(kx),\ u_n(x,0) = 0,\ n \geq 1, \quad \frac{\partial u_n(x,0)}{\partial t} = 0,\ n \geq 0. \quad (9.2.58)$$

Using separation of variables, the solution of the initial value problem for $u_0(x,t)$ is found to be $u_0(x,t) = \cos(\omega t)\cos(kx)$, where $\omega^2 = \gamma^2 k^2 + c^2$, which is the *dispersion relation* for the *linear Klein-Gordon equation* (9.2.56). The equation for $u_1(x,t)$ is uncoupled from that for $u_0(x,t)$ and has homogeneous data. Thus $u_1(x,t) = 0$.

Using the trigonometric identity, $\cos^3(kx) = (3/4)\cos(kx) + (1/4)\cos(3kx)$, the equation for $u_2(x,t)$ may be written as $\partial^2 u_2(x,t)/\partial t^2 - \gamma^2\,\partial^2 u_2(x,t)/\partial x^2 + c^2 u_2(x,t) = (3\sigma/4)\cos^3(\omega t)\cos(kx) + (\sigma/4)\cos^3(\omega t)\cos(3kx)$. Putting $u_2(x,t) = F_1(t)\cos(kx) + F_2(t)\cos(3kx)$, and inserting it into the equation for $u_2(x,t)$, we obtain the ordinary differential equations

$$F_1''(t)+(\gamma^2k^2+c^2)F_1(t) = \frac{3\sigma}{4}\cos^3(\omega t),\ F_2''(t)+(9\gamma^2k^2+c^2)F_2(t) = \frac{\sigma}{4}\cos^3(\omega t). \quad (9.2.59)$$

The initial conditions for $u_2(x,t)$ imply that $F_1(t)$, $F_1'(t)$, $F_2(t)$, and $F_2'(t)$ all vanish at $t = 0$. On using the trigonometric identity above, the equations for $F_1(t)$ and $F_2(t)$ are easily solved by the method of undetermined coefficients.

With these results, the leading terms of the perturbation series solution $u = u(x,t)$ of (9.2.53)–(9.2.54) are found to be

$$u = \cos(\omega t)\cos(kx) + \epsilon^2\left[\frac{9\sigma}{32\omega}t\sin(\omega t) + \frac{3\sigma}{128\omega^2}(\cos(\omega t) - \cos(3\omega t))\right]\cos(kx)$$

$$+\epsilon^2\left[\frac{3\sigma}{128\gamma^2k^2}(\cos(\omega t) - \cos(\lambda t)) + \frac{\sigma}{128c^2}(\cos(\lambda t) - \cos(3\omega t))\right]\cos(3kx) \quad (9.2.60)$$

to $O(\epsilon^3)$, where $\lambda^2 = 9\gamma^2k^2+c^2$. We detect the presence of a *secular term* $f_2(x,t) = (9\epsilon^2\sigma/32\omega)t\sin(\omega t)\cos(kx)$ in the series. When $t = O(\epsilon^{-2})$ this term is of the same order of magnitude as the leading term. As a result, the perturbation series loses its validity as t reaches those values. For, on constructing a perturbation series we assume that terms containing higher powers of ϵ are much smaller (in magnitude) than those with lower powers. Here the range of t is unbounded, and we indicated in Example 9.2 that the conventional perturbation expansion is generally not uniformly valid over unbounded regions.

Consequently, the result (9.2.60) is valid and useful only for times t such that $t < O(\epsilon^{-2})$. We present a simple method for extending the validity of the perturbation result to larger values of t. We note that

$$\cos(\omega t) + \frac{9\epsilon^2\sigma}{32\omega}\,t\sin(\omega t) = \cos\left[\left(\omega - \frac{9\epsilon^2\sigma}{32\omega}\right)t\right] + O(\epsilon^4), \quad (9.2.61)$$

as follows on expanding the right side of (9.2.61) for small ϵ. Using (9.2.61) to replace the first two terms on the right in (9.2.60) yields the improved result

$$u = \cos\left[\left(\omega - \frac{9\epsilon^2\sigma}{32\omega}\right)t\right]\cos(kx) + \epsilon^2\left[\frac{3\sigma}{128\omega^2}(\cos(\omega t) - \cos(3\omega t))\right]\cos(kx)$$

$$+ \epsilon^2\left[\frac{3\sigma}{128\gamma^2k^2}(\cos(\omega t) - \cos(\lambda t)) + \frac{\sigma}{128c^2}(\cos(\lambda t) - \cos(3\omega t))\right]\cos(3kx), \tag{9.2.62}$$

correct to $O(\epsilon^3)$, in which the secular term is no longer present. There may be other secular terms in the series, but they occur for terms of higher order in ϵ, and the validity of the perturbation solution is presumably extended in time when written in the form (9.2.62). We return to this problem when we consider periodic finite-amplitude traveling wave solutions of the nonlinear Klein-Gordon and related equations below.

The method used to remove the *secular term* is essentially the *method of renormalization* discussed in Example 9.2. Secular terms in the perturbation series are removed by summing a part of the perturbation series. In Example 9.2 a full perturbation series was summed, whereas in the present example, since only the leading terms of the perturbation series were determined, the renormalization method was carried out in an approximate manner.

Hyperbolic Equation with Slowly Varying Coefficients

Partial differential equations with slowly varying coefficients can also be treated by using a perturbation approach. In the following example we consider a linear hyperbolic equation with nonconstant but slowly varying coefficients. As we see, the perturbation method, appropriately modified, yields a simple approximate solution.

Example 9.4. The Wave Equation with a Slowly Varying Wave Speed. The *hyperbolic equation*

$$u_{tt}(x,t) - c^2(\epsilon x)u_{xx}(x,t) = 0, \tag{9.2.63}$$

where $c^2(\epsilon x)$ is assumed to have the expansion

$$c^2(\epsilon x) = c_0^2 + \sum_{n=1}^{\infty} c_n[\epsilon x]^n \tag{9.2.64}$$

with constant c_n, reduces to the *wave equation* with wave speed c_0 if $\epsilon = 0$. Since $dc^2(\epsilon x)/dx = O(\epsilon)$ and ϵ is taken to be small, we state that the function $c(\epsilon x)$ gives rise to a slow variation of the wave speed for (9.2.63) around the constant speed c_0. We note that (9.2.63) is not expressed in self-adjoint form and that the (variable) speed of disturbances as determined by the domain of influence for concentrated initial data for (9.2.63) (see Section 8.3) is given by $c(\epsilon x)$. We apply the perturbation method to

an initial value problem for (9.2.63) to determine the effect of the slow variation of $c^2(\epsilon x)$ around c_0^2.

Let the initial conditions for (9.2.63) be

$$u(x,0) = f(x), \qquad u_t(x,0) = 0, \qquad -\infty < x < \infty. \tag{9.2.65}$$

We expand $u(x,t)$ as $u(x,t) = \sum_{n=0}^{\infty} u_n(x,t)\epsilon^n$, and insert it into (9.2.63) and (9.2.65). Equating like powers of ϵ yields as the leading order equations

$$\frac{\partial^2 u_0(x,t)}{\partial t^2} - c_0^2 \frac{\partial^2 u_0(x,t)}{\partial x^2} = 0, \tag{9.2.66}$$

$$\frac{\partial^2 u_1(x,t)}{\partial t^2} - c_0^2 \frac{\partial^2 u_1(x,t)}{\partial x^2} = x c_1 \frac{\partial^2 u_0(x,t)}{\partial x^2}, \tag{9.2.67}$$

where (9.2.64) has been used. The initial conditions for $u_0(x,t)$ and $u_1(x,t)$ are

$$u_0(x,0) = f(x), \ \frac{\partial u_0(x,0)}{\partial t} = 0, \ u_1(x,0) = \frac{\partial u_1(x,0)}{\partial t} = 0. \tag{9.2.68}$$

We first discuss this problem in general terms, noting the difficulties that arise and indicating how they can be resolved. Then we consider a specific choice for $c(\epsilon x)$ and obtain an explicit result for that case.

The solution of (9.2.66) and (9.2.68) is

$$u_0(x,t) = \frac{1}{2} f(x - c_0 t) + \frac{1}{2} f(x + c_0 t). \tag{9.2.69}$$

We insert this result in the equation (9.2.67) for $u_1(x,t)$ and obtain, using the method of undetermined coefficients and the data (9.2.68),

$$u_1(x,t) = \frac{c_1 t}{8c_0}[f(x-c_0t) - f(x+c_0t)]$$

$$-\frac{c_1 x t}{4c_0}[f'(x - c_0 t) - f'(x + c_0 t)] + \frac{c_1 t^2}{8}[f'(x - c_0 t) + f'(x + c_0 t)]. \tag{9.2.70}$$

We observe the presence of *secular terms* in (9.2.70). When $t = O(\epsilon^{-1/2})$ or $xt = O(\epsilon^{-1})$, the term $\epsilon u_1(x,t)$ is of the same order in ϵ as the leading term $u_0(x,t)$, assuming that $f(x)$ and its derivatives are uniformly bounded. Since both the x and the t intervals relevant to this problem are unbounded, the perturbation series would appear to be invalid, in general, since values of x for which $x = O[(\epsilon t)^{-1}]$ must occur if $-\infty < x < \infty$, for any $t > 0$. However, if $f(x)$ vanishes outside a bounded interval, so does $u_1(x,t)$, and large values of x do not play a role as long as t is not too large. In that case, the perturbation theory solution may be valid for a substantial length of time.

The general difficulty with the foregoing perturbation series may be traced to the fact that the solutions (9.2.69) and (9.2.70) are expressed in terms of the characteristics

$$\hat{\phi}_{\pm}(xt) = x \pm c_0 t = \text{constant} \tag{9.2.71}$$

of the *reduced equation* [i.e., (9.2.63) with $\epsilon = 0$] rather than the exact characteristics of (9.2.63), which are given as solutions of

$$\frac{dt}{dx} = \pm \frac{1}{c(\epsilon x)}. \tag{9.2.72}$$

Given a point $(x_0, 0)$ on the x-axis, the difference between the *exact* and *reduced characteristics* issuing from that point becomes appreciable for large values of x and t. Consequently, we may expect that the perturbation solution becomes invalid for x and t large.

To resolve the difficulties with the secular terms that arise in the perturbation result (9.2.69) and (9.2.70), we introduce the independent families of exact characteristic curves $\phi(x, t)$ = constant and $\psi(x, t)$ = constant, determined from (9.2.72), which reduce to the curves $x - c_0 t$ = constant and $x + c_0 t$ = constant, respectively, when we set $\epsilon = 0$ in (9.2.72). Noting the results (9.2.69) and (9.2.70), we look for a solution of (9.2.63) in the form

$$u(x, t) = g(\epsilon x, \epsilon t) f(\phi(x, t)) + h(\epsilon x, \epsilon t) f(\psi(x, t)). \tag{9.2.73}$$

We insert (9.2.73) into (9.2.63) and note that since $\phi(x, t)$ = constant and $\psi(x, t)$ = constant are characteristics, they satisfy the characteristic equations for (9.2.63), and we have

$$\phi_t^2 - c^2 \phi_x^2 = \psi_t^2 - c^2 \psi_x^2 = 0. \tag{9.2.74}$$

We readily conclude that

$$g_\tau + c g_\sigma - \frac{c'}{2} g = 0, \quad h_\tau - c h_\sigma + \frac{c'}{2} h = 0, \tag{9.2.75}$$

where $\sigma = \epsilon x$ and $\tau = \epsilon t$, and that the initial conditions (9.2.65) yield

$$g(\sigma, 0) = h(\sigma, 0) = \frac{1}{2}. \tag{9.2.76}$$

The initial values for the characteristics are found to be

$$\phi(x, 0) = \psi(x, 0) = x. \tag{9.2.77}$$

Using the methods of Chapter 2, the initial value problems for ϕ, ψ, g, and h are readily solved. Rather than present the general solution, we restrict our attention to a specific problem for which the characteristics can be determined explicitly.

Consider (9.2.63) in the interval $0 < x < \infty$ with

$$c(\epsilon x) = c_0 + \epsilon x, \tag{9.2.78}$$

$u(0, t) = 0$ for $t > 0$ and the data (9.2.65) for $x > 0$. [We are now considering an initial and boundary value problem for (9.2.63).] On comparing (9.2.64) with (9.2.78),

we conclude that $c_1 = 2c_0$, $c_2 = 1$, and $c_n = 0$ for $n > 2$. The characteristics $\phi(x,t)$ and $\psi(x,t)$ that correspond to this choice of $c(\epsilon x)$ and the data (9.2.77) are

$$\phi(x,t) = xe^{-\epsilon t} + \frac{c_0}{\epsilon}(e^{-\epsilon t} - 1), \quad \psi(x,t) = xe^{\epsilon t} + \frac{c_0}{\epsilon}(e^{\epsilon t} - 1). \tag{9.2.79}$$

We discuss the solution in the region where $\phi(x,t) > 0$. The equations (9.2.75) for g and h with the data (9.2.76) can be solved by looking for solutions that are independent of σ since $c' = 1$ in this case. We obtain

$$g(\epsilon x, \epsilon t) = \frac{1}{2}\exp\left(\frac{\epsilon t}{2}\right), \qquad h(\epsilon x, \epsilon t) = \frac{1}{2}\exp\left(-\frac{\epsilon t}{2}\right). \tag{9.2.80}$$

Then the approximate solution (9.2.73) is

$$u(x,t) \approx \frac{1}{2}e^{(\epsilon t/2)} f\left[xe^{-\epsilon t} + \frac{c_0}{\epsilon}(e^{-\epsilon t} - 1)\right] + \frac{1}{2}e^{(-\epsilon t/2)} f\left[xe^{\epsilon t} + \frac{c_0}{\epsilon}(e^{\epsilon t} - 1)\right]. \tag{9.2.81}$$

On expanding (9.2.81) in powers of ϵ, it is easily shown that the first two terms in the expansion agree with (9.2.69) and (9.2.70) in the region $\phi(x,t) > 0$. Although the secular terms found in the conventional perturbation expansion are absent in (9.2.81), we do observe a slow exponential growth in t, say, if $f(x)$ is uniformly bounded. Although it would appear that we have used the method of multiple scales to obtain (9.2.81), this problem is somewhat more complicated since we not only introduced ϵx and ϵt as variables but also the characteristics $\phi(x,t)$ and $\psi(x,t)$.

Boundary Perturbation Methods

We have seen that the separation of variables and transform methods for the solution of initial and/or boundary value problems for partial differential equations are useful only for problems given over special regions, such as squares, circles, spheres, or half-spaces. If the boundary of the region varies only slightly from a boundary for which the foregoing methods can be applied successfully, it is possible to use perturbation theory to effect an approximate solution to the problem. We now present a brief discussion of *boundary perturbation methods* for a two-dimensional problem and consider an example.

Let the given problem be specified in a two-dimensional region $\hat{G}$ with $\partial\hat{G}$ as its boundary. We assume that $\partial\hat{G}$ can be expressed in parametric form as

$$x = g(s) + \epsilon\,\hat{g}(s), \qquad y = h(s) + \epsilon\,\hat{h}(s), \tag{9.2.82}$$

where $0 < \epsilon \ll 1$, s is a parameter and the functions $g(s)$, $\hat{g}(s)$, $h(s)$, and $\hat{h}(s)$ are prescribed. We could also consider a more general ϵ dependence such as $x = g(s,\epsilon)$ and $y = h(s,\epsilon)$, where g and h can be expanded in powers of ϵ, but we restrict our discussion to the form (9.2.82) for the sake of simplicity. Assuming that $u(x,y)$ is

specified on the boundary curve $\partial\hat{G}$, we have

$$u(x,y)|_{\partial G} = u(g(s) + \epsilon\hat{g}(s), h(s) + \epsilon\hat{h}(s)) = f(s), \tag{9.2.83}$$

where $f(s)$ is prescribed. Further, $u(x,y)$ is assumed to satisfy a linear second order partial differential equation with no explicit ϵ dependence

$$Lu(x,y) = 0, \qquad (x,y) \in \hat{G}. \tag{9.2.84}$$

Using the perturbation method, we expand $u(x,y)$ as

$$u(x,y) = \sum_{n=0}^{\infty} u_n(x,y)\epsilon^n, \tag{9.2.85}$$

and insert (9.2.85) into (9.2.83)–(9.2.84). Since L is a linear operator we find that $Lu_n(x,y) = 0,\ n \geq 0$. In (9.2.83) we first expand u in powers of ϵ as

$$u(g + \epsilon\hat{g}, h + \epsilon\hat{h}) = u(g,h) + \epsilon\left[\hat{g}u_x(g,h) + \hat{h}u_y(g,h)\right] + O(\epsilon^2), \tag{9.2.86}$$

and then obtain from (9.2.85)

$$u_0(g,h) = f, \quad u_1(g,h) = -\hat{g}\,\frac{\partial u_0(g,h)}{\partial x} - \hat{h}\,\frac{\partial u_0(g,h)}{\partial y}. \tag{9.2.87}$$

We find that the perturbation method reduces the given problem to a collection of problems given over the unperturbed region G with the boundary curve ∂G whose parametric equations are $x = g(s)$ and $y = h(s)$. It is generally assumed that the boundary value problems for the functions $u_n(x,y)$ in the region G can be solved exactly. It is a straightforward matter to extend these results to three-dimensional problems and to cases with other types of boundary conditions. We now consider an example.

Example 9.5. Boundary Perturbations for Laplace's Equation. The solution of the *Dirichlet problem* for *Laplace's equation* in a rectangle was obtained in Example 4.4 via the method of separation of variables. We now consider a small perturbation of the rectangular region $0 < x < l$ and $0 < y < L$. Let the region G be given as the interior of the trapezoid defined as $0 < x < l$ and $\epsilon x < y < L$, so that the boundary line $y = 0$ is replaced by $y = \epsilon x$ with $0 < \epsilon \ll 1$ (it is assumed that $\epsilon l < L$).

The Dirichlet problem is given as

$$u_{xx}(x,y) + u_{yy}(x,y) = 0, \qquad 0 < x < l, \qquad \epsilon x < y < L, \tag{9.2.88}$$

with the boundary conditions

$$\begin{cases} u(0,y) = 0,\ 0 < y < L, & u(l,y) = 0,\ \epsilon\, l < y < L, \\ u(x,L) = 0,\ 0 < x < l, & u(x,\epsilon x) = f(x),\ 0 < x < l, \end{cases} \tag{9.2.89}$$

where $f(x)$ is a prescribed function.

To solve (9.2.88)–(9.2.89) we first expand $u(x, \epsilon x)$ as

$$u(x, \epsilon x) = u(x, 0) + \epsilon x u_y(x, 0) + O(\epsilon^2) = f(x). \tag{9.2.90}$$

Then we expand $u(x, y)$ as in (9.2.85) and find that each of the $u_n(x, y)$ satisfies Laplace's equation. Inserting this expansion into (9.2.90) yields

$$u_0(x, 0) + \epsilon \left[u_1(x, 0) + x \frac{\partial u_0(x, 0)}{\partial y} \right] + O(\epsilon^2) = f(x). \tag{9.2.91}$$

For $u_0(x, y)$ we have the boundary conditions $u_0(x, y) = 0,\ x = 0,\ y = L,\ x = l,\ u_0(x, 0) = f(x),\ 0 < x < l$, in view of (9.2.91), with $\nabla^2 u_0(x, y) = 0,\ 0 < x < l,\ 0 < y < L$. The function $u_1(x, y)$ also vanishes on the three sides of the rectangle (i.e., $x = 0,\ y = L$, and $x = l$), while on the fourth side

$$u_1(x, 0) = -x \frac{\partial u_0(x, 0)}{\partial y}, \qquad 0 < x < l, \tag{9.2.92}$$

and it satisfies Laplace's equation $\nabla^2 u_1(x, y) = 0,\ 0 < x < l,\ 0 < y < L$.

From Example 4.4 we obtain the solution of the problem for $u_0(x, y)$ as

$$u_0(x, y) = \sqrt{\frac{2}{l}} \sum_{k=1}^{\infty} b_k \sinh \left[\frac{\pi k}{l} (y - L) \right] \sin \left(\frac{\pi k x}{l} \right), \tag{9.2.93}$$

where b_k is defined in (4.4.38) with $\hat{l}$ replaced by L. To determine the boundary condition (9.2.92) for $u_1(x, y)$, we differentiate (9.2.93) term by term with respect to y and evaluate it at $y = 0$ to obtain

$$u_1(x, 0) = -\sqrt{\frac{2}{l}} \sum_{k=1}^{\infty} x b_k \frac{\pi k}{l} \cosh \left(\frac{\pi k L}{l} \right) \sin \left(\frac{\pi k x}{l} \right). \tag{9.2.94}$$

Again the results of Example 4.4 can be used to specify $u_1(x, y)$ by putting $f(x)$ equal to the right side of (9.2.94) and setting $g(x) = 0$ in that example.

For the sake of concreteness, we set $f(x) = \sin(\pi x/l)$ in (9.2.89). Then $u_0(x, y)$ has the form

$$u_0(x, y) = \frac{\sinh[\pi(L - y)/l]}{\sinh(\pi L/l)} \sin(\pi x/l), \tag{9.2.95}$$

and the condition (9.2.94) for $u_1(x, y)$ is $u_1(x, 0) = (\pi x/l) \coth(\pi L/l) \sin(\pi x/l)$. We note that $u_1(x, 0)$ vanishes at $x = 0$ and $x = l$ and is a smooth function in the interval $0 < x < l$. Therefore, its Fourier sine series converges uniformly in the given interval and as a result, the solution $u_1(x, y)$ obtained by the methods of Example 4.4 is continuous and therefore bounded in the given region. Thus $\epsilon u_1(x, y)$ represents a small perturbation around $u_0(x, y)$ for small ϵ. Similarly, if $u_1(x, 0)$ in (9.2.94) is a smooth function in $0 < x < l$, we conclude that $\epsilon u_1(x, y)$ is a small perturbation around $u_0(x, y)$ since $u_1(0, 0) = u_1(l, 0) = 0$.

Perturbation Method for Eigenvalue Problems

Eigenvalue problems can also be treated by perturbation methods. A special case has already been considered in Example 8.5. We use the notation of Chapters 4 and 8 in our discussion and consider the problem

$$LM(\mathbf{x};\epsilon) \equiv -\nabla \cdot (p(\mathbf{x})\nabla M(\mathbf{x};\epsilon)) + q(\mathbf{x})M(\mathbf{x};\epsilon) = (\lambda(\epsilon) + \epsilon\, r(\mathbf{x}))\rho(\mathbf{x})M(\mathbf{x};\epsilon) \tag{9.2.96}$$

in the bounded region G with the boundary condition

$$\alpha(\mathbf{x})M(\mathbf{x};\epsilon) + \beta(\mathbf{x})\left.\frac{\partial M(\mathbf{x};\epsilon)}{\partial n}\right|_{\partial G} = 0. \tag{9.2.97}$$

The parameter ϵ is small (i.e., $0 < \epsilon \ll 1$) and $r(\mathbf{x})$ is a given function. It is assumed that the reduced eigenvalue problem with $\epsilon = 0$ is solvable and that the eigenvalues are simple, so that there is one linearly independent eigenfunction for each eigenvalue.

To determine the eigenvalues and eigenfunctions of (9.2.96)–(9.2.97), we expand both $M(\mathbf{x};\epsilon)$ and $\lambda(\epsilon)$ in powers of ϵ and set

$$M(\mathbf{x};\epsilon) = \sum_{n=0}^{\infty} M^{(n)}(\mathbf{x})\,\epsilon^n, \quad \lambda(\epsilon) = \sum_{n=0}^{\infty} \lambda^{(n)}\epsilon^n. \tag{9.2.98}$$

Inserting (9.2.98) into (9.2.96)–(9.2.97) gives

$$LM^{(0)}(\mathbf{x}) = \lambda^{(0)}\rho(\mathbf{x})M^{(0)}(\mathbf{x}), \tag{9.2.99}$$

$$LM^{(1)}(\mathbf{x}) = \rho(\mathbf{x})\left(\lambda^{(0)}M^{(1)}(\mathbf{x}) + r(\mathbf{x})M^{(0)}(\mathbf{x}) + \lambda^{(1)}M^{(0)}(\mathbf{x})\right) \tag{9.2.100}$$

as the equations for the first two $M^{(n)}(\mathbf{x})$. The boundary conditions are

$$\alpha(\mathbf{x})M^{(n)}(\mathbf{x}) + \beta(\mathbf{x})\left.\frac{\partial M^{(n)}(\mathbf{x})}{\partial n}\right|_{\partial G} = 0, \qquad n \geq 0. \tag{9.2.101}$$

Let $\lambda_k^{(0)} \equiv \lambda_k$ and $M^{(0)}(\mathbf{x}) \equiv M_k(\mathbf{x})$ $(k = 1, 2, \ldots)$ represent the (simple) eigenvalues and orthonormalized eigenfunctions of the unperturbed problem (9.2.99) and (9.2.101), where $n = 0$ (these are assumed to be known). The equation (9.2.100) is an inhomogeneous version of (9.2.99), with $M^{(1)}(\mathbf{x})$ required to satisfy the homogeneous boundary condition (9.2.101). This problem has no solution unless the inhomogeneous term satisfies a compatibility condition, as we now show. We replace $M^{(n)}(\mathbf{x})$ and $\lambda^{(n)}$ by $M_k^{(n)}(\mathbf{x})$ and $\lambda_k^{(n)}$ with $n = 0, 1$ in (9.2.99)–(9.2.101). Then (9.2.100) takes the form

$$LM_k^{(1)}(\mathbf{x}) - \lambda_k\rho(\mathbf{x})M_k^{(1)}(\mathbf{x}) = \left(r(\mathbf{x}) + \lambda_k^{(1)}\right)\rho(\mathbf{x})M_k(\mathbf{x}), \tag{9.2.102}$$

since $\lambda_k^{(0)} \equiv \lambda_k$ and $M_k^{(0)}(\mathbf{x}) \equiv M_k(\mathbf{x})$.

Applying the results of Section 4.2, we multiply (9.2.102) by $M_j(\mathbf{x})$ and integrate over G. Since both $M_j(\mathbf{x})$ and $M_k^{(1)}(\mathbf{x})$ satisfy the boundary condition (9.2.101) we obtain, in view of (4.1.28),

$$\iint_G [M_j L M_k^{(1)} - \lambda_k \rho M_j M_k^{(1)}]\, dv = \iint_G [M_k^{(1)} L M_j - \lambda_k \rho M_j M_k^{(1)}]\, dv$$

$$= (\lambda_j - \lambda_k)(M_k^{(1)}, M_j) = \iint_G r\rho M_j M_k\, dv + \lambda_k^{(1)}(M_j, M_k) \qquad (9.2.103)$$

with the inner product (f, g) defined as $(f(\mathbf{x}), g(\mathbf{x})) = \iint_G \rho(\mathbf{x}) f(\mathbf{x}) g(\mathbf{x})\, dv$. Now, if $j = k$ so that $\lambda_j = \lambda_k$, the left side of (9.2.103) must vanish and this implies that $\lambda_k^{(1)}$ must be specified as

$$\lambda_k^{(1)} = -\iint_G r(\mathbf{x})\rho(\mathbf{x}) M_k^2(\mathbf{x})\, dv, \qquad (9.2.104)$$

since $(M_k(\mathbf{x}), M_k(\mathbf{x})) = 1$. For $j \neq k$, the Fourier coefficients $(M_k^{(1)}(\mathbf{x}), M_j(\mathbf{x}))$ of the function $M_k^{(1)}(\mathbf{x})$ are given as

$$(M_k^{(1)}(\mathbf{x}), M_j(\mathbf{x})) = \frac{(r(\mathbf{x})M_k(\mathbf{x}), M_j(\mathbf{x}))}{\lambda_j - \lambda_k}, \qquad j \neq k, \qquad (9.2.105)$$

where we have used the fact that $(M_k(\mathbf{x}), M_j(\mathbf{x})) = 0$ for $j \neq k$.

The eigenfunctions $M_j(\mathbf{x})$ are assumed to form a complete set and $M_k^{(1)}(\mathbf{x})$ can be expanded in the series

$$M_k^{(1)}(\mathbf{x}) = \sum_{j=1}^{\infty} (M_k^{(1)}(\mathbf{x}), M_j(\mathbf{x})) M_j(\mathbf{x}), \qquad k = 1, 2, \ldots. \qquad (9.2.106)$$

Each of the Fourier coefficients is specified as in (9.2.105) except for $(M_k^{(1)}, M_k)$, which remains undetermined. It can be specified if we normalize each of the eigenfunctions $M_k(\mathbf{x}; \epsilon)$. We have

$$1 = (M_k(\mathbf{x}; \epsilon), M_k(\mathbf{x}; \epsilon)) = (M_k^{(0)}(\mathbf{x}), M_k^{(0)}(\mathbf{x}))$$

$$+2\epsilon(M_k^{(1)}(\mathbf{x}), M_k^{(0)}(\mathbf{x})) + O(\epsilon^2) = 1 + 2\epsilon(M_k^{(1)}(\mathbf{x}), M_k^{(0)}(\mathbf{x})) + O(\epsilon^2), \quad (9.2.107)$$

since the eigenfunctions $M_k^{(0)}(\mathbf{x}) \equiv M_k(\mathbf{x})$ are already normalized. On equating like powers of ϵ, we conclude that $(M_k^{(1)}(\mathbf{x}), M_k(\mathbf{x})) = 0$. Consequently, the series (9.2.106) for $M_k^{(1)}(\mathbf{x})$ is completely specified.

For each unperturbed eigenvalue λ_k and eigenfunction $M_k(\mathbf{x})$, we have determined the first perturbation $\epsilon\lambda_k^{(1)}$ and $\epsilon M_k^{(1)}(\mathbf{x})$. Higher approximations can be obtained by applying the finite Fourier transform method to the equations for $M_k^{(n)}(\mathbf{x})$ and proceeding as above. A modification of this procedure is needed to deal with *multiple*

eigenvalues. This is needed because the set of linearly independent eigenfunctions associated with the multiple eigenvalue must be prescribed in a special manner. The Sturm-Liouville eigenvalue problem has been shown to have simple eigenvalues. The results above can easily be expressed in a one-dimensional form. In higher-dimensional problems, multiple eigenvalues often occur. However, the foregoing perturbation method can be applied for every simple eigenvalue in a set of eigenvalues that may include multiple eigenvalues. The following example considers such a case.

Example 9.6. Eigenvalue Perturbations in a Square. We consider a square of side π, with the interior region G defined as $0 < x < \pi$ and $0 < y < \pi$, and examine the *perturbed eigenvalue problem*

$$-(M_{xx}(x,y;\epsilon) + M_{yy}(x,y;\epsilon)) + \epsilon xy M(x,y;\epsilon) = \lambda M(x,y;\epsilon), \ (x,y) \in G, \tag{9.2.108}$$

where $0 < \epsilon \ll 1$, with the boundary condition $M(x,y;\epsilon)\big|_{\partial G} = 0$. The reduced eigenvalue problem with $\epsilon = 0$ was considered in Example 7.6. It was found that the eigenvalues λ_{nm} and the normalized eigenfunctions $M_{nm}(x,y)$ are

$$\lambda_{nm} = n^2 + m^2, \ M_{nm}(x,y) = \frac{2}{\pi}\sin(nx)\sin(my), \tag{9.2.109}$$

for $n, m = 1, 2, \ldots$. We see that λ_{11} is a simple eigenvalue with only one linearly independent eigenfunction $M_{11}(x,y)$. However, λ_{12} equals λ_{21}, so that to this double eigenvalue there correspond two independent eigenfunctions $M_{12}(x,y)$ and $M_{21}(x,y)$. There are infinitely many other multiple eigenvalues.

As indicated previously, it is often of greatest interest to determine the lowest eigenvalue for a given problem. Since the lowest unperturbed eigenvalue λ_{11} is simple, we may apply the perturbation procedure and determine the corresponding perturbed eigenvalue as well as its eigenfunction. The situation is complicated somewhat because of the double-subscript notation for the eigenvalues and the eigenfunctions of the reduced eigenvalue problem. We shall obtain the first perturbation of the eigenvalue λ_{11} but not the perturbation of the eigenfunction $M_{11}(x,y)$.

In the notation of the previous discussion we have $\rho(x,y) = 1$ and $r(x,y) = -xy$. Then with $\lambda_{11}(\epsilon) = \lambda_{11} + \epsilon\lambda_{11}^{(1)} + O(\epsilon^2)$ representing the lowest (perturbed) eigenvalue, we have from (9.2.104),

$$\lambda_{11}^{(1)} = \frac{4}{\pi^2}\int_0^\pi \int_0^\pi xy\sin^2(x)\sin^2(y)\,dx\,dy = \frac{\pi^2}{4}, \tag{9.2.110}$$

so that $\lambda_{11}(\epsilon) = 2 + (\pi^2/4)\epsilon + O(\epsilon^2)$.

Nonlinear Dispersive Wave Motion

The *nonlinear Klein-Gordon equation*

$$w_{tt}(x,t) - \gamma^2 w_{xx}(x,t) + c^2 w(x,t) - \sigma w^3(x,t) = 0 \tag{9.2.111}$$

discussed in Example 9.3 was found to have an approximate solution of the form

$$w(x,t) = \epsilon \cos(kx) \cos\left[\left(\omega - \frac{9\epsilon^2\sigma}{32\omega}\right)t\right] + O(\epsilon^3), \tag{9.2.112}$$

as follows from (9.2.62) and (9.2.52). This can be rewritten as

$$w(x,t) = \frac{\epsilon}{2}\cos\left[kx - \left(\omega - \frac{9\epsilon^2\sigma}{32\omega}\right)t\right] + \frac{\epsilon}{2}\cos\left[kx + \left(\omega - \frac{9\epsilon^2\sigma}{32\omega}\right)t\right] + O(\epsilon^3). \tag{9.2.113}$$

[We recall that k is a constant and $\omega^2(k) = \gamma^2 k^2 + c^2$.]

Each of the terms in (9.2.113) represents a *traveling wave* of the form

$$w(x,t) = a \cos[kx \mp \hat{\omega}(k)t], \tag{9.2.114}$$

with a small *amplitude* term $a = \epsilon/2$ and a *phase* term $\theta = kx \mp \hat{\omega}(k)t$, where

$$\hat{\omega}(k) = \omega(k) - \frac{9\epsilon^2\sigma}{32\omega(k)} = \omega(k)\left[1 - \frac{9a^2\sigma}{8(\gamma^2k^2 + c^2)}\right]. \tag{9.2.115}$$

Each term has the *phase velocity*

$$\frac{dx}{dt} = \pm\frac{\hat{\omega}(k)}{k} = \pm\frac{\omega(k)}{k}\left[1 - \frac{9a^2\sigma}{8(\gamma^2k^2 + c^2)}\right]. \tag{9.2.116}$$

If the amplitude $a = \epsilon/2$ were assumed to be infinitesimal, the term involving a^2 in (9.2.115) would be neglected and the phase velocity would reduce to that for the *linear Klein-Gordon equation* as given in Example 3.8. However, for traveling waves with a small but finite amplitude, where a^2 is not neglected, we find that the phase speed or, correspondingly, the speed of the wave (9.2.114) depends not only on the wave number k but also on the amplitude a.

The dependence of the wave speed on its amplitude has already been observed in our discussion of *nonlinear unidirectional wave motion* in Chapter 2. The dependence of the speed of normal mode solutions of linear hyperbolic PDEs on the wave number k was characterized as *dispersive wave motion* in Section 3.5. The foregoing shows that the nonlinear Klein-Gordon equation has (approximate) traveling wave solutions whose wave speed depends on both the *wave number* and the *amplitude*.

The theory of *nonlinear dispersive wave motion* has undergone much study, especially by *Whitham*. One aspect that has been investigated is whether *periodic finite-amplitude traveling waves* of the form (9.2.114) or some more general form can be constructed for these equations. The amplitude of the waves is assumed to be small, so it can serve as a perturbation parameter. We assume that when a small-amplitude solution is inserted in the given PDE and it is then linearized, a dispersive wave equation results. The nonlinear terms introduce higher harmonics as in (9.2.60) and a dependence of the dispersion relation $\hat{\omega} = \hat{\omega}(k)$ on the amplitude as seen in (9.2.115).

We do not attempt to characterize the general form of nonlinear dispersive wave equations for which periodic traveling wave solutions can be found. Rather, we discuss a specific equation in the following example and consider further examples in the exercises. We refer to the literature for general results. One of the earliest applications of the perturbation method for determining periodic traveling waves for nonlinear equations was given by *Stokes* in his study of water waves. In the following example we construct approximate traveling wave solutions of the *nonlinear Korteweg-deVries equation*, which plays an important role in the *theory of water waves* (see Example 10.15). Exact periodic traveling wave solutions can be found for this equation. We do not exhibit them but apply the perturbation method instead. The main feature in our approach is that not only the solution but the dispersion function is expanded in powers of the small parameter, as suggested by (9.2.115).

Example 9.7. Periodic Traveling Wave Solutions of the Korteweg-deVries Equation. The *Korteweg-deVries equation* is given as

$$u_t(x,t) + (c + u(x,t))u_x(x,t) + \beta u_{xxx}(x,t) = 0, \tag{9.2.117}$$

where c and β are prescribed constants. In its linearized form,

$$u_t(x,t) + cu_x(x,t) + \beta u_{xxx}(x,t) = 0, \tag{9.2.118}$$

the *dispersion relation* is (see Section 3.5)

$$\omega = \omega(k) = ck - \beta k^3, \tag{9.2.119}$$

so that (9.2.118) is of *dispersive type*. If we drop the third derivative term in (9.2.117), we have

$$u_t(x,t) + (c + u(x,t))u_x(x,t) = 0, \tag{9.2.120}$$

a *quasilinear first order wave equation* whose wave speed depends on the amplitude.

The linearized equation (9.2.118) has traveling wave solutions

$$u(x,t) = a\cos(kx - \omega t), \tag{9.2.121}$$

where a = constant and ω is given by (9.2.119). The quasilinear equation (9.2.120) has the (implicit) solutions

$$u = a\cos[kx - k(c+u)t]. \tag{9.2.122}$$

The wave speed for (9.2.121) is $dx/dt = \omega/k = c - \beta k^2$, and the wave speed for (9.2.122) is, formally, $dx/dt = c + u$. In both cases the *wave speed* is perturbed around the constant c, in the linear case by a function of the *wave number* and in the nonlinear case by the *amplitude*. Note that if a is small, then $|u|$ is small in view of (9.2.122). We may also characterize $k(c+u)$ as a nonlinear frequency term, and

then (9.2.121)–(9.2.122) show that the frequency of the traveling wave solution for the full problem (9.2.117) depends on the wave number k and the amplitude a.

As a result, we look for a perturbation solution of (9.2.117) in the form

$$u(x,t) = \sum_{n=1}^{\infty} u_n(\theta) a^n, \tag{9.2.123}$$

where a is a small positive constant and

$$\theta(x,t) = kx - \hat{\omega}t, \quad \hat{\omega} = \sum_{n=0}^{\infty} \omega_n(k) a^n. \tag{9.2.124}$$

We attempt to determine the $u_n(\theta)$ to be periodic functions, so that (9.2.123) is a *periodic traveling wave solution* of (9.2.117). To do so it is necessary to specify the terms in series expansions of the frequency $\hat{\omega}$ appropriately. We have included a dependence on k and a for the frequency $\hat{\omega}$ in (9.2.124).

Inserting (9.2.123) and (9.2.124) into (9.2.117) yields the equations

$$(\omega_0 - ck)u_1'(\theta) - \beta k^3 u_1'''(\theta) = 0, \tag{9.2.125}$$

$$(\omega_0 - ck)u_2'(\theta) - \beta k^3 u_2'''(\theta) = k u_1(\theta) u_1'(\theta) - \omega_1 u_1'(\theta), \tag{9.2.126}$$

$$(\omega_0 - ck)u_3'(\theta) - \beta k^3 u_3'''(\theta) = k u_1(\theta) u_2'(\theta) + k u_2(\theta) u_1'(\theta) - \omega_1 u_2'(\theta) - \omega_2 u_1'(\theta), \tag{9.2.127}$$

on equating like powers of a.

As the solution of (9.2.125) we take $u_1(\theta) = \cos(\theta)$ and obtain $\omega_0(k) = ck - \beta k^3$. This yields a periodic traveling wave solution corresponding to the linearized equation (9.2.118). Neglecting higher powers of a in (9.2.123)–(9.2.124) yields (9.2.121). With this choice for $u_1(\theta)$, the right side of (9.2.126) can be expressed as $k u_1(\theta) u_1'(\theta) - \omega_1(k) u_1'(\theta) = -(k/2)\sin(2\theta) + \omega_1(k)\sin(\theta)$. The term $\omega_1(k)\sin(\theta)$ gives rise to a *secular term* proportional to $\theta \sin(\theta)$ in the expression for $u_2(\theta)$. We seek periodic solutions, so we must remove this term, as will be the case if we set $\omega_1(k) = 0$. Then a solution of (9.2.126) is $u_2(\theta) = (1/12\beta k^2)\cos(2\theta)$.

Continuing with (9.2.127) we obtain for the right side [since $\omega_1(k) = 0$]

$$k u_1(\theta) u_2'(\theta) + k u_2(\theta) u_1'(\theta) - \omega_2(k) u_1'(\theta) = \left[\omega_2(k) - \frac{1}{24\beta k}\right] \sin(\theta) - \frac{1}{8\beta k}\sin(3\theta). \tag{9.2.128}$$

Again the term involving $\sin(\theta)$ must be removed to avoid generating a *secular term* at the next level. Thus we set $\omega_2(k) = 1/24\beta k$, and the solution $u_3(\theta)$ is $u_3(\theta) = (1/192\beta^2 k^4)\cos(3\theta)$.

The full solution to order a^3 is

$$u(\theta) = a\cos(\theta) + a^2 \frac{1}{12\beta k^2}\cos(2\theta) + a^3 \frac{1}{192\beta^2 k^4}\cos(3\theta) + O(a^4), \tag{9.2.129}$$

with θ defined as in (9.2.124) and $\hat{\omega}$ given as $\hat{\omega} = \hat{\omega}(k, a) = ck - \beta k^3 + a^2/24\beta k + O(a^3)$. The frequency $\hat{\omega}$ and, consequently, the wave speed are seen to depend on both the wave number k and the amplitude a.

We conclude this example by noting that on inserting the expression $u = u(\theta)$—with θ given as in (9.2.124)—into the Korteweg-deVries equation, it is possible to find an exact periodic traveling wave solution in terms of elliptic functions. If that solution is expanded for small amplitudes, the preceding results are reproduced. We do not carry out this discussion, which is given in the literature.

Exercises 9.2

9.2.1. Use the perturbation method to solve the following boundary value problem. $u_{xx}(x, y) + u_{yy}(x, y) + \epsilon^2 u(x, y) = 1,\ x^2 + y^2 < 1,\ u(x, y) = 0,\ x^2 + y^2 = 1.$ Obtain the first two terms in the expansion.

9.2.2. Obtain the exact solution of the following boundary value problem for the Helmholtz equation in the unit sphere: $u_{xx}(x, y, z) + u_{yy}(x, y, z) + u_{zz}(x, y, z) + \epsilon^2 u(x, y) = 0,\ x^2 + y^2 + z^2 < 1,\ u(x, y, z) = 1,\ x^2 + y^2 + z^2 = 1.$ Also, solve the problem using perturbation theory and compare the exact solution with the approximate solution up to terms of $O(\epsilon^4)$.

9.2.3. Use the perturbation method to solve $[1 + \epsilon(x^2 + y^2)]u_{xx}(x, y) + u_{yy}(x, y) = 1,\ x^2 + y^2 < 1,\ u(x, y) = 0,\ x^2 + y^2 = 1.$

9.2.4. Apply the perturbation method to obtain an approximate solution of the following problem: $u_{xx}(x, y) + u_{yy}(x, y) - \epsilon^2 u^2(x, y) = 0,\ x^2 + y^2 < 1,\ u(x, y) = 1,\ x^2 + y^2 = 1.$

9.2.5. Let $u(x, y)$ satisfy $u_{xx}(x, y) + [1 + \epsilon y]u_{yy}(x, y) = 0,\ 0 < x, y < \pi$, and the boundary conditions $u(0, y) = u(\pi, y) = u(x, \pi) = 0,\ u(x, 0) = 1,\ 0 < x < \pi$. Use the perturbation method to obtain an approximate solution.

9.2.6. Given the exterior boundary value problem for the modified Helmholtz equation $\nabla^2 u(x, y, z) - \epsilon^2 u(x, y, z) = 0,\ r > 1$, with the conditions $u(x, y, z)\big|_{r=1} = e^{-\epsilon}$, $\lim_{r\to\infty} u(x, y, z) = 0$, where $r^2 = x^2 + y^2 + z^2$, use a perturbation expansion to solve the problem and show that it is not possible to make each term in the series vanish as $r \to \infty$. Determine that the series exhibits secular behavior and use the method of multiple scales to overcome this difficulty and thereby obtain the exact solution of the problem.

9.2.7. Let $K(x, y, z; \xi, \eta, \zeta)$ represent the free space Green's function for the modified Helmholtz equation $\nabla^2 K(x, y, z; \xi, \eta, \zeta) - \epsilon^2 K(x, y, z; \xi, \eta, \zeta) = -\delta(x - \xi)\delta(y - \eta)\delta(z - \zeta)$, with $-\infty < x, y, z, \xi, \eta, \zeta < \infty$. Assuming that ϵ is small, apply the perturbation method to obtain $K(x, y, z; \xi, \eta, \zeta)$ in terms of the Green's function for Laplace's equation. Using the results of Example 6.13, discuss the behavior of the exact and the approximate Green's functions at infinity. Explain the differences

in their behavior at infinity as a result of the occurrence of secular terms in the perturbation series.

9.2.8. Use the perturbation method (for small ϵ) to solve Laplace's equation $\nabla^2 u(r,\theta) = 0$ in the disk $r^2 = x^2 + y^2 < 1$ with the boundary conditions: (a) $\epsilon\, \partial u(1,\theta)/\partial r + u(1,\theta) = 1$; (b) $\partial u(1,\theta)/\partial r + \epsilon\, u(1,\theta) = 1$. Show that the problem of part (b) cannot be solved by using the perturbation method and obtain the exact solution of the problem to show why the regular perturbation theory fails.

9.2.9. Consider the Robin problem for the elliptic equation $-\nabla \cdot (p(\mathbf{x})\nabla u(\mathbf{x})) + q(\mathbf{x})u(\mathbf{x}) = \rho(\mathbf{x})F(\mathbf{x})$ in a region G with the boundary condition $\alpha(\mathbf{x})u(\mathbf{x}) + \beta(\mathbf{x})\, \partial u(\mathbf{x})/\partial n\big|_{\partial G} = B(\mathbf{x})$, where $\alpha(\mathbf{x}) > 0$ and $\beta(\mathbf{x}) > 0$. Assuming that either $\alpha(\mathbf{x})$ or $\beta(\mathbf{x})$ is uniformly small on ∂G, construct a perturbation method whereby the solution of the given third boundary value problem can be approximated by solutions of Dirichlet or Neumann problems. Comment on the possibility that the data for the reduced Neumann problem are incompatible and explain how to resolve that difficulty [see Exercise 9.2.8(b)].

9.2.10. Use the perturbation method to solve the Cauchy problem for the parabolic equation $u_t(x,t) + \epsilon u_x(x,t) = u_{xx}(x,t)$, $-\infty < x < \infty$, $t > 0$, with the initial value $u(x,0) = f(x)$, $-\infty < x < \infty$. Determine that the perturbation series exhibits secular behavior and use the method of multiple scales to deal with this problem. Also, convert the given problem to an initial value problem for the heat equation by introducing the change of variable $\tau = t$, $\sigma = x - \epsilon t$. Obtain the solution of the transformed problem and compare it with the perturbation result.

9.2.11. Solve the Cauchy problem for the weakly damped wave equation $u_{tt}(x,t) + \epsilon u_t(x,t) - c^2 u_{xx}(x,t) = 0$, $-\infty < x < \infty$, $t > 0$, with the initial data $u(x,0) = f(x)$, $u_t(x,0) = 0$, $-\infty < x < \infty$, using perturbation theory. Show that secular terms arise in the perturbation expansion and use the method of multiple scales to eliminate them.

9.2.12. Construct a solution of the Cauchy problem (9.2.53)–(9.2.54) in the form $u(x,t) = \sum_{n=0}^{\infty} \epsilon^{2n} v_n(t) \cos[(2n+1)kx]$. Note that this should lead to the expansion (9.2.60), but show that $v_0(t)$ can be specified so that the $O(\epsilon^2)$ secular term is eliminated. Indicate how $v_1(t)$ must be chosen to eliminate secular behavior at the $O(\epsilon^4)$ level. *Hint*: Expand $v_0(t)$ and $v_1(t)$ in a series in powers of ϵ^2.

9.2.13. Show that if $c(\epsilon x)$ in the hyperbolic equation (9.2.63) is given as in (9.2.78), the equation can be transformed into an equation with constant coefficients in the region $x > 0$. Let $\sigma = (1/\epsilon)\log[(c_0 + \epsilon x)/c_0]$ and $\tau = t$ in (9.2.63) and obtain the equation $v_{\tau\tau}(\sigma,\tau) - v_{\sigma\sigma}(\sigma,\tau) + \epsilon v_\sigma(\sigma,\tau) = 0$. If the data for the wave equation (9.2.63) are $u(0,t) = 0$, $u(x,0) = f(x)$, and $u_t(x,0) = 0$, so that the problem is given in the region $x > 0$, $t > 0$, determine the appropriate region and data for the transformed problem for $v(\sigma,\tau)$. Discuss the approximate solution of the initial and boundary value problem for $v(\sigma,\tau)$ in the region $\sigma - \tau > 0$ and compare with the result (9.2.81). *Hint*: Do not expand the transformed initial data in powers of ϵ on using a perturbation series to solve for $v(\sigma,\tau)$. Show that secular terms arise in the

perturbation series and apply the method of multiple scales in a judicious manner to reproduce the result in the text.

9.2.14. Consider the initial and boundary value problem for the parabolic equation $u_t(x,t) - (c_0 + \epsilon x)^2 u_{xx}(x,t)$, $x > 0$, $t > 0$, where $c_0 > 0$ and $0 < \epsilon \ll 1$, with the data $u(0,t) = 0$, $u(x,0) = f(x)$. Solve by using the perturbation method and show that secular terms occur in the perturbation series. Attempt to eliminate this difficulty by using the method of multiple scales. Use the transformation given in Exercise 9.2.13 to transform the problem into one with constant coefficients that can be solved exactly.

9.2.15. Using the boundary perturbation method, solve the Neumann problem for Laplace's equation $u_{xx}(x,y) + u_{yy}(x,y) = 0$, $0 < x < l$, $\epsilon x < y < L$, with the data $\partial u(0,y)/\partial x = \partial u(l,y)/\partial x = \partial u(x,L)/\partial y = 0$, $\partial u(x,\epsilon x)/\partial n = \cos(\pi x/l)$, $0 < x < l$.

9.2.16. Solve the Dirichlet problem using the boundary perturbation: $u_{xx}(x,y) + u_{yy}(x,y) = 0$, $\epsilon \sin y < x < l$, $y > 0$, with $u(\epsilon \sin(y), y) = u(l,y) = 0$, $y > 0$, $u(x,0) = \sin(\pi x/l)$, $0 < x < l$, $u(x,y)$ bounded as $y \to \infty$.

9.2.17. Apply the boundary perturbation method to solve the following problem: $u_{tt}(x,t) - c^2 u_{xx}(x,t) = 0$, $\epsilon t < x < \infty$, $t > 0$, $u(x,0) = u_t(x,0) = 0$, $x > 0$, $u(\epsilon t, t) = A\cos(\omega t)$, $t > 0$. Show that secular terms result in the perturbation series. Compare the perturbation result with the exact solution given in (6.5.25) when $c_0 = \epsilon$.

9.2.18. Solve the Dirichlet problem for Laplace's equation $\nabla^2 u(x,y) = 0$ in an ellipse $x = (1+a\epsilon)\cos\sigma$, $y = (1+b\epsilon)\sin(\sigma)$ $(0 \le \sigma < 2\pi)$ (i.e., a slightly perturbed unit circle) using the boundary perturbation method. The boundary condition on the ellipse is $u(x,y) = x^2$.

9.2.19. Use the eigenvalue perturbation method to determine up to $O(\epsilon)$ terms the leading eigenvalue and eigenfunction for the following problem. $M''(x) + \lambda(1 - \epsilon \sin^2 x)M(x) = 0$, $0 < x < \pi$, $M(0) = M(\pi) = 0$.

9.2.20. Approximate the leading eigenvalue for the following eigenvalue problem: $-\nabla^2 M(x,y) + \epsilon xy M(x,y) = \lambda M(x,y)$, $0 < x < \pi$, $0 < y < \pi$, $M(0,y) = M(\pi,y) = \partial M(x,0)/\partial y = \partial M(x,\pi)/\partial y = 0$.

9.2.21. Use the results of Example 8.3 to approximate the leading eigenvalue for the following problem in the ball $r^2 = x^2 + y^2 + z^2 \le 1$: $\nabla^2 M(x,y,z) + (\lambda - \epsilon r^2)M(x,y,z) = 0$, $r < 1$, $M(x,y,z) = 0$, $r = 1$.

9.2.22. Develop a perturbation method to obtain the eigenvalues for the following Sturm-Liouville problems: $-d/dx\,(p(x)dM(x)/dx) + q(x)M(x) = \lambda M(x)$, $0 < x < l$, with the boundary conditions (a) $\epsilon M'(0) - M(0) = 0$, $M(l) = 0$; (b) $M'(0) - \epsilon M(0) = 0$, $M(l) = 0$.

9.2.23. Apply the method of Exercise 9.2.22 to obtain the leading eigenvalue for each of the following problems: $M''(x) + \lambda M(x) = 0$, $0 < x < \pi$, with the boundary conditions (a) $\epsilon M'(0) - M(0) = 0$, $M(\pi) = 0$; (b) $M'(0) - \epsilon M(0) = 0$, $M(\pi) = 0$.

9.2.24. Reproduce (9.2.113)–(9.2.116) by solving the nonlinear Klein-Gordon equation (9.2.109) in the form $w(x,t) = \sum_{n=1}^{\infty} w_n(\theta)\epsilon^n + \sum_{n=1}^{\infty} w_n(\hat{\theta})\ \epsilon^n$, where $\theta = kx - \hat{\omega}t$, $\tilde{\theta} = kx - \tilde{\omega}t$ and $\hat{\omega} = \sum_{n=1}^{\infty} \omega_n(k)\epsilon^n$, $\tilde{\omega} = \sum_{n=1}^{\infty} \tilde{\omega}_n(k)\epsilon^n$. Proceed as in Example 9.7 and choose the $\omega_n(k)$ and $\tilde{\omega}_n(k)$ to eliminate secular terms in the expansion of $w(x,t)$.

9.2.25. Obtain traveling wave solutions of the nonlinear hyperbolic equation $w_{tt}(x,t) - \gamma^2 w_{xx}(x,t) - c^2 w(x,t) - \sigma w^3(x,t) = 0$, by using the expansion forms given in Exercise 9.2.24. (Note that the linearized version of this equation does not have a real dispersion relation for all values of k. Consider the problem only for those k that yield traveling waves in the linearized case.)

9.2.26. The nonlinear system $\{E_{tt}(x,t) + P_{tt}(x,t) = \gamma^2 E_{xx}(x,t),\ P_{tt}(x,t) + P(x,t) - \sigma P^3(x,t) = c^2 E(x,t)\}$ occurs in the theory of nonlinear optics. Show that if we consider the linear case and set $\sigma = 0$, and look for a solution $E = E(\theta)$, $P = P(\theta)$, where $\theta = kx - \omega t$, we obtain the dispersion relation $k^2 = \omega^2/\gamma^2 - c^2\omega^2/[\gamma^2(\omega^2 - 1)]$. For the (weakly) nonlinear case look for a solution of the given system in the form $E(\theta) = a\cos\theta + a_1 \cos 3\theta + \cdots$, $P(\theta) = b\cos\theta + b_1 \cos 3\theta + \cdots$, where $\theta = kx - \omega t$ and the constants a, b, a_1, and b_1 are small, but a_1 and b_1 are smaller in magnitude than a and b. Insert these expansions into the nonlinear equations and equate the coefficients of like trigonometric terms. From the coefficients of the $\cos\theta$ terms obtain an expression giving b in terms of a and an equation for $\alpha = \gamma^2 k^2 - \omega^2$. Expand α in powers of a^2 and show that $k^2 \approx \omega^2/\gamma^2 - c^2\omega^2/[\gamma^2(\omega^2 - 1)] + 3\sigma c^6\omega^2/[4\gamma^2(\omega^2 - 1)^4]\, a^2$.

9.3 SINGULAR PERTURBATION METHODS AND BOUNDARY LAYER THEORY

In *singular perturbation theory* we are concerned with the study of PDEs that contain a small parameter that multiplies one or more of the highest derivative terms in the equations. Thus when that parameter is equated to zero, giving rise to the *reduced equation*, either the order or the type (or both the order and type) of the given equation is changed. Generally, this means that a regular perturbation series solution proves inadequate to handle the initial and/or boundary data for the given problem. It can also mean that solutions of the reduced equation are singular at a point, curve, or surface on which solutions of the given problem are not expected to be singular. Rather than discarding the perturbation series completely, it is possible to introduce boundary, initial or internal layers across which the solution of the given problem undergoes a rapid transition from a form that satisfies all the data given for the problem or is nonsingular, to a form represented by the perturbation series. The determination of the boundary layers (initial or interior layers are often be referred to as boundary layers, as well) and the approximate forms of the given equations in those regions forms the subject of *boundary layer theory*. The procedure whereby solutions valid in the boundary layers are identified with the perturbation series solution valid in the so-called outer region(s) is often called the *matching process*. The general process of determining perturbation and boundary layer expansions and matching these results is

sometimes referred to as the *method of matched asymptotic expansions*. By combining the perturbation and boundary layer solutions, a fairly good approximate description of the solution of the given problem can often be found for problems where the exact solution is difficult or impossible to determine or where the solution is not easy to interpret or evaluate.

It should be noted that singular perturbation theory is sometimes taken to encompass any problem where regular perturbation theory is inadequate for any reason. This may not involve the presence of a small parameter multiplying the highest derivative, but may be due to the presence of secular terms that result in the nonuniformity of the solution over an infinite region or the occurrence of a small parameter in the data for the problem (examples of these types have been given in Section 9.2). Nevertheless, we shall restrict our discussion in this section to the type of problem discussed in the preceding paragraph.

Singular Perturbations and Boundary Layers for First Order PDEs

As it is somewhat complicated to present a general theory that encompasses all types of singular perturbation and boundary layer problems, we begin by considering two simple examples for first order PDEs. These problems are analyzed in some detail. Then we consider second and higher-order equations from a general point of view and in a number of examples.

Example 9.8. Singular Perturbation of a First Order Equation. The *initial value problem* for the equation

$$\epsilon(u_t(x,t) + u_x(x,t)) + u(x,t) = \sin(t), \qquad -\infty < x < \infty,\ t > 0, \tag{9.3.1}$$

with

$$u(x,0) = f(x), \tag{9.3.2}$$

where $0 < \epsilon \ll 1$ and $f(x)$ is a prescribed smooth function, has the solution

$$u(x,t) = \frac{1}{1+\epsilon^2}(\sin(t) - \epsilon\cos(t)) + \left[f(x-t) + \frac{\epsilon}{1+\epsilon^2}\right]\exp\left(-\frac{t}{\epsilon}\right), \tag{9.3.3}$$

as is easily verified.

We attempt to solve (9.3.1)—(9.3.2) by using a conventional perturbation series and set

$$u(x,t) = \sum_{n=0}^{\infty} u_n(x,t)\epsilon^n. \tag{9.3.4}$$

Inserting (9.3.4) into (9.3.1) and equating like powers of ϵ yields the system

$$u_0(x,t) = \sin(t),\ u_n(x,t) = -\left(\frac{\partial u_{n-1}(x,t)}{\partial t} + \frac{\partial u_{n-1}(x,t)}{\partial x}\right),\ n \geq 1. \tag{9.3.5}$$

We easily conclude that

$$u_{2n}(x,t) = (-1)^n\sin(t), \quad u_{2n+1}(x,t) = (-1)^{n+1}\cos(t), \quad n \geq 0. \tag{9.3.6}$$

The terms in the perturbation series (9.3.4) are uniquely specified without regard to the initial value $u(x,0)$. Furthermore, the perturbation series

$$u(x,t) = \left[\sum_{n=0}^{\infty}(-1)^n\epsilon^{2n}\right][\sin(t) - \epsilon\cos(t)] = \frac{1}{1+\epsilon^2}[\sin(t) - \epsilon\cos(t)] \quad (9.3.7)$$

certainly does not satisfy the initial condition (9.3.2). The reason is that the *reduced problem* for (9.3.1), where ϵ equated to zero, is not even a differential equation and cannot absorb arbitrary initial values. This reduced problem characterizes the form of all equations arising from the conventional perturbation approach, and none of them is a differential equation.

On comparing the perturbation solution (9.3.7) and the exact solution (9.3.3), we see that difference between both solutions is a term significant only in the region $0 < t \leq O(\epsilon)$. That is, for $t > O(\epsilon)$, the exponential $\exp(-t/\epsilon)$ is small and can be neglected. Then the solution of (9.3.1)–(9.3.2) is well approximated by the perturbation result (9.3.8). However, within a layer of width $O(\epsilon)$ near the x-axis, the exponential term in (9.3.3) is significant and must be retained. It is this term in combination with the perturbation result (9.3.7) that enables the solution to satisfy the initial condition.

The existence of an initial layer of width $O(\epsilon)$, where the conventional perturbation series is not valid, may be inferred from the fact that for $t \approx \epsilon$ we have $\sin(t) \approx \sin(\epsilon) \approx \epsilon$ and $\cos(t) \approx \cos(\epsilon) \approx 1$. Thus the first two terms in the perturbation series behave like $u(x,t) \approx \sin(t) - \epsilon\cos(t) \approx \epsilon - \epsilon = 0$, $t \approx \epsilon$ (i.e., they are of the same order in ϵ). Therefore, the perturbation series (9.3.7) is not well ordered in the region where $t = O(\epsilon)$ and it is not expected to be a valid representation of the solution. Of course, we know that (9.3.7) is not valid near $t = 0$, since it fails to satisfy the initial condition. However, the foregoing argument yields an approximate description of the size of the region where the conventional perturbation method is invalid. This argument is similar to that given in Section 9.2 in connection with secular terms.

The significant conclusion that we have reached is that the perturbation result need not be discarded completely because of its failure to satisfy the initial condition. It need only be replaced by a different or modified approximation in an *initial layer* of width $O(\epsilon)$ near the x-axis. To study the equation in the initial layer, we introduce the *stretching transformation*

$$t = \epsilon^r\tau, \quad (9.3.8)$$

where the positive constant r is to be specified. Since ϵ is small, the τ variable is large even for small or moderate values of t. Thus the region near the x-axis is stretched out. The equation (9.3.8) indicates that we wish to study (9.3.1) in a region where $t = O(\epsilon^r)$. We already have shown that the choice $r = 1$ is appropriate for our problem, but now we wish to show this directly from the equation (9.3.1) by using *boundary layer arguments*.

On using (9.3.8) in (9.3.1) with $\hat{u}(x,\tau) = u(x,\epsilon^r\tau)$, we have

$$\epsilon^{1-r}\hat{u}_\tau(x,\tau) + \epsilon\hat{u}_x(x,\tau) + \hat{u}(x,\tau) = \sin(\epsilon^r\tau) = \epsilon^r\tau - \frac{\epsilon^{3r}\tau^3}{6} + O(\epsilon^{5r}). \quad (9.3.9)$$

Now if $r = 1$, there is a balance between the terms $\hat{u}_\tau$ and $\hat{u}$ in (9.3.9), and they represent the leading terms for small ϵ in that equation. We retain the $\hat{u}_\tau$ term to obtain a differential equation in the initial layer that can absorb an initial condition. We also want to retain the term $\hat{u}$ that occurs in the reduced equation for (9.3.1), so that a smooth transition from the initial layer to the outer region, where the perturbation series (9.3.7) is valid, is expected to result. With $r > 1$, $\hat{u}_\tau$ is the leading term. Although this choice of r leads to equations that can account for the initial condition, the initial layer seems to be too thin since the equations do not retain any part of the reduced equation. For $r < 1$, $\hat{u}$ is the leading term and nothing has been accomplished.

Putting $r = 1$ in (9.3.9), we obtain

$$\hat{u}_\tau(x,\tau) + \hat{u}(x,\tau) + \epsilon\hat{u}_x(x,\tau) = \sin(\epsilon\tau) = \epsilon\tau - \frac{\epsilon^3\tau^3}{6} + O(\epsilon^5). \tag{9.3.10}$$

The initial condition at $\tau = 0$ is $\hat{u}(x,0) = f(x)$.

We solve (9.3.10) by the perturbation method and set

$$\hat{u}(x,\tau) = \sum_{n=0}^{\infty} \hat{u}_n(x,\tau)\epsilon^n. \tag{9.3.11}$$

Inserting (9.3.11) into (9.3.10) and equating like powers of ϵ yields

$$\frac{\partial\hat{u}_0(x,\tau)}{\partial\tau} + \hat{u}_0(x,\tau) = 0, \quad \frac{\partial\hat{u}_1(x,\tau)}{\partial\tau} + \hat{u}_1(x,\tau) = \tau - \frac{\partial\hat{u}_0(x,\tau)}{\partial x} \tag{9.3.12}$$

as the two leading equations. They are effectively ODEs in the variable τ, and the initial conditions for the equations for all the $\hat{u}_n(x,t)$ at $\tau = 0$ are

$$\hat{u}_0(x,0) = f(x), \qquad \hat{u}_n(x,0) = 0, \; n \geq 1. \tag{9.3.13}$$

We remark that it is often the case that the boundary layer equations are ODEs even though the perturbation problem is formulated for a PDE. The first two boundary layer terms $\hat{u}_0(x,\tau)$ and $\hat{u}_1(x,\tau)$ are found to be

$$\hat{u}_0(x,\tau) = f(x)e^{-\tau}, \qquad \hat{u}_1(x,\tau) = \tau - 1 + [1 - \tau f'(x)]e^{-\tau}. \tag{9.3.14}$$

Thus

$$\hat{u}(x,\tau) \approx f(x)e^{-\tau} + \epsilon\{\tau - 1 + [1 - \tau f'(x)]e^{-\tau}\} + O(\epsilon^2). \tag{9.3.15}$$

In general, we would now apply the *matching process* to specify unknown quantities that occur in the outer perturbation expansion. This is carried out by assuming that the perturbation series (9.3.7) and the boundary layer (perturbation) series (9.3.11) have a common region of validity. Then both series are expressed in terms of a common set of variables (say, by expressing the unstretched variable in the perturbation expansion in terms of the stretched variable in the boundary layer expansion, so that t is replaced $\epsilon\tau$ in this example). The modified series are expanded in powers of

ϵ, terms with like powers of ϵ in both series are identified, and unknown quantities are determined. (Exponentially small terms are discarded in this process.) Here, however, the (outer) perturbation series is completely specified so that matching is unnecessary to complete the solution. Nevertheless, we carry out the matching procedure to show how the solution undergoes a transition from its boundary layer form to the form valid in the outer region where $t > O(\epsilon)$.

We have

$$u(x,t) \approx \sum_{n=0}^{\infty} u_n(x,\epsilon\tau)\epsilon^n \approx \sum_{n=0}^{\infty} \hat{u}_n(x,\tau)\epsilon^n \tag{9.3.16}$$

in a common region of validity of both series that is assumed to exist. Thus

$$u_0(x,\epsilon\tau) + \epsilon u_1(x,\epsilon\tau) = \sin(\epsilon\tau) - \epsilon\cos(\epsilon\tau) \approx \epsilon\tau - \epsilon + O(\epsilon^3), \tag{9.3.17}$$

$$\hat{u}_0(x,\tau) + \epsilon\hat{u}_1(x,\tau) = f(x)e^{-\tau} + \epsilon[\tau - 1] + \epsilon[1 - \tau f'(x)]e^{-\tau} \approx \epsilon\tau - \epsilon. \tag{9.3.18}$$

For the purpose of matching we have assumed that τ is large, so that the exponential $e^{-\tau}$ can be neglected, but that $\epsilon\tau$ is small, so that $\sin(\epsilon\tau)$ and $\cos(\epsilon\tau)$ can be approximated by the leading terms in their power series expansions. This procedure can be formalized by assuming, for example, that $\tau = O(\epsilon^{-1/2})$, so that τ is large for small ϵ, but $\epsilon\tau = O(\epsilon^{1/2})$, which is small. Then the region where the matching is carried out corresponds to $t = \epsilon\tau = O(\epsilon^{1/2})$. The entire matching process can be carried out in a systematic manner by using the *method of intermediate limits*. However, we do not require this method, as we consider only leading terms in our analysis, and do not discuss this technique, which can be found in the literature.

On comparing (9.3.17) and (9.3.18), we see that they both agree to the order of ϵ retained. Thus, it follows from our perturbation and boundary layer results that the approximate solution $u(x,t)$ of (9.3.1)–(9.3.3) is

$$u(x,t) \approx \begin{cases} t - \epsilon + [f(x) - tf'(x) + \epsilon]e^{-t/\epsilon}, & 0 \le t \le O(\epsilon), \\ \sin(t) - \epsilon\cos(t) + O(\epsilon^2), & t > O(\epsilon). \end{cases} \tag{9.3.19}$$

In the initial layer of width $O(\epsilon)$, the solution undergoes a rapid transition from a form that satisfies the initial condition to the (outer) perturbation approximation. The solution in the initial layer corresponds to the expansion of the exact solution (9.3.3) for small t, as is easily verified. The fact that f and its derivatives are given as functions of x rather than $x - t$—corresponding to the characteristics $x - t =$ constant of (9.3.1)—as is the case in the exact solution (9.3.3) is not that important. Indeed, since $f(x-t) = f(x) - tf'(x) + O(t^2)$, we can replace $f(x) - tf'(x)$ by $f(x-t)$ for $t \le O(\epsilon)$. Because of the rapid exponential decay of the term that involves f, either of these representations can be used.

We observe that the approximate solution (9.3.19) can be combined into a single *composite expression* that is (uniformly) valid for all t. We write

$$u(x,t) \approx \sin(t) - \epsilon\cos(t) + [f(x) - tf'(x) + \epsilon]e^{-t/\epsilon} + O(\epsilon^2). \tag{9.3.20}$$

The initial condition is satisfied approximately. Within the initial layer of width $O(\epsilon)$, the trigonometric terms can be approximated by $t - \epsilon$, and in the outer region we discard the exponentially small terms so that we retrieve the boundary layer and outer perturbation results. The expression (9.3.20) represents the leading terms of a *composite expansion* which can be constructed in a systematic manner from the terms in the boundary layer and outer expansions. We do not pursue this matter here any further.

An alternative approach that leads to a different boundary layer construction and expansion is based on the following. We have determined that when $t = \epsilon\tau$ the perturbation series (9.3.7) becomes disordered. Then if we substitute $t = \epsilon\tau$ in (9.3.7), we obtain $u(x,t) \approx (\tau - 1)\epsilon + O(\epsilon^3)$, so that $u(x,t) \approx O(\epsilon)$ in that region. Therefore, we put $u(x,t) = \epsilon v(x,\tau)$ in the boundary layer region and obtain, in place of (9.3.10),

$$v_\tau(x,\tau) + v(x,\tau) + \epsilon v_x(x,\tau) = \sin(\epsilon\tau)/\epsilon = \tau + \frac{\epsilon^2\tau^3}{6} + O(\epsilon^4). \qquad (9.3.21)$$

On expanding $v(x,\tau)$ as $v(x,\tau) = \sum_{n=0}^{\infty} v_n(x,\tau)\epsilon^n$, we obtain

$$\frac{\partial v_0(x,\tau)}{\partial\tau} + v_0(x,\tau) = \tau. \qquad (9.3.22)$$

A general solution of this equation is $v_0(x,\tau) = \tau - 1 + F(x)e^{-\tau}$ with arbitrary $F(x)$. Then, since $u(x,t) \approx \epsilon v_0(x,\tau)$ and $u(x,0) = f(x)$, we obtain the boundary layer approximation

$$u(x,t) \approx t - \epsilon + [f(x) + \epsilon]e^{-t/\epsilon}, \qquad 0 \le t \le O(\epsilon). \qquad (9.3.23)$$

This result corresponds to replacing $f(x - t)$ by $f(x)$ in the modified boundary layer approximation given above. However, to obtain (9.3.23) we must put $F(x) = 1 + f(x)/\epsilon$ and, clearly, $F(x)$ should not have any ϵ dependence. Thus, we conclude that the original boundary layer analysis presented above is to be preferred.

To conclude this example, we note that if the term $u(x,t)$ in (9.3.1) is replaced by $-u(x,t)$, we may still use the regular perturbation method to construct a series solution of the form (9.3.4). However, a boundary layer analysis yields as the leading order boundary layer equation $\partial\hat{u}_0(x,\tau)/\partial\tau - \hat{u}_0(x,\tau) = 0$, as is easily seen. The solution of this equation that satisfies the initial condition is $\hat{u}_0(x,\tau) = f(x)e^{\tau}$. It grows exponentially as τ increases so that we do not obtain an initial layer within which a rapid transition of the solution to the form of the outer perturbation expansion takes place. The reason for the breakdown of the perturbation method and the boundary layer theory is that the initial value problem for (9.3.1) with the term $u(x,t)$ replaced by $-u(x,t)$ is *unstable* and has exponentially growing solutions. This is easily shown by finding the exact solution of the initial value problem or by the use of a stability analysis.

The general first order linear equation

$$\epsilon(a(x,t)u_t(x,t)+b(x,t)u_x(x,t))+c(x,t)u(x,t) = d(x,t),\ -\infty < x < \infty,\ t > 0, \tag{9.3.24}$$

where $0 < \epsilon \ll 1$ and $a(x,t)$, $b(x,t)$, $c(x,t)$, and $d(x,t)$ are specified functions, with the initial condition

$$u(x,0) = f(x), \tag{9.3.25}$$

yields a singular perturbation problem that may require for its solution the introduction of additional *internal boundary layers* apart from the *initial layer*. Thus if $c(x,t)$ vanishes on a curve in the region $t > 0$ where $a(x,t)$ and $b(x,t)$ do not vanish, the conventional perturbation series may become singular there. However, the solution of the full problem (9.3.24)–(9.3.25) is not expected to be singular along that curve. Consequently, a modified result that can be obtained by boundary layer methods must be introduced. In the following example a problem of this type is considered.

Example 9.9. An Internal Boundary Layer. The first order PDE

$$\epsilon(u_t(x,t) + u_x(x,t)) + (t-1)^2u(x,t) = 1, \qquad -\infty < x < \infty,\ t > 0, \tag{9.3.26}$$

with the initial condition

$$u(x,0) = 0, \qquad -\infty < x < \infty, \tag{9.3.27}$$

has the exact solution

$$u(x,t) = \frac{1}{\epsilon}\exp\left[-\frac{(t-1)^3}{3\epsilon}\right]\int_0^t \exp\left[\frac{(s-1)^3}{3\epsilon}\right]ds, \tag{9.3.28}$$

as is easily verified. Because of the x-independent initial condition, we may consider (9.3.26)–(9.3.27) to be an ODE problem. It is, nevertheless, of interest to study the phenomena that occur for the perturbation solution of this problem, as they occur for more general problems as well.

We begin by considering the *conventional perturbation series* solution

$$u(x,t) = \sum_{n=0}^{\infty} u_n(x,t)\epsilon^n \tag{9.3.29}$$

of (9.3.26)–(9.3.27). The $u_n(x,t)$ satisfy algebraic rather than differential equations, and we easily obtain as the leading terms of (9.3.29)

$$u(x,t) \approx u_0(x,t) + \epsilon u_1(x,t) = \frac{1}{(t-1)^2} + \frac{2\epsilon}{(t-1)^5}. \tag{9.3.30}$$

Not only does (9.3.30) not satisfy the initial condition (9.3.27) but it also is singular at $t = 1$. The given equation (9.3.26) has no singularity there.

To handle the initial condition we introduce the *stretching transformation*

$$t = \epsilon\tau, \tag{9.3.31}$$

which yields for (9.3.26)

$$\hat{u}_\tau(x,\tau) + \epsilon\hat{u}_x(x,\tau) + \hat{u}(x,\tau) - 2\epsilon\tau\hat{u}(x,\tau) + \epsilon^2\tau^2\hat{u}(x,\tau) = 1, \tag{9.3.32}$$

where $\hat{u}(x,\tau) = u(x,\epsilon\tau)$. To obtain (9.3.31) we can use the approach given in Example 9.8 (see (9.3.8)–(9.3.9)) and first set $t = \epsilon^r\tau$ and conclude that we must have $r = 1$. With

$$\hat{u}(x,\tau) = \sum_{n=0}^{\infty} \hat{u}_n(x,\tau)\epsilon^n \tag{9.3.33}$$

we have

$$\frac{\partial\hat{u}_0(x,\tau)}{\partial\tau} + \hat{u}_0(x,\tau) = 1, \quad \frac{\partial\hat{u}_1(x,\tau)}{\partial\tau} + \hat{u}_1(x,\tau) = 2\tau\hat{u}_0(x,\tau) - \frac{\partial\hat{u}_0(x,\tau)}{\partial x} \tag{9.3.34}$$

as the leading equations. The initial data for the $\hat{u}_n(x,\tau)$ are

$$\hat{u}_n(x,0) = 0, \qquad n \geq 0. \tag{9.3.35}$$

The solutions are $\hat{u}_0(x,\tau) = 1 - e^{-\tau}$ and $\hat{u}_1(x,\tau) = 2(\tau - 1) + (2 - \tau^2)e^{-\tau}$, so that

$$\hat{u}(x,\tau) \approx 1 - e^{-\tau} + \epsilon[2(\tau - 1) + (2 - \tau^2)e^{-\tau}]. \tag{9.3.36}$$

It is readily verified that (9.3.30) and (9.3.36) match one another if (9.3.30) is expanded for small t and (9.3.36) is expanded for large τ.

At t approaches unity, the (outer) perturbation series (9.3.30) blows up. To determine the values of t where (9.3.30) first begins to break down, we note that when $t - 1 = O(\epsilon^{1/3})$, the perturbation series (9.3.30) becomes disordered in that the second term is of the same order in ϵ as the leading term. This suggests that there exists an *internal boundary layer* near $t = 1$ whose width is given as $|t - 1| = O(\epsilon^{1/3})$. Further, we find that within the boundary layer $u(x,t) \approx O[(t-1)^{-2}] = O(\epsilon^{-2/3})$. Therefore, to study the solution of (9.3.26)–(9.3.27) near $t = 1$. we set

$$t - 1 = \epsilon^{1/3}\tau, \qquad u(x,t) = \epsilon^{-2/3}v(x,\tau). \tag{9.3.37}$$

This gives

$$v_\tau(x,\tau) + \tau^2 v(x,\tau) + \epsilon^{1/3}v_x(x,\tau) = 1. \tag{9.3.38}$$

The stretching exponents for ϵ in (9.3.37) could have been determined by the use of an argument similar to that given in the preceding example, whereby a balancing of significant terms in the resulting equation (9.3.38) is carried out. A new feature for this problem is that an independent and the dependent variable were stretched.

This is usually required for nonlinear problems and often arises when dealing with inhomogeneous linear equations.

To specify the solution $v(x,\tau)$ of (9.3.38), we require that it match the outer solution (9.3.30) as we approach the region $O(\epsilon) < t < 1 - O(\epsilon^{1/3})$, where the (outer) perturbation series is valid. That is, we require that as $\tau \to -\infty$, $v(x,\tau)$ tends to the form (9.3.30) expressed as a function of τ.

The boundary layer expansion for $v(x,\tau)$ is given in powers of $\epsilon^{1/3}$, in view of the form of (9.3.38), and we set

$$v(x,\tau) = \sum_{n=0}^{\infty} v_n(x,\tau)\epsilon^{n/3} \tag{9.3.39}$$

and insert (9.3.39) into (9.3.38). The leading equation is

$$\frac{\partial v_0(x,\tau)}{\partial \tau} + \tau^2 v_0(x,\tau) = 1, \tag{9.3.40}$$

whose general solution can be written as

$$v_0(x,\tau) = \exp\left(-\frac{\tau^3}{3}\right)\int_{-\infty}^{\tau} \exp\left(\frac{\sigma^3}{3}\right)\, d\sigma + b\exp\left(-\frac{\tau^3}{3}\right), \tag{9.3.41}$$

with the arbitrary constant b as yet unspecified. We match (9.3.41) with (9.3.30) and integrate by parts to obtain

$$\int_{-\infty}^{\tau} \exp\left(\frac{\sigma^3}{3}\right)\, d\sigma = \frac{1}{\tau^2}\exp\left(\frac{\tau^3}{3}\right) + 2\int_{-\infty}^{\tau} \frac{1}{\sigma^3}\exp\left(\frac{\sigma^3}{3}\right)\, d\sigma, \tag{9.3.42}$$

with τ assumed to be negative. A further integration by parts yields

$$\int_{-\infty}^{\tau} \exp\left(\frac{\sigma^3}{3}\right) d\sigma = \left(\frac{1}{\tau^2} + \frac{2}{\tau^5}\right)\exp\left(\frac{\tau^3}{3}\right) + 10\int_{-\infty}^{\tau} \frac{1}{\sigma^6}\exp\left(\frac{\sigma^3}{3}\right) d\sigma. \tag{9.3.43}$$

Thus

$$v_0(x,\tau) = \frac{1}{\tau^2} + \frac{2}{\tau^5} + 10\exp\left(-\frac{\tau^3}{3}\right)\int_{-\infty}^{\tau} \frac{1}{\sigma^6}\exp\left(\frac{\sigma^3}{3}\right)\, d\sigma, \tag{9.3.44}$$

where we have put $b = 0$, since the solution would otherwise grow exponentially as $\tau \to -\infty$, and that is ruled out by the form of (9.3.30). If we now express (9.3.30) in terms of τ, we have $u(x, 1+\epsilon^{1/3}\tau) \approx \epsilon^{-2/3}\left(1/\tau^2 + 2/\tau^5\right)$. Since $t - 1 = \epsilon^{1/3}\tau$ and $u = \epsilon^{-2/3}v$—in view of (9.3.37)—we find that the *internal boundary layer* expression (9.3.44) matches the *outer solution* (9.3.30).

To complete our discussion of the solution we need to examine what happens in the region above the internal boundary layer [i.e., when $t > 1 + O(\epsilon^{1/3})$]. The outer solution (9.3.30) is expected to be valid there. However, for that to be the case, the

outer solution must match the internal boundary layer solution as $\tau \to \infty$ in $v_0(x,\tau)$. To show this we write $v_0(x,\tau)$ as

$$v_0(x,\tau) = \exp\left(-\frac{\tau^3}{3}\right)\left\{\int_{-\infty}^{a} \exp\left(\frac{\sigma^3}{3}\right)\,d\sigma + \int_{a}^{\tau} \exp\left(\frac{\sigma^3}{3}\right)\,d\sigma\right\} \quad (9.3.45)$$

where both a and τ are positive and $a < \tau$—note that we have put $b = 0$ in (9.3.41). Since $0 < a < \tau$, the first integral in (9.3.45) contributes an exponentially small term to the overall result. Integrating by parts in the second integral as was done previously readily yields an expression that matches the outer solution (9.3.30) as $\tau \to \infty$, since the contributions from the lower limit result in exponentially small terms.

We have shown that a satisfactory description of the solution of (9.3.24)–(9.3.25) can be obtained by combining *perturbation* and *boundary layer methods*. Although an exact solution is available for this problem, the approximate results are generally easier to evaluate. It is also possible to construct a *composite approximation* to the solution. We replace the terms $1 + 2\epsilon(\tau - 1)$ in (9.3.36) by (9.3.41), where we set $b = 0$ and $\tau = (t-1)\epsilon^{-1/3}$, and replace τ by t/ϵ in the remaining terms of (9.3.36). Since the outer solution (9.3.30) is singular at $t = 1$, it cannot be used in the uniformly valid composite result. Consequently, the form of the composite result is not much simpler than that of the exact solution (9.3.28).

Singular Perturbations and Boundary Layers for Hyperbolic PDEs

The following example deals with the singular perturbation of a second order linear hyperbolic equation with constant coefficients. The problem gives rise to an initial layer as well as secular terms.

Example 9.10. The Singular Perturbation of a Hyperbolic Equation. We consider the *second order linear hyperbolic equation*

$$\epsilon(u_{tt}(x,t) - c^2 u_{xx}(x,t)) + u_t(x,t) + a u_x(x,t) = 0, \quad (9.3.46)$$

with constant coefficients and $0 < \epsilon \ll 1$. The initial data on the x-axis are given as

$$u(x,0) = f(x), \qquad u_t(x,0) = g(x). \quad (9.3.47)$$

The *reduced equation* obtained by setting $\epsilon = 0$ in (9.3.46) is of first order, so that we have a *singular perturbation problem.*

We attempt to solve (9.3.46)–(9.3.47) by the use of a regular *perturbation series*

$$u(x,t) = \sum_{n=0}^{\infty} u_n(x,t)\epsilon^n. \quad (9.3.48)$$

Inserting (9.3.48) into (9.3.46) and equating like powers of ϵ yields the recursive system

$$\frac{\partial u_0(x,t)}{\partial t} + a\,\frac{\partial u_0(x,t)}{\partial x} = 0,$$

$$\frac{\partial u_n(x,t)}{\partial t} + a\,\frac{\partial u_n(x,t)}{\partial x} = -\frac{\partial^2 u_{n-1}(x,t)}{\partial t^2} + c^2\,\frac{\partial^2 u_{n-1}(x,t)}{\partial x^2}, \quad n \geq 1, \tag{9.3.49}$$

with the initial conditions

$$u_0(x,0) = f(x), \ \frac{\partial u_0(x,0)}{\partial t} = g(x), \quad u_n(x,0) = 0, \ \frac{\partial u_n(x,0)}{\partial t} = 0, \ n \geq 1. \tag{9.3.50}$$

Since the equations for $u_n(x,t)$ are each of first order, only one initial condition can be assigned for each $u_n(x,t)$ at $t = 0$. Therefore, the initial value problems for the $u_n(x,t)$ cannot be solved in general. The singular nature of the perturbation is thereby brought into evidence. As a result, we cannot expect the perturbation series to be valid near the x-axis. Nevertheless, the series (9.3.48) should provide an approximate description of the solution $u(x,t)$ away from the x-axis, as was found in the preceding examples.

Using the methods of Chapter 2, we easily obtain as the general solutions of the first three equations in (9.3.49) for the $u_n(x,t)$,

$$u_0 = F(x - at), \tag{9.3.51}$$

$$u_1 = t(c^2 - a^2)F''(x - at) + G(x - at), \tag{9.3.52}$$

$$u_2 = (c^2 - a^2)\left[\frac{t^2}{2}F''''(x - at) + 2atF'''(x - at) + tG''(x - at)\right] + H(x - at), \tag{9.3.53}$$

where $F(x)$, $G(x)$, and $H(x)$ are arbitrary functions. [$H(x)$ does not represent the Heaviside function here.]

It is clearly impossible to satisfy the initial conditions (9.3.50) for arbitrary $f(x)$ and $g(x)$. Additionally, if we assume that $F(x)$, $G(x)$, and $H(x)$ together with their derivatives are uniformly bounded for all x, we see from (9.3.51)–(9.3.53) that *secular terms* with coefficients ϵt and $(\epsilon t)^2$ arise in the perturbation expansion (9.3.48). Thus the series (9.3.48) is not expected to be valid for $\epsilon t = O(1)$ or when $t = O(1/\epsilon)$.

We have demonstrated that the perturbation series (9.3.48) is not only invalid near the initial line but that it also breaks down after a sufficiently long time. In addition, we observe from (9.3.51)–(9.3.53) that the terms $u_n(x,t)$ in the perturbation series are waves that travel to the right or left (depending on the sign of a) with speed $|a|$. Now we have already shown that the maximum speed at which disturbances for the hyperbolic equation (9.3.46) can travel is the characteristic speed. This speed equals c (we assume that $c > 0$,) and if $|a| > c$ it can happen that disturbances as described by our perturbation approximation can travel at a speed exceeding the characteristic speed. Since this is theoretically not possible, it would appear that we have to reject the perturbation series completely for this reason if $|a| > c$.

An application of the *stability analysis* of Section 3.5 shows that (9.3.46) is unstable if $|a| > c$. If we insert the normal mode solution (3.5.2) into (9.3.46), we obtain for

$\lambda(k)$, $\epsilon\lambda(k)^2 + \lambda(k) + \epsilon c^2 k + iak = 0$. Noting that $0 < \epsilon \ll 1$ and assuming moderate values of k yields two approximate solutions for $\lambda(k)$,

$$\lambda_1(k) \approx -\frac{1}{\epsilon}, \qquad \lambda_2(k) \approx -iak - \epsilon(c^2 - a^2)k^2. \tag{9.3.54}$$

Thus if $|a| > c$, the real part of $\lambda_2(k)$ is positive and the problem is *unstable*. Therefore, we assume that $|a| < c$ in (9.3.46) so that disturbances associated with the equations (9.3.49) travel at slower speeds than the characteristic speed c. It can be shown that the real parts of $\lambda_1(k)$ and $\lambda_2(k)$ are negative for all $k \neq 0$ if $|a| < c$, so that (9.3.46) is an equation of *dissipative type*.

To specify the unknown functions in the perturbation series, we must relate it to the initial data, and this is done by the use of *boundary layer theory*. The perturbation series is expected to be valid from some time $t > 0$ on, so we assume that there exists an *initial layer* near the x-axis where the solution of (9.3.46)–(9.3.47) undergoes a rapid transition from a form that satisfies the initial data to the perturbation series form.

To determine the appropriate form of (9.3.46) in the initial layer, we introduce the *stretching transformation*

$$t = \epsilon^r \tau, \tag{9.3.55}$$

with $r > 0$ to be specified. We insert (9.3.55) into (9.3.46) to obtain

$$\epsilon^{1-2r}\hat{u}_{\tau\tau}(x,\tau) - \epsilon c^2 \hat{u}_{xx}(x,\tau) + \epsilon^{-r}\hat{u}_\tau(x,\tau) + a\hat{u}_x(x,\tau) = 0, \tag{9.3.56}$$

where we have set $\hat{u}(x,\tau) = u(x,\epsilon^r\tau)$. The initial data for $\hat{u}_x(x,\tau)$ at $\tau = 0$ are $\hat{u}(x,0) = f(x)$, $\hat{u}_\tau(x,0) = \epsilon^r g(x)$.

We intend to solve (9.3.56) by a perturbation method. To specify the constant r in (9.3.55), we argue that the most significant terms in (9.3.56) are those with the lowest power (possibly negative) in ϵ. These terms determine the basic form of the equations (in either homogeneous or inhomogeneous form) that must be satisfied by the terms in the perturbation series. Since each term must satisfy two initial conditions at $\tau = 0$, we require that $\hat{u}_{\tau\tau}(x,\tau)$ be retained as a leading term in the expansion in ϵ. Clearly, this requires that $1 - 2r \leq -r$ or $r \geq 1$, as is seen on comparing terms in (9.3.56). If we choose $r > 1$, $\hat{u}_{\tau\tau}(x,\tau)$ is the only leading term. However, with $r = 1$, $\hat{u}_{\tau\tau}(x,\tau)$ and $\hat{u}_\tau(x,\tau)$ are both of the same order in ϵ, and this choice also yields a balance between the term $\hat{u}_{\tau\tau}(x,\tau)$ that is significant in the boundary layer region, and the term $\hat{u}_\tau(x,\tau)$ that occurs (as $u_t(x,t)$) in the reduced problem in the outer region where the perturbation series (9.3.48) is valid. Therefore, we set $r = 1$. Multiplying through by ϵ in (9.3.56) yields

$$\hat{u}_{\tau\tau}(x,\tau) + \hat{u}_\tau(x,\tau) + \epsilon a\hat{u}_x(x,\tau) - \epsilon^2 c^2 \hat{u}_{xx}(x,\tau) = 0, \tag{9.3.57}$$

with the initial data

$$\hat{u}(x,0) = f(x), \qquad \hat{u}_\tau(x,0) = \epsilon g(x). \tag{9.3.58}$$

To solve (9.3.57)–(9.3.58), we introduce the *boundary layer expansion*

$$\hat{u}(x,\tau) = \sum_{n=0}^{\infty} \hat{u}_n(x,\tau)\epsilon^n \tag{9.3.59}$$

into (9.3.57)–(9.3.58) and obtain the recursive system of equations

$$\frac{\partial^2 \hat{u}_0(x,\tau)}{\partial \tau^2} + \frac{\partial \hat{u}_0(x,\tau)}{\partial \tau} = 0, \tag{9.3.60}$$

$$\frac{\partial^2 \hat{u}_1(x,\tau)}{\partial \tau^2} + \frac{\partial \hat{u}_1(x,t)}{\partial \tau} = -a\,\frac{\partial \hat{u}_0(x,\tau)}{\partial x}, \tag{9.3.61}$$

$$\frac{\partial^2 \hat{u}_n(x,\tau)}{\partial \tau^2} + \frac{\partial \hat{u}_n(x,t)}{\partial \tau} = -a\,\frac{\partial \hat{u}_{n-1}(x,\tau)}{\partial x} + c^2\,\frac{\partial^2 \hat{u}_{n-2}(x,\tau)}{\partial x^2}, \quad n \geq 2, \tag{9.3.62}$$

on equating like powers of ϵ. The initial data are

$$\hat{u}_0(x,0) = f(x), \qquad \hat{u}_n(x,0) = 0, \quad n \geq 1,$$

$$\frac{\partial \hat{u}_0(x,0)}{\partial \tau} = 0, \qquad \frac{\partial \hat{u}_1(x,0)}{\partial \tau} = g(x), \quad \frac{\partial \hat{u}_n(x,0)}{\partial \tau} = 0, \quad n \geq 2. \tag{9.3.63}$$

Each of the boundary layer equations is an ODE in τ with initial data at $\tau = 0$ and is easily solved. We find that for $n = 0$ and $n = 1$,

$$\hat{u}_0(x,\tau) = f(x),\ \hat{u}_1(x,\tau) = (g(x) + af'(x))(1 - e^{-\tau}) - a\tau f'(x). \tag{9.3.64}$$

The exponential $e^{-\tau} = e^{-t/\epsilon}$ in $\hat{u}_1(x,\tau)$ decays rapidly as τ or t increases and thus plays no role in the outer region where $t > O(\epsilon)$ and the perturbation expansion (9.3.48) is valid. To specify the unknown functions in (9.3.48), we need to match (9.3.59) and (9.3.48), each of which are different representations of the unique solution $u(x,t)$. Although (9.3.59) is assumed to be valid in the boundary layer region near the x-axis whose width is of order ϵ, and the perturbation series is expected to be valid in a region where $t > O(\epsilon)$, away from the x-axis, it is assumed that they have a common region of validity.

The perturbation and boundary layer expansions for $u(x,t)$ yield

$$u(x,t) \approx \sum_{n=0}^{\infty} u_n(x,\epsilon\tau)\epsilon^n \approx \sum_{n=0}^{\infty} \hat{u}_n(x,\tau)\epsilon^n, \qquad t = \epsilon\tau. \tag{9.3.65}$$

We assume that $\epsilon\tau$ is so small that the terms $u_n(x,\epsilon\tau)$ can be well approximated by the leading terms in their series expansions in $t = \epsilon\tau$. However, τ must be large enough that the exponentials $e^{-\tau}$ in the $\hat{u}_n(x,\tau)$ terms can be neglected. This is so if $t = \epsilon\tau = O(\epsilon^{1/2})$. Using (9.3.51)–(9.3.52) gives

$$u_0(x,\epsilon\tau) + \epsilon u_1(x,\epsilon\tau) = F(x) - \epsilon a\tau F'(x) + \epsilon G(x) + O(\epsilon^2), \tag{9.3.66}$$

and neglecting the $e^{-\tau}$ terms in (9.3.64) gives

$$\hat{u}_0(x,\tau) + \epsilon\hat{u}_1(x,\tau) = f(x) - \epsilon a\tau f'(x) + \epsilon g(x) + \epsilon a f'(x) + O(\epsilon^2). \quad (9.3.67)$$

Comparing like powers of ϵ in (9.3.66) yields

$$F(x) = f(x), \qquad G(x) = g(x) + af'(x), \quad (9.3.68)$$

on using (9.3.66)–(9.3.67). Consequently, (9.3.48) has the form

$$u = f(x-at)+\epsilon[t(c^2-a^2)f''(x-at)+g(x-at)+af'(x-at)]+O(\epsilon^2). \quad (9.3.69)$$

Further terms in the perturbation series can be determined by carrying this matching procedure out to higher orders.

A *composite expression* for the solution can be given as

$$u(x,t) \approx f(x-at) + \epsilon[t(c^2-a^2)f''(x-at) + g(x-at) + af'(x-at)] - \epsilon[g(x-at) + af'(x-at)]e^{-t/\epsilon}. \quad (9.3.70)$$

In the initial layer where $t = \epsilon\tau$, (9.3.70) can be expressed as (9.3.64). The initial conditions $u(x,0) = f(x)$ and $u_t(x,0) = g(x)$ are satisfied to $O(\epsilon^2)$ and $O(\epsilon)$, respectively.

The perturbation series (9.3.69) breaks down when $t = O(1/\epsilon)$. A leading-order approximation to the solution valid for $t \geq O(1/\epsilon)$ may be obtained as follows. On inserting (9.3.48) into (9.3.46), we obtain

$$\frac{\partial u_0}{\partial t} + a\frac{\partial u_0}{\partial x} + \epsilon\left(\frac{\partial u_1}{\partial t} + a\frac{\partial u_1}{\partial x} + \frac{\partial^2 u_0}{\partial t^2} - c^2\frac{\partial^2 u_0}{\partial x^2}\right) = O(\epsilon^2). \quad (9.3.71)$$

From (9.3.71) we see that

$$\frac{\partial u_0(x,t)}{\partial t} = -a\,\frac{\partial u_0(x,t)}{\partial x} + O(\epsilon), \quad \frac{\partial^2 u_0(x,t)}{\partial t^2} = a^2\,\frac{\partial^2 u_0(x,t)}{\partial x^2} + O(\epsilon) \quad (9.3.72)$$

if we assume that the solutions $u_n(x,t)$ are smooth. Inserting (9.3.72) into (9.3.71) yields, to the same level of approximation,

$$\frac{\partial u_0}{\partial t} + a\,\frac{\partial u_0}{\partial x} + \epsilon\left[\frac{\partial u_1}{\partial t} + a\,\frac{\partial u_1}{\partial x} + (a^2 - c^2)\,\frac{\partial^2 u_0}{\partial x^2}\right] = O(\epsilon^2). \quad (9.3.73)$$

Collecting like powers of ϵ yields equations for $u_0(x,t)$ and $u_1(x,t)$ whose solutions are (9.3.51)–(9.3.52). Clearly, the term $(a^2-c^2)\partial^2 u_0(x,t)/\partial x^2$ is the one that gives rise to the secular term in (9.3.52). We may avoid this secularity by regrouping the terms in (9.3.73) and writing the equation as

$$\left[\frac{\partial u_0}{\partial t} + a\frac{\partial u_0}{\partial x} + \epsilon(a^2 - c^2)\frac{\partial^2 u_0}{\partial x^2}\right] + \epsilon\left(\frac{\partial u_1}{\partial t} + a\frac{\partial u_1}{\partial x}\right) = O(\epsilon^2). \quad (9.3.74)$$

The bracketed term is equated to zero so that $u_0(x,t)$ satisfies a *diffusion equation.* The initial value for $u_0(x,t)$ may be taken to be (9.3.69) evaluated at some time $t = t_0 \leq O(1/\epsilon)$. We remark that results given in Sections 1.2 and 5.7 imply that the solution of (9.3.46) should approximately satisfy a diffusion equation for large t. The present approach is related to the *parabolic equation method* discussed below.

Singular Perturbations and Boundary Layers for Linear Elliptic PDEs: A Specific Example

Next we turn to a consideration of singular perturbation problems for *linear second order elliptic equations* of the form

$$\epsilon(Au_{xx} + 2Bu_{xy} + Cu_{yy} + Du_x + Eu_y + Fu) + au_x + bu_y + cu = g \quad (9.3.75)$$

in a region G, where $0 < \epsilon \ll 1$ and $u = u(x,y)$ is prescribed on the boundary ∂G of G. (That is, we are considering a *Dirichlet problem.*) The coefficients $A, B, \ldots, F, a, b, c$, and g are specified functions of (x,y) and $B^2(x,y) - A(x,y)C(x,y) > 0$ in G so that (9.3.75) is of elliptic type. We assume that $a^2(x,y) + b^2(x,y) > 0$, so that the *reduced equation* for (9.3.75),

$$a(x,y)u_x(x,y) + b(x,y)u_y(x,y) + c(x,y)u(x,y) = g(x,y), \qquad (9.3.76)$$

obtained on setting $\epsilon = 0$ in (9.3.75), is a first order PDE. As a result, we are dealing with a *singular perturbation problem.*

A number of difficulties can arise when trying to solve the Dirichlet problem for (9.3.75) by using a regular perturbation series $u(x,y) = \sum_{n=0}^{\infty} u_n(x,y)\epsilon^n$. A major problem can result from the fact that each term in the series satisfies a first order equation of the general form (9.3.76). Consequently, the data for the given equation (9.3.75) may *overdetermine* the solutions of the equations satisfied by the terms in the perturbation series. Also, *singularities* in the data for (9.3.75) must be carried along the characteristics of the reduced equation (9.3.76), and therefore, they occur in the terms of the regular perturbation series. The elliptic equation (9.3.75) has no real characteristics, and its solutions must therefore be smooth functions. Further, it may happen that *characteristic initial value problems* can occur for the terms in the perturbation series or that these terms may become singular in the interior of the given region for the boundary value problem for (9.3.75). These and other difficulties can be remedied by the introduction of appropriate boundary layers near portions of the boundary or in the interior of the given region.

The construction of a perturbation solution for Dirichlet or other boundary value problems for (9.3.75) over an arbitrary region G is not a simple matter in general. To appreciate some of the problems that can arise in solving (9.3.75), we begin with a specific (simple) example.

Example 9.11. A Singular Perturbation Problem for an Elliptic Equation. We consider the *linear elliptic equation*

$$\epsilon(u_{xx}(x,y) + u_{yy}(x,y)) + u_x(x,y) + bu_y(x,y) = 0 \tag{9.3.77}$$

over the semi-infinite region $0 < x < \infty$ and $0 < y < L$, with $0 < \epsilon \ll 1$ and b taken to be a constant. The boundary conditions are

$$u(x,0) = f(x), \qquad u(0,y) = g(y), \qquad u(x,L) = h(x). \tag{9.3.78}$$

We assume that $f(x)$ and $h(x)$ vanish as $x \to \infty$ and require that $u(x,y) \to 0$ as $x \to \infty$ for $0 \le y \le L$.

As a first step in solving (9.3.77)–(9.3.78) by the *perturbation method*, we expand $u(x,y)$ as

$$u(x,y) = \sum_{n=0}^{\infty} u_n(x,y)\epsilon^n \tag{9.3.79}$$

and insert (9.3.79) into (9.3.77)–(9.3.78). On equating like powers of ϵ, we obtain for $u_0(x,y)$,

$$\frac{\partial u_0(x,y)}{\partial x} + b\,\frac{\partial u_0(x,y)}{\partial y} = 0, \tag{9.3.80}$$

and for the $u_n(x,y)$, with $n \ge 1$,

$$\frac{\partial u_n(x,y)}{\partial x} + b\,\frac{\partial u_n(x,y)}{\partial y} = -\left(\frac{\partial^2 u_{n-1}(x,y)}{\partial x^2} + \frac{\partial^2 u_{n-1}(x,y)}{\partial y^2}\right). \tag{9.3.81}$$

The boundary conditions for $u_0(x,y)$ are identical to those for $u(x,y)$ [i.e., (9.3.78)], while $u_n(x,y) = 0$ on all three boundary lines if $n \ge 1$.

The general solution of (9.3.80) is $u_0(x,y) = F(y - bx)$, where $F(z)$ is an arbitrary function. Also, $u_1(x,y) = F_1(y - bx) - [(1 + b^2)/b]yF''(y - bx)$, where $F_1(z)$ is arbitrary. Alternatively, the solution can be written as $u_1(x,y) = G_1(y - bx) - (1 + b^2)xF''(y - bx)$, where $G_1(z)$ is arbitrary.

The lines $y - bx$ = constant are the characteristic lines for the equations for the $u_n(x,y)$. The solution $u_0(x,y) = F(y - bx)$ is constant on the characteristic lines. If $b \neq 0$, the characteristics intersect two of the boundary lines within the given region for the problem. If $b = 0$, the boundaries $y = 0$ and $y = L$ coincide with characteristic lines. In either case, $u_0(x,y)$ cannot satisfy the full set of boundary conditions since the boundary values $f(x)$, $g(y)$, and $h(x)$ are assigned arbitrarily.

If $b > 0$, for example, we may set

$$u_0(x,y) = g(y - bx),\ x < \frac{y}{b}, \qquad u_0(x,y) = f\left(x - \frac{y}{b}\right),\ x > \frac{y}{b}, \tag{9.3.82}$$

and satisfy the conditions at $x = 0$ and $y = 0$, or we may set

$$u_0(x,y) = h\left(x - \frac{y - L}{b}\right) \tag{9.3.83}$$

and satisfy the condition at $y = L$. We cannot determine from the perturbation series (9.3.79), and the resulting equations for its terms, which of these boundary conditions should be assigned. In any case, since the full set of boundary conditions for $u_0(x, y)$ and the $u_n(x, y)$ cannot be applied, we must introduce *boundary layers* to deal with the lost boundary conditions.

Each of the boundary lines is a candidate for a boundary layer, so we introduce stretching transformations that emphasize neighborhoods of each of the boundary lines. If $b \neq 0$, we set [in (9.3.77)]

$$y = \epsilon\eta, \qquad y - L = \epsilon\eta \tag{9.3.84}$$

near $y = 0$ and $y = L$, respectively. (The stretching exponent for ϵ is found to equal 1 by balancing the second derivative term with terms arising from the reduced equation, in the boundary layer equation.) With $v = v(x, \eta)$ we then have the *boundary layer equation*

$$v_{\eta\eta}(x, \eta) + bv_\eta(x, \eta) + \epsilon v_x(x, \eta) + \epsilon^2 v_{xx}(x, \eta) = 0, \tag{9.3.85}$$

which replaces (9.3.77). With $v(x, \eta)$ expanded in powers of ϵ and $v_0(x, \eta)$ as the leading term, we find that $\partial^2 v_0(x, \eta)/\partial\eta^2 + b\,\partial v_0(x, \eta)/\partial\eta = 0$, with the general solution

$$v_0(x, \eta) = \alpha(x) + \beta(x)e^{-b\eta}, \tag{9.3.86}$$

where $\alpha(x)$ and $\beta(x)$ are arbitrary functions. Now, if $b > 0$, the exponential term in (9.3.86) decreases as η increases. Therefore, the boundary layer for this case must lie near $y = 0$. If $b < 0$, the exponential decreases as η decreases and—in view of (9.3.84)—the boundary layer should be located near $y = L$. In either case the boundary layer width is $O(\epsilon)$.

However, if $b = 0$ and we retain the stretching transformations (9.3.84), we find that $v_0(x, \eta)$ is given as

$$v_0(x, \eta) = \alpha(x) + \beta(x)\eta, \tag{9.3.87}$$

so that no boundary layer effect occurs at either $y = 0$ or $y = L$, since the solution does not undergo exponential decay. In this case we easily conclude that the appropriate stretchings are

$$y = \epsilon^{1/2}\eta, \qquad y - L = \epsilon^{1/2}\eta \tag{9.3.88}$$

and the *boundary layer equation* becomes

$$v_{\eta\eta}(x, \eta) + v_x(x, \eta) + \epsilon v_{xx}(x, \eta) = 0. \tag{9.3.89}$$

Then $v_0(x, \eta)$ satisfies the *parabolic equation*

$$\frac{\partial^2 v_0(x, \eta)}{\partial\eta^2} + \frac{\partial v_0(x, \eta)}{\partial x} = 0. \tag{9.3.90}$$

The boundary layer width is $O(\epsilon^{1/2})$, and there are boundary layers near $y = 0$ and $y = L$, as will be shown below.

Near $x = 0$, we set $x = \epsilon\xi$ in (9.3.77) and obtain the *boundary layer equation*

$$w_{\xi\xi}(\xi, y) + w_\xi(\xi, y) + \epsilon b w_y(\xi, y) + \epsilon^2 w_{yy}(\xi, y) = 0. \tag{9.3.91}$$

If $w(\xi, y)$ is expanded in powers of ϵ, we easily find that the leading term $w_0(\xi, y)$ is given as

$$w_0(\xi, y) = \gamma(y) + \rho(y)e^{-\xi}, \tag{9.3.92}$$

with arbitrary $\gamma(y)$ and $\rho(y)$, so that a boundary layer can be located at $x = 0$ because of the exponentially decaying term in (9.3.92).

For the case $b \neq 0$, we now assume that $b > 0$ and complete the perturbation solution of the problem. Since we have shown that there can be boundary layers near $x = 0$ and $y = 0$ if $b > 0$, we specify $u_0(x, y) = F(y - bx)$ to satisfy the boundary condition at $y = L$. This gives the *outer solution*

$$u_0(x, y) = h\left(x - \frac{y - L}{b}\right). \tag{9.3.93}$$

The *boundary layer solution* $v_0(x, \eta)$ [i.e., (9.3.86)] near $y = 0$ is required to satisfy the *boundary condition* and the *matching condition*

$$v_0(x, 0) = f(x), \qquad \lim_{\eta\to\infty} v_0(x, \eta) = \lim_{y\to 0} u_0(x, y), \tag{9.3.94}$$

respectively. The matching condition states that for large η the boundary layer solution $v_0(x, \eta)$ must agree with the outer solution $u_0(x, y)$ evaluated for small y. We readily find that $v_0(x, \eta)$ is given as

$$v_0(x, \eta) = h\left(x + \frac{L}{b}\right) + \left[f(x) - h\left(x + \frac{L}{b}\right)\right] e^{-b\eta}. \tag{9.3.95}$$

For the *boundary layer* near $x = 0$, we have the *boundary and matching conditions*

$$w_0(0, y) = g(y), \quad \lim_{\xi\to\infty} w_0(\xi, y) = \lim_{x\to 0} u_0(x, y), \tag{9.3.96}$$

with $w_0(\xi, y)$ defined in (9.3.92) and $u_0(x, y)$ given in (9.3.93). Then

$$w_0(\xi, y) = h\left(\frac{L - y}{b}\right) + \left[g(y) - h\left(\frac{L - y}{b}\right)\right] e^{-\xi}. \tag{9.3.97}$$

The matching procedure given in (9.3.94) and (9.3.96) is effectively equivalent to that carried out in the preceding examples. As we are only matching the leading terms in each case, we carry out the matching by the use of a limit process. In the hyperbolic problem of Example 9.10 where two terms in the perturbation and boundary layer expansion were used, the matching was carried out by expressing both series in terms of a common variable. This approach could have been used here as well.

Inspection of the outer and boundary layer solutions shows that the solution $u = u(x, y)$ of (9.3.77)–(9.3.78) with $b > 0$ can be expressed in *composite form* as

$$u \approx h\left(\frac{bx - y + L}{b}\right) + \left[f(x) - h\left(\frac{bx + L}{b}\right)\right] e^{-by/\epsilon} + \left[g(y) - h\left(\frac{L - y}{b}\right)\right] e^{-x/\epsilon}. \tag{9.3.98}$$

The terms containing the exponentials are significant only within their respective boundary layers. Apart from exponentially small terms, $u(x, y)$ satisfies the boundary conditions at $y = 0$, $x = 0$, and $y = L$ and vanishes as $x \to \infty$ since $f(x)$ and $h(x)$ vanish there. Effects due to possibly incompatible boundary values at the points $(x, y) = (0, 0)$ and $(x, y) = (0, L)$ are not considered.

If $b < 0$ in (9.3.77), the perturbation solution has a composite expression similar to that in (9.3.98) except that the boundary layer is shifted to $y = L$. This case is considered in Example 9.12.

If $b = 0$ in (9.3.77), the leading term in the (outer) perturbation expansion is $u_0(x, y) = F(y)$. Since $F(y)$ is constant on $y = 0$ and $y = L$, it cannot satisfy the boundary conditions there. We could choose $F(y) = g(y)$ and thereby satisfy the boundary condition at $x = 0$. However, $u(x, y)$ must vanish as $x \to \infty$, and the solution $u_0(x, y) = g(y)$ does not vary with x and cannot satisfy the condition at infinity unless $g(y) = 0$. Since $g(y)$ is specified arbitrarily, we must put $F(y) = 0$ and obtain as the *outer solution*

$$u(x, y) \approx u_0(x, y) = 0, \tag{9.3.99}$$

for otherwise $u(x, y)$ would not vanish as $x \to \infty$. Consequently, we must use not only the boundary layers at $y = 0$ and $y = L$ but also the boundary layer at $x = 0$, in this case.

The boundary layer at $x = 0$ is the easiest to consider. Using (9.3.96)–(9.3.97), we conclude that

$$w_0(\xi, y) = g(y)e^{-\xi} \tag{9.3.100}$$

since $u_0(x, y) = 0$. Near $y = 0$ and $y = L$ we have the *parabolic boundary layer equation* for $v_0(x, \eta)$,

$$\frac{\partial^2 v_0(x, \eta)}{\partial \eta^2} + \frac{\partial v_0(x, \eta)}{\partial x} = 0, \tag{9.3.101}$$

as given in (9.3.90). In the boundary layer near $y = 0$, where $y = \epsilon^{1/2}\eta$, the boundary condition is

$$v_0(x, 0) = f(x), \tag{9.3.102}$$

while for the boundary layer near $y = L$, where $y - L = \epsilon^{1/2}\eta$, the boundary condition is

$$v_0(x, 0) = h(x). \tag{9.3.103}$$

As $x \to \infty$, we require that $v_0(x, \eta) \to 0$. Also, as $\eta \to \infty$ in the boundary layer near $y = 0$ and $\eta \to -\infty$ in the boundary layer near $y = L$, we must have $v_0(x, \eta) \to 0$, because $v_0(x, \eta)$ must match the outer solution $u_0(x, y)$, which was shown to vanish.

With x thought of as the time variable, (9.3.101) is a *backward heat equation.* Thus we cannot assign values for $v_0(x,\eta)$ at $x = 0$ and solve (9.3.101) for $x > 0$, for such a problem would not be well posed. Instead, we must place conditions on $v_0(x,\eta)$ at some $x = x_0 > 0$ and solve for $x < x_0$. The only conditions given are at infinity, as we have shown. We may think of the problem as being given with data at $x = x_0$ and then let $x_0 \to -\infty$. The solution is then given for all $x < \infty$. The result is equivalent to a *steady-state problem* for the heat equation where the effects of the initial temperature have died out. We can think of that problem as being given for all $t > -\infty$. The solution to the steady-state problem for the heat equation and to our problem for (9.3.101) is unique if we assume that the solution is uniformly bounded.

By adapting the results of Example 5.5, we conclude that the solution of (9.3.101)–(9.3.102) with $v_0(-\infty,\eta) = 0$ is

$$v_0(x,\eta) = \frac{\eta}{\sqrt{4\pi}} \int_x^\infty \exp\left[-\frac{\eta^2}{4(\sigma - x)}\right] \frac{f(\sigma)}{(\sigma - x)^{3/2}}\, d\sigma. \tag{9.3.104}$$

For the problem (9.3.101) and (9.3.103) the solution is

$$v_0(x,\eta) = \frac{\eta}{\sqrt{4\pi}} \int_x^\infty \exp\left[-\frac{\eta^2}{4(\sigma - x)}\right] \frac{h(\sigma)}{(\sigma - x)^{3/2}}\, d\sigma. \tag{9.3.105}$$

As $|\eta| \to \infty$ both solutions decay exponentially and vanish as $x \to \infty$.

A *composite expression* for the solution may be given as

$$u(x,y) \approx g(y) \exp\left(-\frac{x}{\epsilon}\right) + \frac{y}{\sqrt{4\epsilon\pi}} \int_x^\infty \exp\left[-\frac{y^2}{4\epsilon(\sigma - x)}\right] \frac{f(\sigma)}{(\sigma - x)^{3/2}}\, d\sigma$$

$$+ \frac{y - L}{\sqrt{4\epsilon\pi}} \int_x^\infty \exp\left[-\frac{(y - L)^2}{4\epsilon(\sigma - x)}\right] \frac{h(\sigma)}{(\sigma - x)^{3/2}}\, d\sigma. \tag{9.3.106}$$

We do not consider the behavior of $u(x,y)$ at $(0,0)$ and $(0,L)$. The solution is exponentially small when $x > O(\epsilon)$, $y > O(\sqrt{\epsilon})$, and $|y - L| > O(\sqrt{\epsilon})$.

We remark that near $x = 0$ and in the case where $b \neq 0$, we were led to consider ODEs in the boundary layer regions, as had been the case in the problems considered previously. However, with $b = 0$ the boundary layer equations valid near $y = 0$ and $y = L$ turned out to be parabolic PDEs. This can be attributed to the fact that these boundary lines coincided with the characteristics of the reduced equation for (9.3.77) with $b = 0$. The occurrence of parabolic boundary layer equations when portions of the boundary coincide with the characteristics of the reduced equation is generally the case, as shown below.

Singular Perturbations and Boundary Layers for Elliptic PDEs: A General Discussion

To study the singular perturbation problem for the *linear elliptic equation* (9.3.75) in a general case, it is convenient to transform the coordinates so as to simplify the form of the *reduced equation*

$$a(x,y)u_x(x,y) + b(x,y)u_y(x,y) + c(x,y)u(x,y) = g(x,y). \tag{9.3.107}$$

We assume that the characteristic curves for (9.3.107) do not intersect in the region given for (9.3.75). Then the characteristics and their orthogonal trajectories may be introduced as a new set of coordinates in (9.3.75). The transformed equation can be written as

$$\epsilon(Au_{xx}+2Bu_{xy}+Cu_{yy}+Du_x+Eu_y+Fu)-u_x(x,y)+c(x,y)u(x,y) = g(x,y), \tag{9.3.108}$$

where we have retained the original notation for the coefficients and the variables x, y and $u(x,y)$ for the sake of convenience. [The coefficients $A, B, \ldots, F$ in (9.3.108) can be functions of (x,y)]. The type of an equation is invariant under nonsingular transformations so that (9.3.108) is again of *elliptic type*.

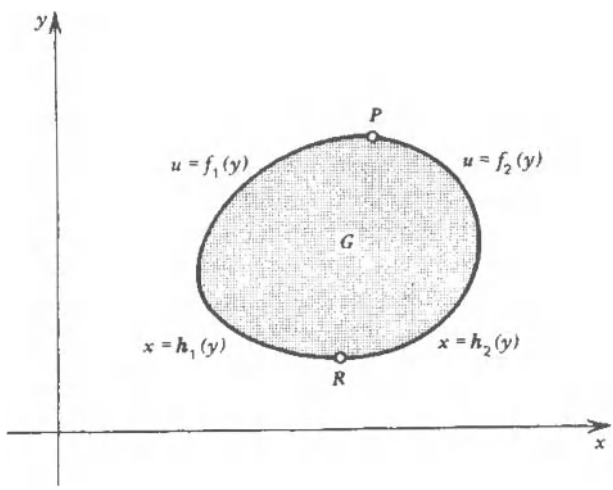

Figure 9.1 The convex region G.

We now consider the *Dirichlet problem* for (9.3.108) in a bounded region G, and note that our discussion applies to the transformed region obtained in going from (9.3.75) to (9.3.108). To begin, we assume that G is a *convex region* as shown in Figure 9.1. The points P and R in the diagram are points where the tangent line to the (smooth) boundary curve is horizontal. To the left of those points the curve is given as $x = h_1(y)$, with $u(h_1(y), y) = f_1(y)$ as the boundary condition. To right of P and

R we have $x = h_2(y)$ and $u(h_2(y), y) = f_2(y)$. The characteristics for the reduced equation for (9.3.108) are the lines y = constant. Thus the characteristics that pass through the interior of G intersect the boundary curve twice. The two characteristics that are tangent to ∂G intersect the boundary curve ∂G at the points P and R.

The *conventional perturbation series* $u(x, y) = \sum_{n=0}^{\infty} u_n(x, y)\epsilon^n$ when inserted into (9.3.108) yields for $u_0(x, y)$,

$$-\frac{\partial u_0(x, y)}{\partial x} + c(x, y)\, u_0(x, y) = g(x, y). \tag{9.3.109}$$

The $u_n(x, y)$, $n \geq 1$, satisfy equations of the same form. Now $u_0(x, y)$ must satisfy the conditions $u_0(h_1(y), y) = f_1(y)$, $u_0(h_2(y), y) = f_2(y)$, and this is generally impossible, since $u_0(x, y)$ is specified uniquely either by its values on $x = h_1(y)$ or on $x = h_2(y)$. Consequently, a *boundary layer* must be placed near $x = h_1(y)$ or $x = h_2(y)$, and its location is not known in advance.

To specify the boundary layer location, we stretch the neighborhoods of the boundary curves $x = h_1(y)$ and $x = h_2(y)$. The *stretching transformations* are

$$x - h_i(y) = \epsilon\xi, \qquad i = 1, 2, \qquad y = \eta. \tag{9.3.110}$$

The boundary layer width is easily shown to be $O(\epsilon)$, and we replace the variables (x, y) by (ξ, η) and $u(x, y)$ by $v(\xi, \eta)$. The notation for the coefficients in (9.3.108) is retained, so that $A(x, y)$ is written as $A(\xi, \eta)$, for example. Then the *boundary layer equation* obtained from (9.3.108) takes the form

$$(A(0, \eta) - 2h_i'(\eta)B(0, \eta) + h_i'^2(\eta)C(0, \eta))\, v_{\xi\xi}(\xi, \eta) - v_\xi(\xi, \eta) = O(\epsilon), \; i = 1, 2. \tag{9.3.111}$$

Since (9.3.108) is of elliptic type, both A and C must either be positive or negative, and the quadratic form

$$Q(\lambda) \equiv A - 2\lambda B + \lambda^2 C \tag{9.3.112}$$

must be nonzero. As a result, the coefficient of $v_{\xi\xi}(\xi, \eta)$ in (9.3.112), where $\lambda = h_i'(\eta)$, is either strictly positive or strictly negative. As an example, we assume that $A = C = \alpha$, with α as a constant, and that $C = D = E = F = 0$ in (9.3.108). Then $Q(\lambda) = \alpha(1 + \lambda^2)$. The sign of α determines that of $Q(\lambda)$. If $\alpha > 0$, so is $Q(\lambda)$, and if $\alpha < 0$ so is $Q(\lambda)$.

To solve the boundary layer equation, we expand $v(\xi, \eta)$ as

$$v(\xi, \eta) = \sum_{n=0}^{\infty} v_n(\xi, \eta)\epsilon^n, \tag{9.3.113}$$

and insert (9.3.113) into (9.3.111). Then $v_0(\xi, \eta)$ satisfies the equation

$$Q(h_i'(\eta))\, \frac{\partial^2 v_0(\xi, \eta)}{\partial \xi^2} - \frac{\partial v_0(\xi, \eta)}{\partial \xi} = 0. \tag{9.3.114}$$

The general solution of (9.3.114) is

$$v_0(\xi,\eta) = \alpha(\eta) + \beta(\eta)\exp\left(\frac{\xi}{Q}\right). \tag{9.3.115}$$

Now ξ increases and decreases with x in view of (9.3.110). Thus, if $Q > 0$, the boundary layer must be placed near $x = h_2(y)$, while if $Q < 0$, the boundary layer must be near $x = h_1(y)$. With these choices, the exponential in (9.3.115) decreases on moving away from the boundary into the region G. Since $Q(\lambda)$ is either greater than or less than zero for all λ, we conclude from (9.3.113) that $Q > 0$ if A and C are positive in and on the boundary of G and that $Q < 0$ if A and C are negative there. Thus, the boundary layer must be placed at $x = h_2(y)$ if A and C are positive and at $x = h_1(y)$ if A and C are negative.

Assuming that A and C are positive, we have determined that a boundary layer (if it is necessary) must be located near the boundary curve $x = h_2(y)$. Thus the (outer) equation (9.3.109) must be solved subject to the boundary condition given at $x = h_1(y)$ [i.e., $u_0(h_1(y), y) = f_1(y)$]. The solution is easily found to be

$$u_0(x,y) = f_1(y)\exp\left[\int_{h_1(y)}^{x} c(s,y)\,ds\right] - \int_{h_1(y)}^{x} g(\sigma,y)\exp\left[\int_{\sigma}^{x} c(s,y)\,ds\right]d\sigma. \tag{9.3.116}$$

For the boundary layer function $v_0(\xi,\eta)$, we require that $v_0(0,\eta) = f_2(\eta)$. Additionally, the matching condition gives $\lim_{\xi\to-\infty} v_0(\xi,\eta) = \lim_{x\to h_2(\eta)} u_0(x,\eta)$. Using (9.3.115) yields

$$v_0(\xi,\eta) = u_0(h_2(\eta),\eta) + [f_2(\eta) - u_0(h_2(\eta),\eta)]\exp\left[\frac{\xi}{Q(h_2'(\eta))}\right]. \tag{9.3.117}$$

Combining (9.3.116) and (9.3.117), we obtain a *composite approximation* for $u(x,y)$ as

$$u(x,y) \approx u_0(x,y) + [f_2(y) - u_0(h_2(y),y)]\exp\left[\frac{x - h_2(y)}{\epsilon Q(h_2'(y))}\right]. \tag{9.3.118}$$

Since $h_2'(y)$ blows up at the points P and R on the boundary curve (where the tangent lines are horizontal), these results are not valid near P and R. (The behavior of the solution near those points is considered below.) We obtain a similar result for the solution if A and C are negative.

A different approach is needed if the region G has a portion of its boundary that coincides with a line y = constant. A situation of this type is pictured in Figure 9.2. The point P separates the curves $x = h_1(y)$ and $x = h_2(y)$. The transition from these curves to the line $y = y_0$ at the points R and T is assumed to be smooth. In the region $y > y_0$ within G and away from the point P, the result (9.3.118) again yields an approximate solution if A and C are positive. The line $y = y_0$ is a *characteristic* for the *reduced equation* (9.3.109), and the equation for $u_0(x,y)$ cannot be solved there subject to arbitrary data $u_0(x,y_0) = f(x)$. A boundary layer must be introduced near $y = y_0$.

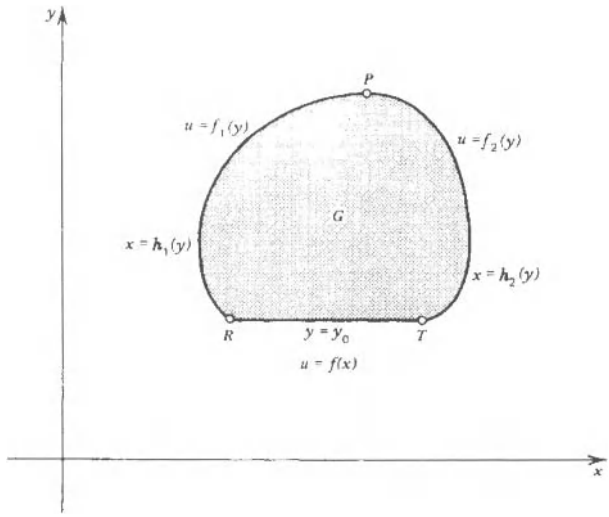

Figure 9.2 A characteristic boundary curve.

We set $y - y_0 = \epsilon^r \eta$ to emphasize the neighborhood of $y = y_0$ (with the exponent r to be determined) and insert this into (9.3.108). This gives, with $w(x, \eta) = u(x, y_0 + \epsilon^r \eta)$,

$$\epsilon^{1-2r} C(x, y_0) w_{\eta\eta}(x, \eta) - w_x(x, \eta) + c(x, y_0) w(x, \eta) + O(\epsilon^{1-r}) = g(x, y_0). \tag{9.3.119}$$

To achieve a balance between the reduced equation and the term $w_{\eta\eta}(x, \eta)$, we must set $r = \frac{1}{2}$. On expanding $w(x, \eta)$ in powers of $\epsilon^{1/2}$, we obtain for the leading term $w_0(x, \eta)$,

$$C(x, y_0) \frac{\partial^2 w_0(x, \eta)}{\partial \eta^2} - \frac{\partial w_0(x, \eta)}{\partial x} + c(x, y_0) w_0(x, \eta) = g(x, y_0). \tag{9.3.120}$$

which is a *parabolic partial differential equation.* We have the boundary and matching conditions

$$w_0(x, 0) = f(x), \quad \lim_{\eta \to \infty} w_0(x, \eta) = \lim_{y \to y_0} u_0(x, y) \tag{9.3.121}$$

because the boundary layer and the outer solution (9.3.116) must agree. An additional condition should be placed on $w_0(x, \eta)$ relating to its behavior near $x = h_1(y)$ or $x = h_2(y)$. Thinking of x as a timelike variable in (9.3.120), we see that (9.3.120) corresponds to a forward parabolic equation since $C(x, y_0) > 0$ by assumption. Thus it is appropriate to specify the values of $w_0(x, \eta)$ at $x = h_1(y)$ since we are considering a region where $x > h_1(y)$. However, near $x = h_1(y)$ away from the point

R, the outer solution $u_0(x, y)$ is valid, so that we may consider the (initial) condition near $x = h_1(y)$ to be equivalent to the condition (9.3.121). That is, $w_0(x, \eta)$ should agree with $u_0(x, y)$ away from the line $y = y_0$.

The solution of (9.3.120)–(9.3.121) is not straightforward if the coefficients C and c are not constant, and special techniques or possibly numerical methods must be used to solve for $w_0(x, \eta)$. If C and c are constants and $g = 0$, the solution can easily be obtained on following the procedure given in the preceding example, but we do not present it here.

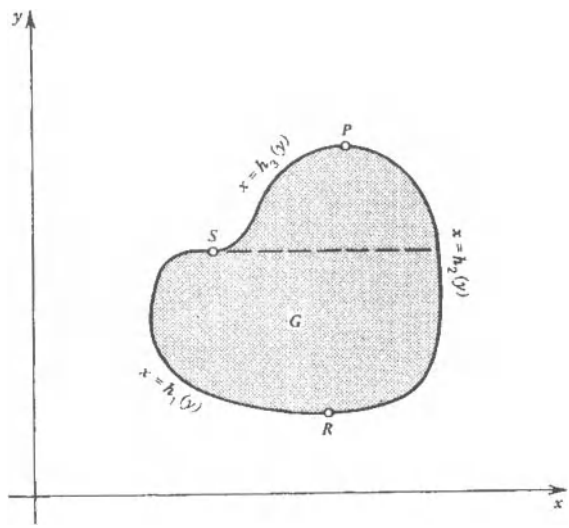

Figure 9.3 The nonconvex region G.

We conclude our discussion of (9.3.108) by examining the behavior of the solution of the Dirichlet problem for (9.3.108) near points on the boundary where the tangent line is horizontal. As shown in Figure 9.3, we are interested in points where the curve is either concave up or concave down or has a point of inflection. This contrasts with the situation considered previously where an entire portion of the boundary was horizontal. The boundary of G is comprised of three curves $x = h_1(y)$, $x = h_2(y)$, and $x = h_3(y)$. These curves connect the points P, R, and S, as shown in the figure. At the point P in the figure, the curve is concave down, at R it is concave up, and at S the curve has a point of inflection and a zero slope.

If Dirichlet boundary conditions are given for (9.3.108) and A and C are both positive, we find, as was shown earlier, that a boundary layer must be introduced along $x = h_2(y)$. The outer equation (9.3.109) can be solved with the data given on $x = h_1(y)$ and $x = h_3(y)$. We can even solve for $u_0(x, y)$ with data given at the point of inflection. However, the derivatives of $h_1(y)$, $h_2(y)$, and $h_3(y)$ are singular at one or more of the points P, R, and S. Thus the boundary layer solution breaks down at P and R and the second term in the ordinary perturbation series is singular

at S, since it contains derivatives of $h_1(y)$ or $h_3(y)$. The outer solution cannot be determined on the horizontal line that extends from S in Figure 9.3, and an internal boundary layer must be introduced along that line.

Let (x_0, y_0) represent the coordinates of the boundary point P, R, or S. We assume the boundary curve can be expressed as $y = F(x)$ near (x_0, y_0). Since the slope vanishes at (x_0, y_0), we have

$$y - y_0 = \frac{1}{2}F''(x_0)(x - x_0)^2 + \cdots \tag{9.3.122}$$

near P or R, and

$$y - y_0 = \frac{1}{6}F'''(x_0)(x - x_0)^3 + \cdots \tag{9.3.123}$$

near S, with $y_0 = F(x_0)$. We assume that $F''(x_0)$ and $F'''(x_0)$ are not zero in these equations.

To emphasize the neighborhood of the point (x_0, y_0), we introduce a (double) *coordinate stretching*

$$x - x_0 = \epsilon^r \xi, \qquad y - y_0 = \epsilon^s \eta, \tag{9.3.124}$$

with r and s to be determined. Inserting (9.3.124) into (9.3.108) and with $v(\xi, \eta) = u(x_0 + \epsilon^r \xi, y_0 + \epsilon^s \eta)$, we obtain

$$\epsilon^{1-2r} A v_{\xi\xi}(\xi, \eta) + 2\epsilon^{1-r-s} B v_{\xi\eta}(\xi, \eta) + \epsilon^{1-2s} C v_{\eta\eta}(\xi, \eta) - \epsilon^{-r} v_\xi(\xi, \eta) + \cdots = 0, \tag{9.3.125}$$

where A, B, and C are evaluated at (x_0, y_0) and only the leading terms in ϵ have been retained.

Now both r and s must be positive; otherwise, the neighborhood of (x_0, y_0) is not emphasized. Also, the term $v_\xi(\xi, \eta)$ must be retained in (9.3.125) to maintain a balance between terms from the reduced equation and higher derivative terms. Finally, we require that r and s be such that an approximation to the boundary curve $y = F(x)$ of the form (9.3.122) or (9.3.123) be retained near (x_0, y_0). Thus if (9.3.124) is inserted in (9.3.122), we have

$$\epsilon^s \eta = \frac{1}{2}\epsilon^{2r} F''(x_0)\xi^2 + \cdots, \tag{9.3.126}$$

and from (9.3.123) we obtain

$$\epsilon^s \eta = \frac{1}{6}\epsilon^{3r} F'''(x_0)\xi^3 + \cdots. \tag{9.3.127}$$

Consequently, for points of the type P or R we must put $2r = s$, while for points of type S, $3r = s$ in order to balance the leading terms in (9.3.126) and (9.3.127). It is then easily seen that the only way to balance $v_\xi(\xi, \eta)$ with a second derivative term in (9.3.125) is to set $1 - 2s = -r$.

Thus for points of type P or R we have $r = \frac{1}{3}$ and $s = \frac{2}{3}$, while for points of type S, $r = \frac{1}{5}$ and $s = \frac{3}{5}$. In either case, if $v(\xi, \eta)$ is expanded in appropriate powers of ϵ, we find that the leading term satisfies the equation

$$C(x_0, y_0)\frac{\partial^2 v_0(\xi, \eta)}{\partial \eta^2} - \frac{\partial v_0(\xi, \eta)}{\partial \xi} = 0. \tag{9.3.128}$$

To leading order in ϵ, the boundary curve has the form $\eta = \frac{1}{2}F''(x_0)\xi^2$ near points of type P and R, and it has the form $\eta = \frac{1}{6}F'''(x_0)\xi^3$ near points of type S. To facilitate the study of the boundary layer equation, we introduce the new variables

$$\alpha = \eta - \frac{1}{\kappa + 1}a_\kappa \xi^{\kappa+1}, \qquad \beta = \xi, \qquad \kappa = 1, 2, \tag{9.3.129}$$

where $a_1 = F''(x_0)$ and $a_2 = \frac{1}{2}F'''(x_0)$. In these coordinates, $\alpha = 0$ corresponds to the boundary curve. With $w_k(\alpha, \beta) = v_k(\alpha(\xi, \eta), \beta(\xi, \eta))$, it is possible to study the solution in the appropriate boundary layers, but we do not discuss this problem further.

If (9.3.75) is of *hyperbolic* or *parabolic type*, perturbation and boundary layer methods may again be used in the manner given to solve initial and boundary value problems. A particular hyperbolic problem was considered in Example 9.10, where the special difficulty arising from an excess of initial conditions for the reduced equations was examined. A general discussion of these problems akin to that given for the elliptic problem is not given here.

Parabolic Equation Method

A common problem that arises when using perturbation methods to solve elliptic, hyperbolic, and parabolic differential equations is where to place the boundary layer(s) if the data for the given problem result in an overdetermined problem for the reduced equation. It can be resolved, as has been done previously, by introducing boundary layer coordinates along the relevant parts of the boundary. Those portions of the boundary along which the boundary layer equations have solutions that decay exponentially away from the boundary are the regions where boundary layers can be located. An alternative method for determining the location of the boundary layer for *elliptic* and *hyperbolic equations* is now presented.

To begin, (9.3.75) is replaced by

$$\epsilon(Au_{xx}(x, y) + 2Bu_{xy}(x, y) + Cu_{yy}(x, y)) + au_x(x, y) + bu_y(x, y) = 0, \tag{9.3.130}$$

which may be of *elliptic* or *hyperbolic type*. The additional terms in (9.3.75) have been neglected for the sake of simplicity since they do not play a significant role in our discussion. Now (9.3.130) states that

$$a(x, y)u_x(x, y) + b(x, y)u_y(x, y) = O(\epsilon). \tag{9.3.131}$$

Assuming that $a(x,y)$, $b(x,y)$, $u(x,y)$ and their derivatives are smooth functions, we may express $u_{xx}(x,y)$, $u_{xy}(x,y)$, and $u_{yy}(x,y)$ either in terms of $u_{xx}(x,y)$ and $u_x(x,y)$ or in terms of $u_{yy}(x,y)$ and $u_y(x,y)$ correct to $O(\epsilon)$. For example, if $a(x,y) \neq 0$, we have

$$u_x(x,y) = -\frac{b(x,y)}{a(x,y)}\, u_y(x,y) + O(\epsilon), \tag{9.3.132}$$

so that on differentiating (9.3.132) and simplifying via (9.3.132), we obtain

$$u_{xx} = \left(\frac{b(x,y)}{a(x,y)}\right)^2 u_{yy} + \left[\frac{b(x,y)}{a(x,y)}\left(\frac{b(x,y)}{a(x,y)}\right)_y - \left(\frac{b(x,y)}{a(x,y)}\right)_x\right] u_y + O(\epsilon), \tag{9.3.133}$$

$$u_{xy} = -\frac{b(x,y)}{a(x,y)}\, u_{yy} - \left(\frac{b(x,y)}{a(x,y)}\right)_y u_y + O(\epsilon). \tag{9.3.134}$$

Inserting (9.3.133)–(9.3.134) into (9.3.130) yields

$$\epsilon\left\{\left[A\left(\frac{b}{a}\right)^2 - \frac{2Bb}{a} + C\right] u_{yy} + \left[\left(\frac{Ab}{a} - 2B\right)\left(\frac{b}{a}\right)_y - A\left(\frac{b}{a}\right)_x\right] u_y\right\}$$
$$+ a u_x + b u_y = O(\epsilon^2). \tag{9.3.135}$$

Similarly, if $b(x,y) \neq 0$, we obtain

$$\epsilon\left\{\left[C\left(\frac{a}{b}\right)^2 - \frac{2Ba}{b} + A\right] u_{xx} + \left[\left(\frac{Ca}{b} - 2B\right)\left(\frac{a}{b}\right)_x - C\left(\frac{a}{b}\right)_y\right] u_x\right\}$$
$$+ a u_x + b u_y = O(\epsilon^2). \tag{9.3.136}$$

The *parabolic equations* (9.3.135) and (9.3.136) may be solved by using a conventional perturbation series. The two leading terms in the series satisfy equations identical to those for the leading terms in the regular perturbation solution of (9.3.130), as is easily shown. Therefore, we expect the solutions of (9.3.135) or (9.3.136) to coincide, approximately, with the leading terms of the outer solution of (9.3.130) in regions where that solution is valid. As we shall see, the solutions of (9.3.135) and (9.3.136) remain valid in certain regions where the outer solution breaks down. Also, information about the location of boundary layers for (9.3.130) can be determined from the parabolic equations (9.3.135) and (9.3.136). The replacement of the given equation (9.3.130) by (9.3.135) or (9.3.136) and the analysis of the perturbation problem based on the latter equations is referred to as the *parabolic equation method.*

For example, if the data for (9.3.130) are not smooth, the discontinuities or singularities are propagated into the interior along the characteristics of the reduced equation (9.3.131). If the (9.3.130) is elliptic, the solution must be smooth in the interior, and if it is hyperbolic, the singularities must occur across the characteristics of (9.3.130). These characteristics generally do not coincide with those of the reduced

equation. The parabolic equations (9.3.135) and (9.3.136) generally smooth out the singularities in the data, so that its solutions are smooth across the characteristics of the reduced equation.

To determine the location of *boundary layers* for the elliptic or hyperbolic equation (9.3.130), we consider initial value problems for (9.3.135) or (9.3.136) along the boundary of the region. Those portions of the boundary for which the initial value problems for the parabolic equations (9.3.135) or (9.3.136) are not well posed determine where the boundary layers must be located.

To see how this works, we consider a simpler form of the *elliptic equation* (9.3.75) or (9.3.130),

$$\epsilon(A(x,y)u_{xx}(x,y) + 2B(x,y)u_{xy}(x,y) + C(x,y)u_{yy}(x,y)) - u_x(x,y) = 0 \tag{9.3.137}$$

and the convex region G considered previously. Since $a(x,y) = -1$ and $b(x,y) = 0$ in relation to (9.3.130), the *parabolic equation* (9.3.135) becomes

$$\epsilon\, C(x,y)u_{yy}(x,y) - u_x(x,y) = 0, \tag{9.3.138}$$

where we have dropped the $O(\epsilon^2)$ terms. Referring to the diagram of the convex region G given in Figure 9.1, we see that if $C(x,y) > 0$, the initial value problem for (9.3.138)—with x treated as the time variable—is well posed for data given on $x = h_1(y)$. But if $C(x,y) < 0$, the data must be assigned on $x = h_2(y)$ for the initial value problem to be well posed. Noting the relation between the parabolic equation solution and the outer solution, we conclude that the boundary layer must be placed along $x = h_2(y)$ if $C(x,y) > 0$ and along $x = h_1(y)$ if $C(x,y) < 0$. This result is consistent with that given earlier based on a direct boundary layer construction. It is not difficult to show that the present and the earlier method for determining the location of boundary layers both lead to the same result. However, this is not demonstrated here.

Given the *hyperbolic equation* with constant coefficients

$$\epsilon(u_{tt}(x,t) - c^2u_{xx}(x,t)) + u_t(x,t) - au_x(x,t) = 0, \tag{9.3.139}$$

a related parabolic equation is

$$\epsilon(a^2 - c^2)u_{xx}(x,t) + u_t(x,t) - au_x(x,t) = 0 \tag{9.3.140}$$

since $u_t(x,t) = au_x(x,t) + O(\epsilon)$. The parabolic equation (9.3.140) cannot absorb arbitrary initial data $u(x,0) = f(x)$ and $u_t(x,0) = g(x)$. Thus a boundary layer is generally required at $t = 0$. However, if $a^2 > c^2$, we find that even if appropriate initial values are determined for (9.3.140), the problem is not well posed in the region $t > 0$ since the coefficients of $u_{xx}(x,t)$ and $u_t(x,t)$ are both positive. A similar problem was encountered in Example 9.10, where it was shown that if $|a| > c$, the initial value problem for (9.3.139) is unstable. We assume, therefore, that $|a| < c$ and note that the parabolic equation approach can be used to determine whether the given problem is *stable* or *unstable*.

On considering an initial and boundary value problem for (9.3.139) in the interval $0 < x < l$ and with $t > 0$, we find that a boundary layer is required either along $x = 0$ or $x = l$ if the conventional perturbation approach is used. Now once initial data for the parabolic equation (9.3.140) have been determined, it is possible to solve (9.3.140) with the given boundary conditions on $x = 0$ and $x = l$. That is, a boundary layer need not be introduced at all on either boundary line. Nevertheless, it is usually simpler to solve the given problem with the use of boundary layers.

Since we are concerned with the solution near the lines $x = 0$ and $x = l$, we replace (9.3.140) with the parabolic equation

$$\epsilon\left[1 - \left(\frac{c}{a}\right)^2\right] u_{tt}(x,t) + u_t(x,t) - a u_x(x,t) = 0, \tag{9.3.141}$$

where we have used $u_x(x,t) = (1/a)u_t(x,t) + O(\epsilon)$. If $a > 0$, the coefficients of $u_{tt}(x,t)$ and $u_x(x,t)$ have the same sign since $c > |a|$, whereas if $a < 0$ they have opposite signs. Thus if (initial) data are given at $x = 0$ and (9.3.141) is to be solved for $x > 0$, the problem is well posed only if $a < 0$. Consequently, a boundary layer must occur near $x = 0$ in the case $a > 0$. Similarly, if data are assigned at $x = l$ and (9.3.141) is solved for $x < l$, the problem is well posed only if $a > 0$, so that a boundary layer near $x = l$ occurs when $a < 0$. We note that if $a = 0$, the lines $x = 0$ and $x = l$ are characteristics of the reduced equation $u_t(x,t) = 0$ and (9.3.140) is, in fact, the parabolic boundary layer equation. Again, these results concerning the location of the boundary layers are identical with those obtained by means of a boundary layer analysis.

Parabolic Equation Method: Specific Examples

Example 9.12. The Parabolic Equation Method for an Elliptic PDE. We begin by reexamining the problem of Example 9.11. That is, we consider the boundary value problem for the *elliptic equation*

$$\epsilon(u_{xx}(x,y) + u_{yy}(x,y)) + u_x(x,y) + b u_y(x,y) = 0,$$

$$u(x,0) = f(x), \quad u(0,y) = g(y), \quad u(x,L) = h(x) \tag{9.3.142}$$

within the region $0 < x < \infty$ and $0 < y < L$, with $0 < \epsilon \ll 1$, b = constant. Dirichlet conditions are specified on three sides of the region, and $u(x,y)$ is required to vanish as $x \to \infty$ in the interval $0 \le y \le L$.

Since $u_x(x,y) = -b u_y(x,y) + O(\epsilon)$ and $u_y(x,y) = -(1/b)u_x(x,y) + O(\epsilon)$ if $b \neq 0$, the *parabolic equations* for (9.3.142) are

$$\epsilon(1 + b^2)u_{yy}(x,y) + u_x(x,y) + b u_y(x,y) = 0, \tag{9.3.143}$$

$$\epsilon\left(1 + \frac{1}{b^2}\right) u_{xx}(x,y) + u_x(x,y) + b u_y(x,y) = 0. \tag{9.3.144}$$

From (9.3.143) we see that a problem with initial data given at $x = 0$ and $u(x, y)$ to be determined for $x > 0$ is not well posed. Also, (9.3.144) shows that if $b > 0$, a problem with initial data at $y = 0$ to be solved in the region $y > 0$ is not well posed. If $b < 0$ and initial data are given at $y = L$ and (9.3.144) is to be solved for $y < L$, the resulting problem is not well posed. If $b = 0$, (9.3.144) is not valid and in this case there are parabolic boundary layers at $y = 0$ and $y = L$.

We conclude from this discussion that a boundary layer should occur at $x = 0$ in all cases. If $b > 0$, a boundary layer occurs at $y = 0$, while if $b < 0$, a boundary layer occurs at $y = L$. These results are consistent with those obtained in Example 9.11 by a different method.

We now suppose that $b < 0$ in (9.3.142) so that the boundary layer must be located at $y = L$. (The cases with $b > 0$ and $b = 0$ were considered in Example 9.11.) It is then possible to solve the parabolic equation (9.3.144) with the data given along the lines $y = 0$ and $x = 0$. Even though $x = 0$ was shown to be the site of a boundary layer, the initial and boundary value problem for (9.3.144) is well posed. Thus we consider the equation

$$\epsilon\lambda^2 u_{xx}(x, y) + u_x(x, y) - c^2 u_y(x, y) = 0, \tag{9.3.145}$$

where we have set $\lambda^2 = 1 + (1/b)^2$ and $b = -c^2$, for convenience of notation. The initial and boundary conditions are

$$u(x, 0) = f(x), \qquad u(0, y) = g(y). \tag{9.3.146}$$

Even though (9.3.77)–(9.3.78) was formulated in the region $0 < x < \infty$ and $0 < y < L$ so that $g(y)$ in (9.3.146) is unspecified for $y > L$, we may solve (9.3.145)–(9.3.146) in the first quadrant $0 < x < \infty$ and $0 < y < \infty$ by extending $g(y)$ arbitrarily over the interval $y > L$. Because of the causality property for (9.3.145), the solution $u(x, y)$ at any value of y with $0 < y \leq L$ depends only on data given for smaller (nonnegative) values of y. As a result, the values assigned to $g(y)$ for $y > L$ do not affect the solution in $0 < y \leq L$.

To solve (9.3.145), we set $u(x, y) = \exp(-x/2\epsilon\lambda^2 - y/4\epsilon\lambda^2 c^2)v(x, y)$, and find that $v(x, y)$ satisfies the initial and boundary value problem

$$v_y = \frac{\epsilon\lambda^2}{c^2} v_{xx}, \quad v(x, 0) = f(x) \exp\left(\frac{x}{2\epsilon\lambda^2}\right), \; v(0, y) = g(y) \exp\left(\frac{y}{4\epsilon\lambda^2 c^2}\right), \tag{9.3.147}$$

in the region $x > 0$, $y > 0$. Using the results of Example 5.5, we find that $u(x, y)$ can be expressed as

$$u(x, y) = \frac{cx}{\sqrt{4\pi\epsilon\lambda^2}} \int_0^y \left\{ \exp\left[-\frac{(y + c^2 x - s)^2}{4\epsilon c^2 \lambda^2 (y - s)} \right] \frac{g(s)}{(y - s)^{3/2}} \right\} ds$$

$$+ \int_0^\infty \left[G\left(\frac{y}{c^2} + x - s, y\right) - G\left(\frac{y}{c^2} + x + s, y\right) \exp\left(\frac{s}{\epsilon\lambda^2}\right) \right] f(s)\, ds, \tag{9.3.148}$$

where $G(x,y)$ is given as $G(x,y) = c/\sqrt{4\pi\epsilon\lambda^2 y}\, \exp\left(-c^2x^2/4\epsilon\lambda^2 y\right)$, which is the fundamental solution of the parabolic equation in (9.3.147).

In the first integral in (9.3.148) we have $y - s \geq 0$ and $x \geq 0$. Thus when $x > 0$, we see that the contribution from the entire integral is exponentially small if $\epsilon \ll 1$, since the exponential is uniformly small. This term yields the boundary layer behavior near $x = 0$. From the expression for $G(x,y)$ given above, we see that as $\epsilon\lambda^2 y/c^2 \to 0$, $G(x,y) \to \delta(x)$ in view of the Dirac delta function behavior of the fundamental solution discussed previously. Thus as $\epsilon \to 0$, we find that $G[(y/c^2) + x \pm s, y] \to \delta[(y/c^2) + x \pm s]$. In the region where $y > 0$, $x \geq 0$, and $s \geq 0$, the delta function $\delta[(y/c^2) + x + s]$ vanishes. Even though $e^{s/\epsilon\lambda^2}$ grows exponentially as $\epsilon \to 0$ with positive s, its product with the term $G[(y/c^2) + x + s, y]$ may be shown to vanish as $\epsilon \to 0$.

Thus if $\epsilon \ll 1$, we find that (9.3.148) reduces to (on replacing c^2 by $-b$)

$$u(x,y) \approx f\left(x - \frac{y}{b}\right), \qquad x > 0,\ 0 \leq y < L, \tag{9.3.149}$$

to leading order in ϵ. The expression (9.3.149) represents the leading term in the outer solution of (9.3.142) that satisfies the boundary condition at $y = 0$. The term involving $g(y)$ in (9.3.148) is significant only in a boundary layer near $x = 0$. To satisfy the boundary condition at $y = L$, a boundary layer must be introduced as was done in Example 9.11. The solution can then be expressed in a composite form in the manner of (9.3.98). Instead, we use the results of Example 9.11 to construct a *composite expression* for the leading perturbation solution analogous to (9.3.98) for the case $b > 0$. We obtain

$$u(x,y) \approx f\left(x - \frac{y}{b}\right) + \left[g(y) - f\left(-\frac{y}{b}\right)\right] \exp\left(-\frac{x}{\epsilon}\right) + \left[h(x) - f\left(x - \frac{L}{b}\right)\right] \exp\left[-\frac{b(y-L)}{\epsilon}\right]. \tag{9.3.150}$$

The terms containing the exponentials are significant only within their respective boundary layers. Apart from exponentially small terms, $u(x,y)$ satisfies the boundary conditions at $y = 0$, $x = 0$, and $y = L$ and vanishes as $x \to \infty$ since $f(x)$ and $h(x)$ vanish there. Effects due to possibly incompatible boundary values at the points $(x,y) = (0,0)$ and $(x,y) = (0,L)$ are not considered.

Example 9.13. The Parabolic Equation Method for a Hyperbolic PDE. We consider a *signaling problem* for the *hyperbolic equation*

$$\epsilon(u_{tt}(x,t) - c^2 u_{xx}(x,t)) + u_t(x,t) - a u_x(x,t) = 0, \qquad x,t > 0, \tag{9.3.151}$$

with the initial and boundary conditions

$$u(x,0) = 0,\ u_t(x,0) = 0,\ x > 0, \quad u(0,t) = g(t),\ t > 0. \tag{9.3.152}$$

We assume that $|a| < c$, so that the problem is *stable*.

The relevant parabolic approximations for (9.3.151) are

$$u_t(x,t) - au_x(x,t) - \epsilon(c^2 - a^2)u_{xx}(x,t) = 0, \tag{9.3.153}$$

$$u_t(x,t) - au_x(x,t) - \epsilon\left(\left(\frac{c}{a}\right)^2 - 1\right)u_{tt}(x,t) = 0. \tag{9.3.154}$$

Since $c^2 > a^2$, the initial and boundary value problem for (9.3.153) is well posed. However, if initial data are given on $x = 0$ for (9.3.154), the problem is well posed for $a < 0$ and is not well posed for $a > 0$, as is immediately seen. This is consistent with the fact that the characteristics of the reduced equation $u_t(x,t) - au_x(x,t) = 0$ are the lines $x + at =$ constant. For $a < 0$, these characteristics do not simultaneously intersect the initial line $t = 0$ and the boundary line $x = 0$, so that the boundary condition in (9.3.152) can be assigned for the perturbation series solution of (9.3.151) without encountering any difficulty, and the perturbation series solution in the sector $x + at < 0$ is completely determined. Even though there are two initial conditions given at $t = 0$ and the equations for the perturbation terms are all of first order, we may equate all these terms to zero in the sector $x + at > 0$ since the initial conditions are both homogeneous. This can also be done if $a > 0$. However, in that case, the characteristics $x + at =$ constant intersect both the initial and boundary lines, so that a boundary layer is needed at $x = 0$ to accommodate the boundary condition assigned there.

As indicated, since the initial data are homogeneous, there is no need for an initial layer near $t = 0$ in this problem. We may use $u(x,0) = 0$ as the initial condition when solving (9.3.151) either by the conventional perturbation and boundary layer approach or by the use of the parabolic equation (9.3.153), as we now demonstrate. We rewrite (9.3.153) as

$$u_t(x,t) - au_x(x,t) - \epsilon\gamma^2 u_{xx}(x,t) = 0, \qquad x,\ t > 0, \tag{9.3.155}$$

with $\gamma^2 = c^2 - a^2$. The initial and boundary conditions are

$$u(x,0) = 0,\ x > 0, \quad u(0,t) = g(t),\ t > 0. \tag{9.3.156}$$

To solve (9.3.155)–(9.3.156), we proceed exactly as for the initial and boundary value problem for (9.3.145) treated earlier. Taking into account the slight change in notation, we find that

$$u(x,t) = \frac{\epsilon x}{\sqrt{4\pi\gamma^2}} \int_0^t \exp\left[-\frac{[x + a(t-\tau)]^2}{4\epsilon\gamma^2(t-\tau)}\right] \frac{g(\tau)}{[\epsilon(t-\tau)]^{3/2}}\, d\tau. \tag{9.3.157}$$

Now if $a > 0$, the argument of the exponential in (9.3.157) vanishes only at $x = 0$, in the region $x \geq 0$ and $t \geq 0$. Since $0 < \epsilon \ll 1$, we find that $u(x,t)$ is exponentially small for $x > 0$ and $t \geq 0$ and is only significantly greater than zero in a small (boundary) layer near $x = 0$. Thus the *outer solution* is given as $u(x,t) \approx 0$, $x > 0$, $t \geq 0$, apart from a layer of thickness $O(\epsilon)$ near $x = 0$.

If $a < 0$, the solution (9.3.157) is significantly different from zero for values of (x, t) in the first quadrant for which $x = -a(t - \tau)$. Since $t - \tau > 0$ and $t,\ \tau \geq 0$, this yields the characteristics of the reduced equation $u_t(x, t) - au_x(x, t) = 0$ that occupy the sector between the line $x + at = 0$ and the t-axis.

Now if $a = 0$ in (9.3.157), we know that as $x \to 0$, the expression (9.3.157) tends to $g(t)$ with the basic contribution from the integral coming from the value $\tau = t$. Similarly, we expect that as $\epsilon \to 0$ in (9.3.157) the main contribution to the integral comes from the value $\tau = x/a + t$ and that $u(x, t)$ tends to $g[x/a + t]$. We note that (9.3.157), has been written in terms of $\epsilon(t - \tau)$ and ϵx, to indicate that small values of x and $t - \tau$ are equivalent to considering small values of ϵ. Further, in the sector between $x + at = 0$ and the x-axis, the exponential term in (9.3.157) and, consequently, the integral is small. Therefore, we conclude that for small ϵ and $a < 0$ we have

$$u(x,t) \approx \begin{cases} g(x/a + t), & x + at < 0, \\ 0, & x + at > 0, \end{cases} \qquad x,\ t \geq 0. \tag{9.3.158}$$

This result is consistent with that obtained from the conventional perturbation method, as is easily shown. However, the expression (9.3.157) remains valid and useful for large values of x and t, as well as in the case where the boundary function $g(t)$ is not smooth or if $g(0) \neq 0$.

If the initial data $u(x, 0)$ and $u_t(x, 0)$ for (9.3.151) are not homogeneous, an initial layer of width $O(\epsilon)$ is required. The value of the solution at the edge of this initial layer may then be used as an initial condition for the parabolic equation (9.3.153).

Singular Perturbation of an Elliptic PDE in an Exterior Region

Example 9.14. The Singular Perturbation of the Helmholtz and Modified Helmholtz Equations. We begin by considering a boundary value problem for the *modified Helmholtz equation* in the region exterior to the unit disk given as $r^2 = x^2 + y^2 > 1$,

$$\epsilon^2(u_{xx}(x, y) + u_{yy}(x, y)) - u(x, y) = 0, \qquad r > 1, \tag{9.3.159}$$

with the boundary conditions

$$u(x, y)|_{r=1} = 1, \qquad u(x, y) \to 0 \text{ as } r \to \infty. \tag{9.3.160}$$

The exact solution of this problem is

$$u(x, y) = \frac{K_0(r/\epsilon)}{K_0(1/\epsilon)}, \tag{9.3.161}$$

where K_0 is the zero-order modified Bessel function of the second kind.

If we introduce the regular perturbation expansion

$$u(x,y) = \sum_{n=0}^{\infty} u_n(x,y)\epsilon^{2n} \tag{9.3.162}$$

into (9.3.159), we find that each term in the expansion equals zero. Thus the *outer solution* of this singular perturbation problem is $u(x,y) = 0$. This solution satisfies the boundary condition at infinity but not on the finite boundary $r = 1$. To deal with that boundary condition, we require a boundary layer at $r = 1$ and introduce the stretching transformation

$$r - 1 = \epsilon\sigma \tag{9.3.163}$$

into (9.3.159), which is transformed to polar coordinates and where we set $u = u(r)$, since the problem has no angular dependence. With $v(\sigma) = u(1 + \epsilon\sigma)$ we obtain $v''(\sigma) - v(\sigma) + \epsilon v'(\sigma) + O(\epsilon^2) = 0$ as the boundary layer equation, and expand $v(\sigma)$ as $v(\sigma) = v_0(\sigma) + \epsilon v_1(\sigma) + O(\epsilon^2)$. Then the boundary condition at $r = 1$ requires that $v_0(0) = 1$, while the remaining terms in the expansion vanish at $\sigma = 0$. The vanishing of the outer solution implies that each term in the expansion of $v(\sigma)$ must vanish as $\sigma \to \infty$. We easily find that $v(\sigma) = e^{-\sigma} - \frac{1}{2}\epsilon\sigma e^{-\sigma} + O(\epsilon^2)$. In terms of the radial variable r, we have

$$u(r) \approx \left[1 - \frac{1}{2}(r-1)\right] \exp\left[-\frac{1}{\epsilon}(r-1)\right], \tag{9.3.164}$$

which agrees with the asymptotic approximation of the exact solution (9.3.161) near $r = 1$ that is obtained by using the formula (6.7.34) with $1/\epsilon \to \infty$. Outside the boundary layer, the exact solution is exponentially small for small ϵ, so that its magnitude is smaller than any power of ϵ. Therefore, the regular perturbation expansion (9.3.162) of the solution must be zero.

The situation is quite different if we consider a similar exterior boundary value problem for the *Helmholtz* or *reduced wave equation*,

$$\epsilon^2(u_{xx}(x,y) + u_{yy}(x,y)) + u(x,y) = 0, \qquad r > 1, \tag{9.3.165}$$

with $u(x,y) = 1$ at $r = 1$. As $r \to \infty$, we require that $u(x,y)$ satisfy the radiation condition (7.4.51), where ω is replaced by $1/\epsilon$. The exact solution of this problem is

$$u(x,y) = \frac{H_0^{(1)}(r/\epsilon)}{H_0^{(1)}(1/\epsilon)}, \tag{9.3.166}$$

where $H_0^{(1)}$ is the zero-order Hankel function of the first kind.

The regular perturbation expansion (9.3.162) for this problem again yields $u(x,y) = 0$ as the outer solution. [We observe that $u(x,y) = 0$ satisfies the radiation condition.] To deal with the boundary condition at $r = 1$, we introduce the stretching transformation (9.3.163) and find that the leading term $v_0(\sigma)$ in the boundary layer expansion satisfies the equation $v_0''(\sigma) + v_0(\sigma) = 0$. At $\sigma = 0$ we have $v_0(\sigma) = 1$.

As $\sigma \to \infty$ the boundary layer term must match the outer solution $u(x, y) = 0$, so we require that $v_0(\sigma) \to 0$ as $\sigma \to \infty$. However, the boundary layer equation has only trigonometric (oscillatory) solutions, and the only solution that tends to zero at infinity is the zero solution, which does not satisfy the boundary condition at $\sigma = 0$.

Consequently, no boundary layer exists for the Helmholtz problem and the methods of this section are inappropriate for this problem. Indeed, the asymptotic approximation of the exact solution (9.3.166) [see (6.7.38)] for small ϵ is $u(r) \approx (1/\sqrt{r}) \exp[i(r-1)/\epsilon]$. Thus, the outer solution $u(x, y) = 0$ has no asymptotic validity, since the solution of the given problem does not decay exponentially for large r and small ϵ but oscillates rapidly and decays algebraically. A different approach is required for dealing with this and related problems, and this is given in Section 10.1.

Exercises 9.3

9.3.1. Obtain the exact and the regular perturbation solution of the initial value problem $\epsilon(u_t(x,t)+u_x(x,t))+u(x,t) = \sin(x)$, $-\infty < x < \infty$, $t > 0, u(x,0) = \sin x$, $-\infty < x < \infty$. (Note that the solution of the reduced equation satisfies the initial condition.) Construct a boundary layer solution, if it is needed, to obtain an approximate solution valid for all $t > 0$.

9.3.2. Construct a perturbation and exact solution for the following problem and compare both results. $\epsilon(u_t(x,t) - u_x(x,t)) - u(x,t) = 0$, $x > 0$, $-\infty < t < \infty$, $u(0,t) = 1$, $-\infty < t < \infty$.

9.3.3. Give a complete discussion of the perturbation solution of the following problem: $\epsilon(u_t(x,t) + u_x(x,t)) + u^2(x,t) = 0$, $-\infty < x < \infty$, $t > 0$, $u(x,0) = \sin(x)$, $-\infty < x < \infty$.

9.3.4. Obtain a perturbation and boundary layer solution of the following problem: $\epsilon(u_t(x,t)+u_x(x,t))+tu(x,t) = 1$, $-\infty < x < \infty$, $t > 0, u(x,0) = \sin x$, $-\infty < x < \infty$. *Hint*: The initial layer has width $O(\sqrt{\epsilon})$.

9.3.5. Derive the exact solution of the problem in Example 9.9 and compare it with the perturbation result for $t \approx 1$ and $t \approx 0$.

9.3.6. Let $f(x) = \sin x$ and $g(x) = 0$ in the problem of Example 9.10 and obtain the boundary layer solution near $t = 0$ and the outer solution (9.3.70). At $t = T > 0$, let (9.3.70) serve as initial data for the parabolic equation for $u_0(x,t)$ in (9.3.74) and let $u_1(x,t) = 0$ in that equation. Determine $u_0(x,t)$ for $t > T$ as a solution of the parabolic equation and characterize the range of values of t in which each of the approximate solutions obtained is valid.

9.3.7. Discuss the initial value problem $v_t(x,t) + cv_x(x,t) = \frac{1}{2}Dv_{xx}(x,t)$, $-\infty < x < \infty$, $t > 0, u(x,0) = \delta(x)$, $-\infty < x < \infty$, (see Section 1.1) from the point of view of a singular perturbation problem if we assume that the diffusion coefficient D

is small. Show that even though the reduced equation can absorb the initial condition, a boundary layer must be introduced near $x = ct$ since the solutions of the regular perturbation equations are singular there.

9.3.8. Consider the singular perturbation problem for the ODE $\epsilon y''(x) + a(x)y'(x) + b(x)y(x) = 0$, $0 < x < 1$, with the boundary conditions $y(0) = \alpha$, $y(1) = \beta$. Show that it is not possible, in general, to solve this problem using regular perturbation methods. (a) Assume that $a(x) > 0$ in $0 \leq x \leq 1$ and obtain an approximate solution of the problem by introducing a boundary layer at $x = 0$. (b) Assume that $a(x) < 0$ in $0 \leq x \leq 1$ and solve the problem approximately by introducing a boundary layer at $x = 1$.

9.3.9. Let $f(x) = x^2e^{-x}$, $h(x) = 0$, and $g(y) = \sin(\pi y/L)$ in the problem of Example 9.11. Discuss the perturbation and boundary layer solutions of the given problem in the following cases. (a) $b = 1$; (b) $b = 0$; (c) $b = -1$.

9.3.10. Use perturbation and boundary layer methods to obtain an approximate solution of the following problem: $\epsilon(u_{xx}(x,y) + u_{yy}(x,y)) + u_x(x,y) + u_y(x,y) = 0$, $x > 0$, $y > 0$, with $u(x,0) = xe^{-x}$, $u(0,y) = ye^{-y}$, and $\lim_{\{x\to\infty, y\to\infty\}} u(x,y) = 0$.

9.3.11. Obtain an approximate solution of the following singular perturbation problem: $\epsilon(u_{xx}(x,y) + u_{yy}(x,y)) - u(x,y) = 1$, $0 < x < \pi$, $0 < y < \pi$, $u(x,0) = u(0,y) = u(x,\pi) = u(\pi,y) = 0$.

9.3.12. Given the Dirichlet problem $\epsilon a(u_{xx}(x,y) + u_{yy}(x,y)) - u_x(x,y) + u(x,y) = 0$, $x^2 + y^2 < 1$, $u(x,y) = x$, $x^2 + y^2 = 1$ (a = constant), determine the location of the boundary layer and obtain an approximate solution, away from $(x,y) = (0,-1)$ and $(x,y) = (0,1)$ if (a) $a = -1$; (b) $a = 1$.

9.3.13. Obtain a perturbation and boundary layer solution of the following problem: $\epsilon(u_{xx}(x,y) + u_{yy}(x,y)) - u_x(x,y) = 0$, $x^2 + y^2 < 1$, $y > 0$, $u(x,0) = 1$, $-1 < x < 1$, $u(x,y) = y$, $x^2 + y^2 = 1$, $y > 0$. Discuss the solution away from the points $(-1,0)$, $(1,0)$, and $(0,1)$.

9.3.14. Consider the singular perturbation problem for the elliptic equation $\epsilon(u_{xx}(x,y) + u_{yy}(x,y)) - u_x(x,y) = 0$, $x^2 + y^2 < 1$, with the Dirichlet boundary condition $u(x,y) = f(x,y)$, $x^2 + y^2 = 1$. (a) Show that the point $(0,-1)$ on the circle is a point of type R. (b) Show that the point $(0,1)$ is of type P.

9.3.15. Use perturbation and boundary layer methods to obtain an approximate solution of the following parabolic problem. $\epsilon u_{xx}(x,t) + u_x(x,t) - u_t(x,t) = 0$, $0 < x < l$, $t > 0$, with the initial and boundary conditions $u(x,0) = f(x)$, $u(0,t) = g(t)$, $u(l,t) = h(t)$.

9.3.16. Obtain an approximate solution of the singular perturbation problem $\epsilon u_{xx}(x,t) + au_x(x,t) - u_t(x,t) = 0$, $x > 0$, $t > 0$, with the initial and boundary conditions $u(x,0) = f(x)$, $u(0,t) = g(t)$ if (a) $a = 1$; (b) $a = -1$.

9.3.17. Use the parabolic equation method to locate the boundary layers for the elliptic equation $\epsilon(u_{xx}(x,y)+u_{yy}(x,y))-u_x(x,y)=0$ if Dirichlet conditions are assigned on the boundaries of the following regions: (a) $0<x<\pi,\ 0<y<\pi$; (b) $0<x<\pi,\ 0<y<\infty$; (c) $x^2+y^2<1$.

9.3.18. Complete the solution of the problem for (9.3.142) given in Example 9.12 by obtaining a boundary layer solution near $y=L$. Combine the result of this exercise and of the example in the text to yield a composite approximate solution of the problem in the manner of (9.3.98).

9.3.19. Solve the following initial and boundary value problem using the parabolic equation method: $\epsilon(u_{tt}(x,t)-c^2u_{xx}(x,t))-u_t(x,t)-au_x(x,t)=0,\ 0<x<l,\ t>0,\ u(x,0)=u_t(x,0)=0,\ 0<x<l,\ u(0,t)=1,\ u(l,t)=0,\ t>0$. Assume that $|a|<c$ and consider these cases. (a) $a<0$; (b) $a=0$; (c) $a>0$.

9.3.20. Consider the problem $\epsilon(u_{xx}(x,y)+u_{yy}(x,y))+u_x(x,y)=e^{-x},\ 0<x<\infty,\ 0<y<L,\ u(x,0)=u(0,y)=u(x,L)=0$, with $u(x,y)\to 0$ as $x\to\infty$. Apply the parabolic equation approach and obtain $\epsilon u_{yy}(x,y)+u_x(x,y)=(1+\epsilon)e^{-x}$. Solve the problem using the parabolic equation method and the perturbation and boundary layer method and compare results.

9.3.21. Use perturbation and boundary layer methods to solve $\epsilon u_{xx}(x,y)+u_{yy}(x,y)+u_x(x,y)=0,\ 0<x<\pi,\ 0<y<\pi,\ u(x,0)=u(0,y)=u(x,\pi)=0,\ u(\pi,y)=f(y)$.

9.3.22. Apply boundary layer and perturbation methods to solve the following initial value problem: $\epsilon(u_{ttt}(x,t)-a^2u_{xxt}(x,t))+u_{tt}(x,t)-c^2u_{xx}(x,t)=0,\ -\infty<x<\infty,\ t>0,\ u(x,0)=f(x),\ u_t(x,0)=g(x),\ u_{tt}(x,0)=k(x)$, with $a>c$. *Hint*: Proceed as in Example 9.10.

9.3.23. Apply perturbation and boundary layer methods to the following problem. $u_{tt}(x,t)+\epsilon^2u_{xxxx}(x,t))=e^{-t},\ 0<x<l,\ t>0, u(x,0)=u_t(x,0)=0,\ 0<x<l, u(0,t)=u_x(0,t)=u(l,t)=u_x(l,t)=0,\ t>0$. Find the outer solution and show that there are parabolic boundary layers at $x=0$ and $x=l$.

9.3.24. Determine the location of the boundary layers for the following singular perturbation problem: $\epsilon(u_{xx}(x,y)+u_{yy}(x,y))+u_x(x,y)=0,\ y-x^3<0,$ $u(x,x^3)=\begin{cases}0, & x\le 0,\\ e^{-x}, & x>0.\end{cases}$ *Hint*: $(x,y)=(0,0)$ is a point of type S.

9.3.25. Obtain the exact and approximate solution of the modified Helmholtz equation in the region exterior to the unit sphere: $\epsilon^2\nabla^2u(x,y,z)-u(x,y,z)=0,\ r>1$, with the boundary conditions (9.3.160) where $r^2=x^2+y^2+z^2$. Proceed as in Example 9.14.

9.4 MAPLE METHODS

Regular Perturbation Expansions

We have created a procedure $RegPerturb$ that automates the construction of regular perturbation expansions for ordinary and partial differential equations that contain a small parameter. Initial value, boundary value, and initial and boundary value problems can be treated. The output exhibits the equations that each of the terms in the expansion must satisfy. These may be ODEs, PDEs, or algebraic equations. (The number of terms considered is prescribed by an argument in the procedure.) Appropriate initial and/or boundary conditions for each term are also displayed. The procedure is given as $RegPerturb(Equation, [IC], [BC], u, [indxvars], [xrange], [trange], \epsilon^r, n)$. Instead of discussing the procedure in general terms, we apply it to the examples presented in this chapter.

We begin with Example 9.1. As was shown there, we can put $u = u(r)$ and consider the perturbation problem for the ODE, $ODE1 = u''(r) + (1/r)u'(r) + \epsilon^2 u(r) = 0$, $0 < r < 1$, with the boundary conditions $u(0)$ = bounded and $u(1) = 1$. We use the procedure $RegPerturb(ODE1, [\,], [u = bounded, u = 1], u, [r], [r = 0..1], [\,], \epsilon^2, 1)$. There is only one independent variable r and there are boundary conditions at $r = 0$ and $r = 1$. There is no t dependence and no initial condition so that the second and seventh arguments are empty and given as $[\,]$. The dependent variable is u. The perturbation expansion is given in terms of the parameter ϵ^2 as $u(r) = \sum_0^\infty u_n(r)\epsilon^{2n}$. The last argument in the procedure $n = 1$ indicates that only the boundary value problems for $u_0(r)$ and $u_1(r)$ will be exhibited in the output of the procedure. The result for $u_1(r)$ is given as $d^2u_1(r)/dr^2 + (1/r)du_1(r)/dr + u_0(r) = 0$, $u_1(0)$ = bounded, $u_1(1) = 0$. This agrees with (9.2.31). The ODEs can be solved recursively using the Maple procedure $dsolve$.

The perturbation problem of Example 9.2 is a Cauchy problem given over the infinite interval $-\infty < x < \infty$ as $PDE1 = u_t(x,t) + \epsilon u(x,t) = u_{xx}(x,t)$, $u(x,0) = f(x)$. The procedure $RegPerturb(PDE1, [u = f(x)], [\,], u, [x], [\,], [t = 0..\infty], \epsilon, 1)$, determines the initial value problems for the first two terms in the perturbation expansion (9.2.38). In this case there is an x and t dependence in the problem. However, the initial condition for the problem is given over the entire x-axis so that no specific boundary conditions are prescribed, except for requiring the solution to be bounded. Thus the third and sixth arguments are empty and given as $[\,]$. The problem for $u_1(x,t)$ is determined to be $\partial u_1(x,t)/\partial t - \partial^2 u_1(x,t)/\partial x^2 = -u_0(x,t)$, $u_1(x,0) = 0$.

To apply the method of multiple scales as was done in Example 9.2, we use $RegPerturb(u_t + \epsilon(u_\tau + u) - u_{xx} = 0, [u = f(x)], [\,], u, [x, \tau], [\,], [t = 0..\infty], \epsilon, 1)$. We invoke the PDE (9.2.44) and introduce the slow time variable in the fifth argument of the procedure. The initial value problem for the term $u_1(x,t,\tau)$ in the perturbation expansion is determined to be $\partial u_1(x,t,\tau)/\partial t - \partial^2 u_1(x,t,\tau)/\partial x^2 = -\left[\partial u_0(x,t,\tau)/\partial \tau + u_0(x,t,\tau)\right]$, $u_1(x,0,0) = 0$.

The perturbation problem for the nonlinear Klein-Gordon equation considered in Example 9.3 leads to the consideration of the PDE, $PDE2 = u_{tt}(x,t) - \gamma^2 u_{xx}(x,t) + c^2 u(x,t) - \epsilon^2 \sigma u^3(x,t) = 0$, with the initial conditions $u(x,0) = \cos(kx)$ and

$u_t(x,0) = 0$. The procedure $RegPerturb(PDE2, [u = \cos(kx), u_t = 0], [\,], u, [x, \tau], [\,], [t = 0..\infty], \epsilon, 3)$ finds the problems that the first four terms in the perturbation expansion (9.2.55) must satisfy. The problem for $u_3(x,t)$ is given as $\partial^2 u_3(x,t)/\partial t^2 - \gamma^2\, \partial^2 u_3(x,t)/\partial x^2 + c^2 u_3(x,t) - 3\sigma u_0(x,t)^2 u_1(x,t) = 0$, with the initial conditions $u_3(x,0) = 0$, $\partial u_3(x,0)/\partial t = 0$.

The perturbation problem of Example 9.4 involves the PDE, $PDE3 = u_{tt}(x,t) - \left(c_0^2 + \sum_{n=1}^{\infty} c_n(\epsilon x)^n\right) u_{xx}(x,t) = 0$, with the initial conditions $u(x,0) = f(x)$ and $u_t(x,0) = 0$. Invoking the procedure $RegPerturb(PDE3, [u = f(x), u_t = 0], [\,], u, [x, \tau], [\,], [t = 0..\infty], \epsilon, 1)$ yields $\partial^2 u_1(x,t)/\partial t^2 - c_0^2\, \partial^2 u_1(x,t)/\partial x^2 = x c_1\, \partial^2 u_0(x,t)/\partial x^2$, $u_1(x,0) = 0$, $\partial u_1(x,0)/\partial t = 0$, in agreement with the results (9.2.67) and (9.2.68).

Boundary perturbations can also be treated with the procedure. We consider Example 9.5 for Laplace's equation. The boundary condition at $y = \epsilon x$ is expressed as $u(x, \epsilon x) = u(x,0) + \epsilon x u_y(x,0) = f(x)$. The procedure is given as $RegPerturb(u_{xx} + u_{yy} = 0, [u + \epsilon x u_y = f(x), u = 0], [u = 0, u = 0], u, [x], [x = 0..l], [y = 0..L], \epsilon, 1)$. Each term in $u_n(x,y)$ in the perturbation expansion satisfies Laplace's equation and vanishes on the three sides $x = 0$, $y = L$, $x = l$ of the unperturbed rectangle. On the fourth side, the procedure yields the boundary condition for $u_1(x,y)$ as $u_1(x,0) = -x\, \partial u_0(x,0)/\partial y$.

The *eigenvalue perturbation problem* of Example 9.6 for the function $M(x,y)$ in the square of side π, with Dirichlet boundary conditions on the sides, are dealt with by using the procedure $RegPerturb(M_{xx} + M_{yy} - \epsilon x y M + \lambda(\epsilon) M = 0, [M = 0, M = 0], [M = 0, M = 0], M, [x], [x = 0..\pi], [y = 0..\pi], \epsilon, 1)$. The problem for the term $M_1(x,y)$ in the procedure is $\partial^2 M_1(x,y)/\partial x^2 + \partial^2 M_1(x,y)/\partial y^2 - xyM_0(x,y) + \lambda'(0) M_0(x,y) + \lambda(0) M_1(x,y) = 0$, with the boundary conditions $M_1(x,0) = 0, M_1(x,\pi) = 0, M_1(0,y) = 0, M_1(\pi,y) = 0$. The result agrees with that found in Example 9.6.

The problem of Example 9.7 can be solved completely using Maple methods, but we do not present this discussion here.

Singular Perturbations and Boundary Layer Methods

When dealing with singular perturbation problems, perturbation expansions continue to play a significant role. Initially, the given (singular) perturbation problem is expanded in a perturbation series and this determines the outer solution. Once the boundary or initial layer equation is determined after carrying out a stretching procedure, it is also expanded in a perturbation series. The procedure $RegPerturb$ can be used to determine the problems that each term in the perturbation expansion of the given equation or the boundary layer equation must satisfy.

A procedure $BoundLayer$ has been constructed that determines the boundary layer equation associated with a given stretching transformation of a PDE. It is given as $BoundLayer(PDE, DVar, [IndVars], \epsilon, BLLoc, r, StretchVar)$. The second and third arguments of the procedure specify the dependent and independent variables, respectively, in the PDE. The fourth argument $BLLoc$ is an equation that determines the location of the boundary, initial, or interior layer. Thus, if an initial

layer at $t = 0$ is required, the fourth argument is $t = 0$. The fifth argument specifies the stretching exponent r. The final argument prescribes the stretched variable. For example, if we have an initial layer at $t = 0$ with the stretching $t = \epsilon^r \tau$, then $StretchVar = \tau$. If a double stretching is required, the procedure must be applied twice.

We begin by reexamining Example 9.8. The Cauchy problem is given as $SPPDE1 = \epsilon(u_t(x,t)+u_x(x,t))+u(x,t) = \sin(t)$, $u(x,0) = f(x)$. The procedure $RegPerturb(SPPDE1, [u = f(x)], [\,], u, [x], [\,], [t = 0..\infty], \epsilon, 1)$ yields for the first two terms in the expansion of the outer solution, $u_0(x,t) = \sin(t)$, $u_0(x,0) = f(x)$, and $u_1(x,t) = -(\partial u_0(x,t)/\partial t + \partial u_0(x,t)/\partial x), u_1(x,0) = 0$, in agreement with the results obtained in the example. The initial conditions cannot be satisfied and an initial layer at $t = 0$ is required.

We invoke the procedure $BoundLayer(SPPDE1, u, [x,t], \epsilon, t = 0, r, \tau)$. The stretching transformation (9.3.8) is used and the output of the procedure is the initial layer equation $\epsilon u_x(\tau, x, \epsilon, r) + \epsilon^{1-r} u_\tau(\tau, x, \epsilon, r) + u(\tau, x, \epsilon, r) = \sin(\epsilon^r \tau)$. The dependent variable in the initial layer $\hat{u}(x,\tau)$ of Example 9.8 is given by the procedure in the form $u(\tau, x, \epsilon, r)$. On setting the stretching exponent $r = 1$, we obtain the initial layer equation (9.3.10). If we put $r = 1$ in the procedure, the required initial layer equation is obtained directly. It is given as $IL1 = \epsilon \hat{u}_x(x,\tau) + \hat{u}_\tau(x,\tau) + \hat{u}(x,\tau) = \sin(\epsilon\tau)$. Then the procedure $RegPerturb(IL1, [\hat{u} = f(x)], [\,], \hat{u}, [x], [\,], [\tau = 0..\infty], \epsilon, 2)$ yields the problems (9.3.12)–(9.3.13).

Next, we consider Example 9.9 and the internal layer that occurs in that problem. The PDE is $SPPDE2 = \epsilon(u_t(x,t) + u_x(x,t)) + (t-1)^2 u(x,t) = 1$, and there is an internal layer at $t = 1$. Then $BoundLayer(SPPDE2, u, [x,t], \epsilon, t = 1, 1/3, \tau, -2/3)$, which contains an additional eighth argument, yields the stretching of both and independent and the dependent variable given as in (9.3.37) as $t-1 = \epsilon^{1/3}\tau$ and $u = \epsilon^{-2/3} v$. The internal boundary layer equation is given as $IL2 = \epsilon^{1/3} v_x(x,\tau) + v_\tau(x,\tau) + \tau^2 v(x,\tau) = 1$. The procedure $RegPerturb(IL2, [v = f(x)], [\,], v, [x], [\,], [\tau = 0..\infty], \epsilon^{1/3}, 2)$ determines the problems satisfied by the first three terms in the perturbation expansion (9.3.39). Even though an initial condition $v(x,0) = f(x)$ is specified in the procedure, it plays no role in the solution of the problem, as indicated in our discussion of the solution in Example 9.9. (The initial condition must be entered for the procedure to work.)

Parabolic Equation Method

Examples 9.11 and 9.12 both deal with the *elliptic equation* $SPPDE3 = \epsilon[u_{xx}(x,y)+u_{yy}(x,y)]+u_x(x,y)+bu_y(x,y) = 0$, with $0 < x < \infty$, $0 < y < L$, and b = constant, and the boundary conditions $u(x,0) = f(x)$, $u(0,y) = g(y)$, $u(x,L) = h(x)$. The perturbation problem (9.4.12)–(9.4.13) was solved using the *boundary layer method* in Example 9.11 and the *parabolic equation method* in Example 9.12. We conclude our discussion of Maple methods by applying both methods to the solution of the problem.

On applying the procedure *BoundLayer* to the PDE (9.4.12) at the boundaries $y = 0$, $y = L$, and $x = 0$, with the stretching exponent $r = 1$ and $b \neq 0$, we

reproduce the results of Example 9.11. If $b = 0$ in (9.4.12), the choice $r = 1/2$ results in parabolic boundary layer equations at $y = 0$ and $y = L$, as before.

The procedure *ParabolicEqMeth* carries out the parabolic equation method. The reduced equation for the given PDE must be of first order. The output exhibits two forms of the parabolic equation, if possible. The procedure has the form $ParabolicEqMeth(PDE, depvar, [indvars], \epsilon)$, with *depvar* as the dependent variable and [*indvars*] as a list of the independent variables. On invoking the procedure $ParabolicEqMeth(SPPDE3, u, [x, y], \epsilon)$, we obtain the two parabolic equations $\epsilon(1+b^2)u_{yy}(x,y)+u_x(x,y)+bu_y(x,y) = 0$, and $\epsilon(1+1/b^2)u_{xx}(x,y)+u_x(x,y)+bu_y(x,y) = 0$, as given in (9.3.143)–(9.3.144) in Example 9.12.

Our discussion of Maple methods has been restricted to the presentation of the procedures *RegPerturb*, *BoundLayer*, and *ParabolicEqMeth* that we have constructed and their use in analyzing regular and singular perturbation problems. The results obtained via these methods can then be analyzed further by using built-in Maple procedures, as is the case for the examples presented in this chapter.

Exercises 9.4

9.4.1. (a) Solve the ODE (9.2.28) with the boundary conditions given in the text using the Maple procedure *dsolve* and obtain the solution (9.2.29). (b) Apply the procedure *RegPerturb* to the problem of part (a) with the last argument in the procedure given as the number 2. Solve the three resulting equations recursively using *dsolve*. (c) Expand the solution found in part (a) using the Maple procedure *taylor* and compare with the result obtained from part (b).

9.4.2. Use *RegPerturb* to analyze the problem of Example 9.2 in the manner discussed in the text.

9.4.3. Apply *RegPerturb* to the PDE (9.2.53) with general initial conditions and obtain the equations and initial conditions that the first four terms in the perturbation expansion of the solution must satisfy.

9.4.4. Apply *RegPerturb* to the perturbation problem of Example 9.4 and obtain the problems satisfied by $u_0(x,t)$, $u_1(x,t)$, and $u_2(x,t)$.

9.4.5. Apply *RegPerturb* to the boundary perturbation problem of Example 9.5 and obtain the problems satisfied by $u_0(x,t)$ and $u_1(x,t)$.

9.4.6. Use *RegPerturb* to obtain the problems for the first two terms for the eigenvalue perturbation problem of Example 9.6.

9.4.7. Solve the problem of Example 9.7 using Maple methods.

9.4.8. Use the procedures *RegPerturb* and *BoundLayer* to obtain the results for Example 9.8 given in the text.

9.4.9. Apply *BoundLayer* and *RegPerturb* to the problem of Example 9.9.

9.4.10. Apply the procedures *BoundLayer* and *ParabolicEquationMeth* to the problems of Examples 9.11 and 9.12 in the manner presented in the text.

CHAPTER 10

ASYMPTOTIC METHODS

This chapter continues our presentation of methods that yield approximate solutions of initial and/or boundary value problems for PDEs. We refer to them as *asymptotic methods*, for reasons indicated in Section 9.1, even though they do not all fit precisely into that category.

We begin with a consideration of equations with a *large parameter* and concentrate mostly on the (elliptic) *reduced wave equation* or the *Helmholtz equation*. Linear and nonlinear problems are considered. A common feature in all the problems is that the solutions contain rapidly oscillating terms.

Next we consider methods for describing the *propagation of singularities of hyperbolic equations*. They permit the solutions of hyperbolic equations to be analyzed near singular regions, such as wave fronts, without having to solve the full problems. Even though the methods and results appear to have a somewhat different character than those considered in the remainder of this chapter and in Chapter 9, they are related to the techniques of Section 10.1, as will be seen.

Finally, we present a method for simplifying linear and nonlinear equations and systems of equations by approximating them, in a systematic way, by simpler PDEs. The simpler, more easily solvable equations generally model one or more of the

Partial Differential Equations of Appplied Mathematics, Third Edition. By Erich Zauderer

basic physical features associated with the full equations or systems. The study of simplified models of given problems is fundamental in applied mathematics.

10.1 EQUATIONS WITH A LARGE PARAMETER

Linear Reduced Wave Equation

In the study of boundary value problems for the linear *reduced wave equation*

$$\nabla^2 u(\mathbf{x}) + k^2 n^2(\mathbf{x}) u(\mathbf{x}) = 0, \tag{10.1.1}$$

a case of greatest interest in the context of wave propagation occurs when the (constant) parameter k is large. The function $n(\mathbf{x})$ in (10.1.1) is known as the *index of refraction*. The connection between (10.1.1) and the wave equation is given in Exercise 2.4.7.

Assuming that k is large, we apply a perturbation procedure to (10.1.1) and expand $u(\mathbf{x})$ in inverse powers of k as

$$u(\mathbf{x}) = \sum_{j=0}^{\infty} u_j(\mathbf{x}) k^{-j}. \tag{10.1.2}$$

Inserting (10.1.2) into (10.1.1) and equating like powers of k, we immediately conclude that $u_0(\mathbf{x})$ and all further terms in the expansion must vanish. Putting $\epsilon = 1/k$, we recognize that (10.1.1) is, in fact, a *singular perturbation problem* and that asymptotic expansions of (10.1.1) must have a form different from that given in (10.1.2) (see Example 9.14). To determine appropriate expansion forms, we consider some exact solutions of (10.1.1) in the case where the index of refraction n is a constant.

For the two-dimensional problem, the simplest solutions of (10.1.1) are the *plane wave* solutions

$$u(x, y) = \exp[ikn(x \cos\theta + y \sin\theta)], \tag{10.1.3}$$

where θ and n are constants. With $r^2 = (x - \xi)^2 + (y - \eta)^2$, another solution of (10.1.1) in two dimensions is

$$u(x, y) = J_0(knr), \tag{10.1.4}$$

where J_0 is the zero-order Bessel function. For large values of k, assuming that nr is not small, the asymptotic expansion of the Bessel function $J_0(knr)$ (see Exercise 5.7.1) yields

$$\begin{aligned} u(x, y) = J_0(knr) &\approx \left(\frac{2}{\pi knr}\right)^{1/2} \cos\left(knr - \frac{\pi}{4}\right) \\ &= \left(\frac{1}{2\pi knr}\right)^{1/2} \exp\left(iknr - \frac{i\pi}{4}\right) + \left(\frac{1}{2\pi knr}\right)^{1/2} \exp\left(-iknr + \frac{i\pi}{4}\right). \end{aligned} \tag{10.1.5}$$

Each of the exponential terms in (10.1.5) represents an asymptotic *cylindrical wave* solution of (10.1.1).

In the three-dimensional case with $r^2 = (x-\xi)^2+(y-\eta)^2+(z-\zeta)^2$, a *spherical wave* solution of (10.1.1) is

$$u(x,y,z) = \frac{1}{r}\, e^{iknr}. \tag{10.1.6}$$

The exact and asymptotic solutions (10.1.3)–(10.1.6) of (10.1.1) indicate that when the solutions of (10.1.1) are expanded asymptotically for large k, they have the form of a *rapidly varying exponential term* multiplied by an *amplitude term.* On considering further terms in the asymptotic expansion of the Bessel function $J_0(knr)$, we find that the amplitude terms in (10.1.5) are given as series in inverse powers of k. On expanding the exponentials in (10.1.5) in powers of k, we see that the solution $u(x,y)$ contains both negative and positive powers of k in its expansion. This explains why the series (10.1.2) does not lead to a useful result.

Eiconal and Transport Equations of Geometrical Optics

Noting these results, we look for *asymptotic solutions* of (10.1.1) in the form

$$u(\mathbf{x}) = v(\mathbf{x})\, e^{ik\phi(\mathbf{x})}, \tag{10.1.7}$$

where $\phi(\mathbf{x})$ is the *phase term* and $v(\mathbf{x})$ is the *amplitude term.* (Although we are generally interested in real solutions of (10.1.1), such as given by (10.1.4), the determination of asymptotic solutions is greatly simplified by using the complex representation (10.1.7). The real and imaginary parts of (10.1.7) yield real asymptotic solutions of (10.1.1) and $v(\mathbf{x})$ is clearly an amplitude term, if it is real valued.) Inserting (10.1.7) into (10.1.1) gives

$$\nabla^2 u + k^2 n^2 u = \{k^2[n^2 - (\nabla\phi)^2]v + ik[2\nabla\phi\cdot\nabla v + v\nabla^2\phi] + \nabla^2 v\}\, e^{ik\phi} = 0. \tag{10.1.8}$$

Since we assume that $k \gg 1$, we equate the coefficient of the highest power of k in (10.1.8) to zero and obtain

$$(\nabla\phi(\mathbf{x}))^2 = n^2(\mathbf{x}), \tag{10.1.9}$$

which is known as the *eiconal equation.* It is a nonlinear first order PDE. (It was studied in Section 2.4 in the two-dimensional case.) The phase term $\phi(\mathbf{x})$ in (10.1.7) is specified from (10.1.9) subject to appropriate conditions that result from the data given for the reduced wave equation (10.1.1).

Once $\phi(\mathbf{x})$ is specified, the amplitude $v(\mathbf{x})$ must be determined from

$$ik[2\nabla\phi(\mathbf{x})\cdot\nabla v(\mathbf{x}) + v(\mathbf{x})\nabla^2\phi(\mathbf{x})] + \nabla^2 v(\mathbf{x}) = 0 \tag{10.1.10}$$

in view of (10.1.8)–(10.1.9). Dividing by ik in (10.1.10) and letting $k \to \infty$ reduces the order of (10.1.10) from second to first order. Since k is assumed to be large, we

conclude that the problem of determining $v(\mathbf{x})$ from (10.1.10) is a *singular perturbation problem*. Therefore, difficulties of the type encountered in Section 9.3 are likely to occur in the present problem.

Nevertheless, we look for an asymptotic solution of (10.1.10) in the form

$$v(\mathbf{x}) = \sum_{j=0}^{\infty} v_j(\mathbf{x})(ik)^{-j}. \tag{10.1.11}$$

(The expansion in powers of ik rather than k is carried out for convenience.) Inserting (10.1.11) into (10.1.10) and equating like powers of k yields the recursive system of equations

$$2\nabla\phi(\mathbf{x}) \cdot \nabla v_0(\mathbf{x}) + v_0(\mathbf{x})\nabla^2\phi(\mathbf{x}) = 0, \tag{10.1.12}$$

$$2\nabla\phi(\mathbf{x}) \cdot \nabla v_j(\mathbf{x}) + v_j(\mathbf{x})\nabla^2\phi(\mathbf{x}) = -\nabla^2 v_{j-1}(\mathbf{x}) \qquad j \geq 1. \tag{10.1.13}$$

These first order PDEs for the $v_j(\mathbf{x})$ are known as *transport equations* since they describe the variation of the amplitude terms $v_j(\mathbf{x})$ along the rays or characteristics determined from the eiconal equation, as shown later. The (initial) conditions for the $v_j(\mathbf{x})$ are determined from the data for (10.1.1).

Much of the terminology associated with the large k solutions of the reduced wave equation (10.1.1) is drawn from the *theory of optics*. The propagation of light waves, when described in terms of (10.1.1), corresponds to the case where k is large. The asymptotic solution (10.1.7) characterizes *geometrical optics* (or *ray optics*), since it is given in terms of geometry of the rays or characteristics determined from the *eiconal equation*. Often, geometrical optics is taken to mean the results determined strictly from the eiconal equation without regard to the transport equations (10.1.12)–(10.1.13). Solutions based directly on (10.1.1), or some approximation thereof that involves second order differential equations, yield results that characterize *wave optics*. The theory of wave optics is required when geometrical optics fails to describe the solution of (10.1.1) correctly, either because the solution breaks down in the geometrical optics description or because it fails to account for certain effects due to the diffraction of light. However, a *geometrical theory of diffraction* has been developed by *J. B. Keller* that shows how to take into account diffraction effects within the context of a geometrical description of the solution, similar to that given by geometrical optics. These asymptotic analyses of wave propagation are examined in some detail in this section.

Exact and Asymptotic Representations of the Free-Space Green's Function

We consider the free-space Green's functions for the two-dimensional reduced wave equation with a constant n. The exact result is expanded asymptotically. This yields useful information about the nature of the asymptotic solutions of (10.1.1) in general and serves to motivate the discussion that follows.

The *free-space Green's function* $K(x, y; \xi, \eta)$ for the reduced wave equation (10.1.1) in two dimensions satisfies the equation

$$\nabla^2 K(x, y; \xi, \eta) + k^2 n^2 K(x, y; \xi, \eta) = -\delta(x - \xi)\delta(y - \eta), \tag{10.1.14}$$

and the radiation condition at infinity

$$\lim_{r \to \infty} \sqrt{r} \left(\frac{\partial K(x, y; \xi, \eta)}{\partial r} - ikn K(x, y; \xi, \eta) \right) = 0, \tag{10.1.15}$$

with $r^2 = (x - \xi)^2 + (y - \eta)^2$, and n assumed to be constant. In Example 6.13 this Green's function was found to be

$$K(x, y; \xi, \eta) = \frac{i}{4} H_0^{(1)}(knr), \tag{10.1.16}$$

where $H_0^{(1)}$ is the zero-order Hankel function of the first kind.

Noting the asymptotic representation of the Hankel function in (6.7.36), the leading term in the asymptotic expansion of $K(x, y; \xi, \eta)$ with $knr \gg 1$ is

$$K(x, y; \xi, \eta) \approx \frac{i}{4} \left(\frac{2}{\pi knr} \right)^{1/2} \exp\left(iknr - \frac{i\pi}{4} \right). \tag{10.1.17}$$

Since with $r \neq 0$, $K(x, y; \xi, \eta)$ is a solution of the (homogeneous) reduced wave equation (10.1.1), we conclude that (10.1.17) is the leading term in an asymptotic solution of (10.1.1) of the form (10.1.7) and (10.1.11). We demonstrate this by constructing an asymptotic solution of (10.1.1) that depends only on r.

Let

$$K(x, y; \xi, \eta) = e^{ik\phi(r)} \sum_{j=0}^{\infty} v_j(r)(ik)^{-j} \tag{10.1.18}$$

with $r > 0$. The eiconal equation for $\phi(r)$ becomes $(\nabla\phi(\mathbf{x}))^2 = [\phi'(r)]^2 = n^2$. The solution of this equation for which the radiation condition (10.1.15) can be satisfied by the expansion (10.1.18) is clearly given as $\phi(r) = nr$. The transport equation for $v_0(r)$ is

$$2\nabla\phi \cdot \nabla v_0 + v_0 \nabla^2 \phi = 2\phi' v_0' + \left(\phi'' + \frac{\phi'}{r} \right) v_0 = 2n v_0'(r) + \frac{n}{r} v_0(r) = 0, \tag{10.1.19}$$

and its (general) solution is

$$v_0(r) = \frac{c_0}{\sqrt{r}}, \tag{10.1.20}$$

where c_0 is an arbitrary constant that may depend on k. With $\phi(r)$ and $v_0(r)$ given as above, we find that $v_0(r)e^{ik\phi(r)}$ agrees with the asymptotic form of $K(x, y; \xi, \eta)$ given in (10.1.17). However, the constant c_0 in (10.1.20) cannot be specified by the foregoing (direct) asymptotic method since it depends on the behavior of the Green's function $K(x, y; \xi, \eta)$ at the source point $r = 0$. Now both the Green's function

(10.1.16) and its asymptotic representation (10.1.17) are singular at $r = 0$. But the Green's function (10.1.16) has a logarithmic singularity at $r = 0$, while its asymptotic form (10.1.17) is algebraically singular at $r = 0$. Since the derivation of (10.1.17) from (10.1.16) required that $r > 0$, we cannot expect (10.1.17) or the equivalent result obtained from the direct asymptotic method to be valid at $r = 0$.

Thus, although the asymptotic expansion (10.1.18) generates the correct form for the Green's function $K(x, y; \xi, \eta)$ with $r > 0$, it cannot be related directly to the behavior of the Green's function at $r = 0$, so that the arbitrary constants that occur in the full expansion (10.1.18) can be specified. An indirect approach such as the boundary layer method must be used to construct a modified expansion for $K(x, y; \xi, \eta)$ valid near $r = 0$. By matching this expansion with (10.1.18), the arbitrary constants in the latter expansion can be specified. The boundary layer method in this case yields the exact solution (10.1.16), as is easily shown. However, in the case of a variable index of refraction $n(\mathbf{x})$, the boundary layer method yields useful results since the exact solution is not, in general, available.

Exact and Asymptotic Representations of the Half-Plane Green's Function

Next, we use the free-space Green's function (10.1.16) to construct a Green's function for the Dirichlet problem for the reduced wave equation (10.1.1) in the half-plane $x > 0$. The Green's function $K(x, y; \xi, \eta)$ satisfies (10.1.14)–(10.1.15) with the boundary condition

$$K(0, y; \xi, \eta) = 0. \tag{10.1.21}$$

The source point (ξ, η) lies in the right half-plane, so that $\xi > 0$. The exact solution of this problem, obtained by the *method of images*, is

$$K(x, y; \xi, \eta) = \frac{i}{4} H_0^{(1)}(knr) - \frac{i}{4} H_0^{(1)}(kn\hat{r}), \tag{10.1.22}$$

where $\hat{r}^2 = (x + \xi)^2 + (y - \eta)^2$ and $K(x, y; \xi, \eta)$ satisfies the radiation condition at infinity.

Introducing the oscillatory term $e^{-i\omega t}$, where the frequency ω is a constant, we form the function $v(x, y, t) = K(x, y; \xi, \eta)e^{-i\omega t}$. We represent the index of refraction as $n(x, y) = c_0/c(x, y)$, where c_0 is a constant reference speed (generally taken to be the speed of light), and $c(x, y)$ is the speed of wave propagation, and define the wave number k as ω/c_0. Then $v(x, y, t)$ satisfies the two-dimensional wave equation $v_{xx}(x, y, t) + v_{yy}(x, y, t) = [1/c^2(x, y)]v_{tt}(x, y, t)$ when $(x, y) \neq (\xi, \eta)$. [This result is valid for a solution $K(x, y; \xi, \eta)$ of (10.1.14) even if $n(x, y) \neq$ constant, as has been assumed above.]

We set

$$K(x, y; \xi, \eta)e^{-i\omega t} = \frac{i}{4} H_0^{(1)}(knr)e^{-i\omega t} - \frac{i}{4} H_0^{(1)}(kn\hat{r})e^{-i\omega t} \equiv u_I\, e^{-i\omega t} + u_S\, e^{-i\omega t}, \tag{10.1.23}$$

where u_I and u_S are termed the *incident* and *scattered waves*, respectively. If we expand the Green's function (10.1.22) asymptotically for $r, \hat{r} > 0$ in the right half-plane, we have

$$K(x,y;\xi,\eta) \approx \frac{i}{4}\left(\frac{2}{\pi knr}\right)^{1/2} e^{iknr - i\pi/4} - \frac{i}{4}\left(\frac{2}{\pi kn\hat{r}}\right)^{1/2} e^{ikn\hat{r} - i\pi/4}. \tag{10.1.24}$$

Introducing this expression into (10.1.23) yields the asymptotic representation of the *incident* and *scattered waves*. [That is, the asymptotic representations of $u_I(x,y)$ and $u_S(x,y)$ are the first and second terms on the right of (10.1.24), respectively.]

This terminology is employed because in the asymptotic representation of (10.1.23), the term $v_I(x,y,t) = u_I(x,y)\exp(-i\omega t)$ represents a *cylindrical wave* traveling away from the source point (ξ,η). That part of the wave that hits the boundary $x = 0$ is moving toward the boundary. Thus it is characterized as a wave *incident* upon the boundary. The cylindrical wave $v_S(x,y,t) = u_S(x,y)\exp(-i\omega t)$ is a wave traveling away from the (fictitious) reflected or image source point $(-\xi,\eta)$, and that part of the wave that intersects the boundary $x = 0$ moves away from it toward the interior of the half-plane $x > 0$. As its existence is due to the presence of the boundary, we refer to it as a *scattered* wave. (Section 2.4 discusses cylindrical waves.)

Noting these results, we formulate the asymptotic (half-plane) Green's function problem as follows. We seek a function $K(x,y;\xi,\eta)$ given as

$$K(x,y;\xi,\eta) \approx u_I(x,y) + u_S(x,y), \tag{10.1.25}$$

where $u_I(x,y)$ and $u_S(x,y)$ are asymptotic solutions of the reduced wave equation (10.1.1). The incident wave $u_I(x,y)$ is assumed to be completely specified—it is the free-space Green's function (10.1.16)–(10.1.17)—and the scattered wave $u_S(x,y)$ is to be determined. The boundary condition (10.1.21) for $K(x,y;\xi,\eta)$ implies that

$$u_S(0,y) = -u_I(0,y). \tag{10.1.26}$$

The scattered wave $u_S(x,y)$ is required to be an *outgoing wave* at the boundary $x = 0$. That is, $u_S(x,y)e^{-i\omega t}$ must represent wave motion away from $x = 0$ into the interior of the region $x > 0$. The *outgoing condition* takes the place of the radiation condition in the asymptotic formulation of this problem.

We now demonstrate that the solution of the foregoing Green's function problem, carried out to leading terms only, agrees with (10.1.24). The asymptotic form of u_I is given in (10.1.17), and u_S is expressed as $u_S(x,y) \approx V(x,y)e^{ik\phi(x,y)}$. In view of (10.1.26) we must have $\phi(0,y) = n\sqrt{\xi^2 + (y-\eta)^2}$ and $V(0,y) = -(i/4)\,(2/\pi kn)^{1/2}\,[\xi^2 + (y-\eta)^2]^{-1/2}\exp(-i\pi/4)$; that is, we equate both the phase and the amplitude terms in (10.1.26).

There are two possible solutions of the eiconal equation satisfied by the (scattered) phase term $\phi(x,y)$ with the given initial condition. They are $\phi_{\pm}(x,y) = n\sqrt{(x \pm \xi)^2 + (y-\eta)^2}$, but we reject the minus sign since it does not yield an outgoing cylindrical wave for $u_S(x,y)$. In fact, the phase is then identical to that of

the incoming wave $u_I(x, y)$. Selecting the plus sign in the phase term and defining $\hat{r}^2 = (x + \xi)^2 + (y - \eta)^2$, we specify the phase as $\phi(x, y) = n\hat{r}$ and conclude that $V(x, y) = -(i/4)\,(2/\pi kn\hat{r})^{1/2} \exp(-i\pi/4)$ on using (10.1.19)–(10.1.20) and the given boundary condition. The asymptotic result (10.1.25) agrees with (10.1.24), which was obtained from the exact solution.

Ray Equations for the Asymptotic Phase Term

We now present a general discussion of the *eiconal equation* (10.1.9) for the phase term $\phi(\mathbf{x})$ and its solution. It is a nonlinear first order PDE and it is solved using the *method of characteristics,* as developed in Section 2.4.

We consider the three-dimensional problem, but the results are easily specialized to the two-dimensional case. With $\phi = \phi(\mathbf{x})$ and $p = \phi_x$, $q = \phi_y$, and $r = \phi_z$, we set

$$F(x, y, z, \phi, p, q, r) = p^2 + q^2 + r^2 - n^2, \tag{10.1.27}$$

where $n = n(\mathbf{x})$, so that $F = 0$ corresponds to the eiconal equation. Using the results of Exercise 2.4.13, we find that the characteristic equations for x, y, z, ϕ, p, q, and r are

$$\frac{dx}{ds} = 2p, \qquad \frac{dy}{ds} = 2q, \qquad \frac{dz}{ds} = 2r, \tag{10.1.28}$$

$$\frac{d\phi}{ds} = 2(p^2 + q^2 + r^2) = 2n^2, \tag{10.1.29}$$

$$\frac{dp}{ds} = \frac{\partial(n^2)}{\partial x}, \qquad \frac{dq}{ds} = \frac{\partial(n^2)}{\partial y}, \qquad \frac{dr}{ds} = \frac{\partial(n^2)}{\partial z}, \tag{10.1.30}$$

where s is a parameter along the characteristics and the fact that $F = 0$ along the characteristic curves was used to obtain (10.1.29).

The characteristic (base) curves $\mathbf{x} = \mathbf{x}(s) = [x(s), y(s), z(s)]$ are called *rays*, and the surfaces of constant phase $\phi(\mathbf{x}) =$ constant are called *wave fronts.* Since the vector $[p, q, r] = [\phi_x, \phi_y, \phi_z] = \nabla\phi(\mathbf{x})$ is normal to the wave front $\phi(\mathbf{x}) =$ constant, the equations (10.1.28), which can be written as

$$\frac{d\mathbf{x}}{ds} = 2[p, q, r] = 2\nabla\phi(\mathbf{x}), \tag{10.1.31}$$

show that the rays are orthogonal to the wave fronts.

The full first order system of characteristic equations (10.1.28)–(10.1.30) must be solved simultaneously, particularly if the index of refraction n is a function of $\mathbf{x}$. However, it is possible to obtain a separate second order system of equations for the rays $\mathbf{x} = \mathbf{x}(s)$. On differentiating (10.1.31) with respect to s and using (10.1.30), we obtain

$$\frac{d^2\mathbf{x}}{ds^2} = 2\nabla(n^2(\mathbf{x})). \tag{10.1.32}$$

This vector equation, together with

$$\left(\frac{dx}{ds}\right)^2 + \left(\frac{dy}{ds}\right)^2 + \left(\frac{dz}{ds}\right)^2 = 4(p^2 + q^2 + r^2) = 4n^2(\mathbf{x}), \qquad (10.1.33)$$

determines the rays $\mathbf{x}(s)$ and the variation of the parameter s along the rays. These equations for $\mathbf{x}(s)$ are known as the *ray equations*. They show that the rays are specified in terms of the *index of refraction* $n(\mathbf{x})$. In particular, if the index $n(\mathbf{x})$ is a constant, $\mathbf{x}(s)$ is linear in s, so that the rays are straight lines.

Rays in a Stratified Medium

In the following we specialize the ray equations to the case of two dimensions (x, y) and assume that the index of refraction depends only on x [i.e., $n = n(x)$]. That is, we consider wave propagation in an inhomogeneous but *stratified medium* since $n = n(x)$.

The *ray equations* (10.1.32)–(10.1.33) for $x(s)$ and $y(s)$ take the form

$$\frac{d^2x}{ds^2} = \frac{d(2n^2(x))}{dx}, \qquad \frac{d^2y}{ds^2} = 0, \qquad \left(\frac{dx}{ds}\right)^2 + \left(\frac{dy}{ds}\right)^2 = 4n^2(x). \quad (10.1.34)$$

From the equation for $y(s)$ we conclude that $y'(s) = 2a$, where a is an arbitrary constant. We multiply the equation for $x(s)$ by $x'(s)$ and obtain

$$\frac{dx}{ds}\frac{d^2x}{ds^2} = \frac{d}{ds}\left\{\frac{1}{2}\left[\frac{dx}{ds}\right]^2\right\} = \frac{dx}{ds}\frac{d(2n^2(x))}{dx} = \frac{d(2n^2(x))}{ds}, \qquad (10.1.35)$$

which implies that $[x'(s)]^2 = 4n^2(x) + \text{constant}$. Making use of the third equation in (10.1.34) and the fact that $y'(s) = 2a$, we obtain $x'(s) = \pm 2(n^2(x) - a^2)^{1/2}$. But $x'(s) = x'(y)y'(s) = 2ax'(y)$, so that the ray equations reduce to

$$\frac{dy}{dx} = \pm\frac{a}{\sqrt{n^2(x) - a^2}}. \qquad (10.1.36)$$

If we specify a point (x_0, y_0) that the ray must pass through, as well as the slope of the ray at that point, the constant a and the sign in (10.1.36) are determined. The resulting ray equation has a unique solution.

In the context of *geometrical optics*, it is of interest to determine the structure of the rays in a stratified medium without determining the full field. We consider two cases.

First we assume that all rays issue from the y-axis with the same (positive) slope. Thus, we set (with a fixed θ) $y(0) = y_0$, $y'(0) = \tan(\theta)$, $0 < \theta < \pi/2$. We use the prescribed initial slope $y'(0)$ to determine the arbitrary constant a in (10.1.36) and find that $a = n(0)\sin(\theta)$. As a result, the initial value problem for the ray equation becomes [since $y'(0) > 0$]

$$\frac{dy}{dx} = \frac{n(0)\sin(\theta)}{\sqrt{n^2(x) - n^2(0)\sin^2(\theta)}}, \qquad y(0) = y_0, \qquad (10.1.37)$$

and the (formal) solution is

$$y(x) = \int_0^x \frac{n(0)\sin(\theta)ds}{\sqrt{n^2(s) - n^2(0)\sin^2(\theta)}} + y_0. \tag{10.1.38}$$

Each value of y_0 determines a ray and we consider the rays only in the right half-plane $x > 0$.

To proceed we must specify the behavior of the index of refraction $n(x)$. The simplest case occurs if $n(x)$ is monotonically increasing or nondecreasing. Then the denominator in (10.1.38) is never zero, and the solution is valid for all $x > 0$. As y_0 ranges from minus to plus infinity, the rays cover the entire right half-plane.

If $n(x)$ decreases monotonically to zero as $x \to \infty$, the slope $y'(x)$ increases monotonically until x reaches the value $x = x_c$, at which point $n(x_c) = n(0)\sin(\theta)$ and the slope $y'(x_c)$ is infinite. Each of the rays is refracted (or bent) continuously as x increases and at the point $x = x_c$, each ray has a vertical tangent line and begins to reverse its direction. Since none of the rays penetrate the region $x > x_c$, it is called the *refraction shadow region.* While the slope $y'(x)$ of each ray is infinite at $x = x_c$, the integral (10.1.38) is convergent at $x = x_c$ and $y(x_c)$ serves as the initial value for the (turned) ray with negative slope. The equation for the new set of rays is

$$y(x) = \int_0^{x_c} \frac{n(0)\sin(\theta)\,ds}{\sqrt{n^2(s) - n^2(0)\sin^2(\theta)}} - \int_{x_c}^x \frac{n(0)\sin(\theta)\,ds}{\sqrt{n^2(s) - n^2(0)\sin^2(\theta)}} + y_0. \tag{10.1.39}$$

Each ray has a negative slope for $x < x_c$ and an infinite slope at $x = x_c$.

If $n(x)$ decreases monotonically to a value $n_1 > 0$, we see that if $n_1 > n(0)\sin(\theta)$, the slope of the rays never becomes infinite and there is no shadow region. The rays cover the entire right half-plane. However, if we determine the angle θ_1 from the equation $n_1 = n(0)\sin(\theta_1)$, we find that for $\theta > \theta_1$ a shadow region will exist. Thus for decreasing $n(x)$ if the initial ray slope is sufficiently large, a refraction shadow will occur.

Next we consider an initial value problem for (10.1.36) in which all the rays issue from the origin in the (x, y)-plane. The initial values are $y(0) = 0$, $y'(0) = \tan(\theta)$, $0 < \theta < \pi$. We assume that $n(x)$ is a smooth function defined as follows. For $x \le 0$, $n(x) = n(0)$ and for $x \ge x_1 > 0$, $n(x) = n(x_1) \equiv n_1$, with $n(0) > n_1$. Thus, $n(x)$ is constant for $x < 0$ and $x > x_1$. As x increases from 0 to x_1, $n(x)$ decreases monotonically from $n(0)$ to n_1.

The rays whose initial direction angle θ satisfies $\pi/2 < \theta < \pi$ are straight lines whose equation is $y(x) = x\tan(\theta)$, $\pi/2 < \theta < \pi$. The ray that corresponds to $\theta = \pi/2$ is the line $x = 0$ with $y > 0$. For the remaining direction angles we obtain the rays in the first quadrant as

$$y(x) = \int_0^x \frac{n(0)\sin(\theta)\,ds}{\sqrt{n^2(s) - n^2(0)\sin^2(\theta)}} + y_0, \qquad 0 < \theta < \frac{\pi}{2}. \tag{10.1.40}$$

Now the rays with an initial direction angle θ for which $n(0)\sin(\theta) < n_1$ never achieve a vertical slope as x increases from zero. As these rays pass the line $x = x_1$,

their slope becomes constant and they have the equation

$$y(x) = \int_0^{x_1} \frac{n(0)\sin(\theta)\,ds}{\sqrt{n^2(s) - n^2(0)\sin^2(\theta)}} + \left[\frac{n(0)\sin(\theta)}{\sqrt{n_1^2 - n^2(0)\sin^2(\theta)}}\right](x - x_1). \tag{10.1.41}$$

The rays for which $\sin(\theta) > n_1/n(0)$ have an infinite slope at the point $x_\theta < x_1$ determined from $n(x_\theta) = n(0)\sin(\theta)$. At that point the rays turn backward. Their equation is given by (10.1.39), where we replace x_c by x_θ and put $y_0 = 0$. There is a limiting diffracted ray whose angle θ_c is determined from $\sin(\theta_c) = n_1/n(0)$. All rays with $\theta > \theta_c$ are refracted backward while those with $\theta < \theta_c$ have a finite slope for all positive x. The refraction shadow is the region above the extended limiting ray with negative slope. Since this ray turns around at $x = x_1$ where its slope is infinite, the shadow region lies to the left of the line $x = x_1$ (see Figure 10.1). We remark that each of the turned rays has a constant slope on the part of the ray that lies in the second quadrant.

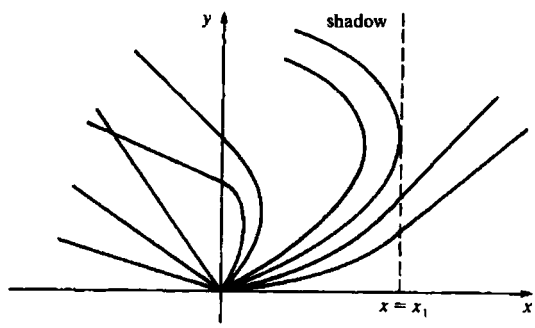

Figure 10.1 Rays in a stratified medium.

General Initial Value Problems for the Ray Equations

We reconsider the *ray equations* (10.1.32)–(10.1.33). With $d\sigma^2 = dx^2 + dy^2 + dz^2$ so that σ is an arc length parameter, (10.1.33) implies that $d\sigma = 2n\,ds$, so that s does not generally represent arc length on the rays. In terms of σ the ray equations (10.1.32) become

$$n(\mathbf{x})\frac{d}{d\sigma}\left(n(\mathbf{x})\frac{d\mathbf{x}}{d\sigma}\right) = \nabla\left(\frac{1}{2}n^2(\mathbf{x})\right), \tag{10.1.42}$$

whereas (10.1.33) is replaced by

$$\left(\frac{dx}{d\sigma}\right)^2 + \left(\frac{dy}{d\sigma}\right)^2 + \left(\frac{dz}{d\sigma}\right)^2 = 1. \tag{10.1.43}$$

Replacing the parameter s by σ and noting that $d/ds = 2n\, d/d\sigma$, we have

$$\frac{d\phi}{d\sigma} = n, \tag{10.1.44}$$

for the phase ϕ, instead of (10.1.29). Integrating along a ray $\mathbf{x}(\sigma)$ from σ_0 to σ yields

$$\phi = \phi_0 + \int_{\sigma_0}^{\sigma} n\, d\sigma, \tag{10.1.45}$$

where ϕ and n are functions of $\mathbf{x}(\sigma)$ and ϕ_0 is a constant that may vary from ray to ray. Thus once the rays are known the phase ϕ can be determined.

The (standard) *initial value problem* for the phase ϕ requires that it be specified on a given initial surface. Let the initial surface be expressed parametrically as $\mathbf{x} = \mathbf{R}(\alpha, \beta)$ and let $\phi = \phi_0(\alpha, \beta)$ on that surface. To determine ϕ we must find the rays that pass through the initial surface and then integrate along the rays as indicated in (10.1.45). However, since the ray equations are of second order, an initial point, as well as an initial direction, must be assigned for each ray on the initial surface. Now (10.1.31) shows that the direction of a ray $\mathbf{x}$ (i.e., $d\mathbf{x}/ds$ or $d\mathbf{x}/d\sigma$) is given in terms of p, q, and r. Thus, to solve for ϕ uniquely, we must prescribe not only $\mathbf{x}$ and ϕ initially, but also p, q, and r, consistent with the fact that the full set of characteristic equations (10.1.28)–(10.1.30) for $\mathbf{x}$, ϕ, p, q, and r is what we are solving.

To determine p, q, and r initially, and thereby $d\mathbf{x}/d\sigma$, we follow the procedure given in Section 2.4 and obtain conditions equivalent to the *strip conditions* found there. From the initial condition $\phi[\mathbf{R}(\alpha, \beta)] = \phi_0(\alpha, \beta)$ we obtain

$$\frac{\partial \phi_0}{\partial \alpha} = \nabla\phi \cdot \frac{\partial \mathbf{R}}{\partial \alpha} = \frac{1}{2}\frac{d\mathbf{x}}{ds} \cdot \frac{\partial \mathbf{R}}{\partial \alpha} = n\frac{d\mathbf{x}}{d\sigma} \cdot \frac{\partial \mathbf{R}}{\partial \alpha}, \tag{10.1.46}$$

and similarly,

$$\frac{\partial \phi_0}{\partial \beta} = \nabla\phi \cdot \frac{\partial \mathbf{R}}{\partial \beta} = n\frac{d\mathbf{x}}{d\sigma} \cdot \frac{\partial \mathbf{R}}{\partial \beta}, \tag{10.1.47}$$

where (10.1.31) and $d/ds = 2n\, d/d\sigma$ have been used. An additional condition, equivalent to setting $F = 0$ in (10.1.27) and identical with (10.1.43), is

$$\frac{d\mathbf{x}}{d\sigma} \cdot \frac{d\mathbf{x}}{d\sigma} = 1. \tag{10.1.48}$$

This yields three equations for the three components of the vector $d\mathbf{x}/d\sigma$ evaluated on the initial surface. We assume that $\sigma = \sigma_0$ at the point where each ray intersects the surface $\mathbf{x} = \mathbf{R}(\alpha, \beta)$, so that $d\mathbf{x}/d\sigma$ evaluated at $\sigma = \sigma_0$ is determined from (10.1.46)–(10.1.48).

To simplify the determination of the possible solutions for $d\mathbf{x}/d\sigma$ at $\sigma = \sigma_0$, we assume that α and β measure arc length in an orthogonal coordinate system on the initial surface $\mathbf{x} = \mathbf{R}(\alpha, \beta)$. Then $d\mathbf{x}/d\sigma$, $d\mathbf{R}/d\alpha$, and $d\mathbf{R}/d\beta$ are all unit vectors and $d\mathbf{R}/d\alpha$ and $d\mathbf{R}/d\beta$ are mutually orthogonal. Let θ and ω denote the angles

between the vector $d\mathbf{x}(\sigma_0)/d\sigma$ and the vectors $d\mathbf{R}/d\alpha$ and $d\mathbf{R}/d\beta$, respectively, with $0 \leq \theta,\ \omega \leq \pi$. Then (10.1.46)–(10.1.48) yield

$$\cos(\theta) = \frac{1}{n}\frac{\partial \phi_0}{\partial \alpha}, \qquad \cos(\omega) = \frac{1}{n}\frac{\partial \phi_0}{\partial \beta}. \tag{10.1.49}$$

The preceding equations generally determine exactly two directions for each ray on the initial surface. To see this, we choose an arbitrary point P on that surface and introduce a (local) Cartesian coordinate system with its origin at that point. Let the unit basis vectors $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ of the coordinate system correspond to the vectors $d\mathbf{R}/d\alpha$, $d\mathbf{R}/d\beta$, and a unit normal vector to the initial surface, respectively. Then our definition of θ and ω implies that

$$\cos(\theta) = \frac{d\mathbf{x}}{d\sigma} \cdot \mathbf{i}, \qquad \cos(\omega) = \frac{d\mathbf{x}}{d\sigma} \cdot \mathbf{j} \tag{10.1.50}$$

at the point P on the initial surface. The angles θ and ω are uniquely specified from (10.1.49). If the angle γ is defined as (with $0 \leq \gamma \leq \pi$)

$$\cos(\gamma) = \frac{d\mathbf{x}}{d\sigma} \cdot \mathbf{k}, \tag{10.1.51}$$

we conclude from (10.1.48) that

$$\cos^2(\theta) + \cos^2(\omega) + \cos^2(\gamma) = 1. \tag{10.1.52}$$

Now by definition we have $0 \leq \theta,\ \omega,\ \gamma \leq \pi$. If θ and ω are such that $\cos(\gamma) \neq 0$ in (10.1.52), the quadratic expression (10.1.52) yields two possible values of $\cos(\gamma)$, one positive and one negative, with identical absolute values. If we choose the value of γ, say $\gamma = \gamma_0$, for which $\cos(\gamma_0) > 0$, we must have $0 < \gamma_0 < \pi/2$ and the vector $d\mathbf{x}(\sigma_0)/d\sigma$ points toward the same side of the initial surface as the normal vector in view of (10.1.51). If we set $\gamma = \gamma_1 = \pi - \gamma_0$, we have $\cos(\gamma_1) = -\cos(\gamma_0) < 0$ and $\pi/2 < \gamma_1 < \pi$. In this case $d\mathbf{x}(\sigma_0)/d\sigma$ and the normal vector have opposite directions. The vectors $d\mathbf{x}(\sigma_0)/d\sigma$ that correspond to γ_0 and γ_1, respectively, and the normal vector to the initial surface are all coplanar at each point P.

If the initial data are such that $\cos(\gamma) = 0$ at a point on the initial surface, it follows from (10.1.51) that $d\mathbf{x}(\sigma_0)/d\sigma$ lies in the tangent plane of the surface at that point. If the rays are tangent to the initial surface at every point, there are two cases that are referred to as *characteristic initial value problems* for ϕ. In one case the rays are completely contained within the initial surface. Since the rays do not leave the initial surface $\mathbf{x} = \mathbf{R}(\alpha, \beta)$, the equation (10.1.44) that describes the variation of ϕ along the rays is, in fact, a condition on the variation of the initial value ϕ_0 along the rays on the initial surface. If the conditions on the data are met, we find that our result does not determine the phase function ϕ outside the initial surface.

A more interesting case occurs if the rays are merely tangent to the initial surface but do not lie within it. Then the initial surface is an envelope or a *caustic surface* of the system of rays. Again the initial value ϕ_0 must be such that (10.1.44) is satisfied

when $\sigma = \sigma_0$, that is, at the point of tangency of the rays. One ray issues from each point on the initial surface, and ϕ is determined from (10.1.45) along each ray. A more complicated situation that occurs often in asymptotic problems is when the rays are tangent to the initial surface only along a curve or some other subregion but are not tangent elsewhere on the surface. Although it is possible to determine the rays and the phase function ϕ for many of these problems, the fact that the rays intersect at the initial surface leads to difficulties with the full asymptotic solution.

In the noncharacteristic case [i.e., when $\cos(\gamma) \neq 0$], two possible ray directions are determined at each point on the initial surface. To obtain a unique solution for the phase ϕ, one of the two directions determined from (10.1.49)–(10.1.52) must be chosen. We select the set of rays issuing either on one side of the initial surface or on the other side. Thereby, $\mathbf{x}(\sigma)$ and $d\mathbf{x}/d\sigma$ are uniquely specified at each point of the initial surface, and the ray equations determine $\mathbf{x}(\sigma)$ uniquely. Finally, the phase $\phi(x, y, z)$ is found from (10.1.45) with integration taken along the ray that passes through the point (x, y, z).

Transport Equations: Rays and Wave Fronts

The *transport equations* (10.1.12)–(10.1.13) can be solved by integrating along the rays determined by solving the eiconal equation. We have from (10.1.31)

$$2\nabla\phi \cdot \nabla v_j = \frac{d\mathbf{x}}{ds} \cdot \nabla v_j = 2n \frac{d\mathbf{x}}{d\sigma} \cdot \nabla v_j = 2n \frac{dv_j}{d\sigma}, \tag{10.1.53}$$

where $dv_j/d\sigma$ is directional derivative along the ray $\mathbf{x}(\sigma)$. Thus the transport equations reduce to ODEs along the rays and are given as

$$2n \frac{dv_0}{d\sigma} + v_0 \nabla^2 \phi = 0, \tag{10.1.54}$$

$$2n \frac{dv_j}{d\sigma} + v_j \nabla^2 \phi = -\nabla^2 v_{j-1}, \qquad j \geq 1, \tag{10.1.55}$$

where all functions are evaluated on the ray $\mathbf{x}(\sigma)$.

To solve these equations, we express $\nabla^2\phi$ in terms of its variation along a ray. We consider a region R bounded laterally by a *tube of rays* and capped by two segments of *wave fronts* $\phi =$ constant, S_0 and S_1, as shown in Figure 10.2. The lateral boundary is denoted by S. Let $\mathbf{N}$ be the exterior unit normal on the boundary of R. Applying the divergence theorem in R gives

$$\iiint_R \nabla^2\phi \, dV = \iiint_R \nabla \cdot \nabla\phi \, dV = \int_{S_0} \nabla\phi \cdot \mathbf{N} \, da + \int_S \nabla\phi \cdot \mathbf{N} \, da + \int_{S_1} \nabla\phi \cdot \mathbf{N} \, da. \tag{10.1.56}$$

T he surface integral over S vanishes since $\nabla\phi$ has the direction of the rays and $\mathbf{N}$ is orthogonal to that direction on S. Assuming that the parameter σ increases as we move from S_0 to S_1, we observe that $\nabla\phi$ has the direction of $\mathbf{N}$ on S_1, while it has the

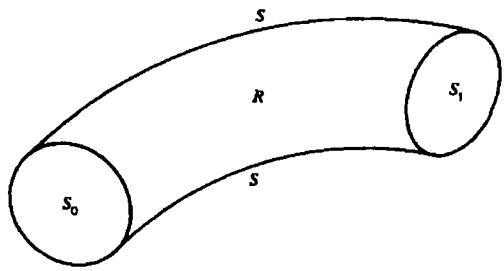

Figure 10.2 A ray tube.

direction of $-\mathbf{N}$ on S_0, since S_0 and S_1 are portions of phase surfaces. Also, since $|\nabla\phi| = n$, in view of (10.1.9), we have

$$\iint_R \nabla^2\phi \, dV = \int_{S_1} n \, da - \int_{S_0} n \, da. \tag{10.1.57}$$

To determine a simple expression for the element of surface area da on a wave front, we introduce a reference wave front $\phi = 0$ and a coordinate system (ξ, η) on that wave front. We then determine all the rays orthogonal to that wave front. Each surface ϕ = constant is parallel to the wave front $\phi = 0$ and is orthogonal to the rays. With σ as arc length on the rays, we may introduce a (ray) coordinate system (ξ, η, σ) in space. Further, we can replace σ by the wave front coordinate ϕ, and in view of (10.1.44), we have

$$d\phi = n \, d\sigma \tag{10.1.58}$$

along the rays. In the (ξ, η, ϕ) coordinate system, the surfaces ϕ = constant are orthogonal to the rays, which are given as ξ = constant and η = constant. Consequently, if $\mathbf{x} = \mathbf{r}(\xi, \eta, \phi)$ are the equations of transformation from Cartesian coordinates to (ξ, η, ϕ) coordinates, we find that the area element da on a wave front is given as $da = J \, d\xi \, d\eta$, where $J = |\mathbf{r}_\xi \times \mathbf{r}_\eta|$, as is well known from vector analysis. The volume element dV in (ξ, η, ϕ) coordinates has the form $dV = da \, d\sigma = (1/n) \, J \, d\xi \, d\eta \, d\phi$, since σ is arc length and the distance between two infinitesimally close wave fronts ϕ = constant and $\phi + d\phi$ = constant is given by $d\sigma = (1/n) \, d\phi$, in view of (10.1.58).

We now assume that the surface elements S_0 and S_1 correspond to the wave fronts ϕ = constant and $\phi + d\phi$ = constant, respectively. Then as the region R (i.e., the ray tube) shrinks down on a particular ray, we have in (ξ, η, ϕ) coordinates

$$\iint_R \nabla^2\phi \, dV \approx \nabla^2\phi \, \frac{1}{n} \, J \, d\xi \, d\eta \, d\phi, \tag{10.1.59}$$

$$\int_{S_1} n \, da - \int_{S_0} n \, da \approx \left[nJ\big|_{\phi+d\phi} - nJ\big|_{\phi} \right] d\xi \, d\eta, \tag{10.1.60}$$

so that

$$\nabla^2\phi \approx \frac{n}{J} \frac{nJ|_{\phi+d\phi} - nJ|_{\phi}}{d\phi}. \tag{10.1.61}$$

In the limit as $d\phi \to 0$ we obtain, on using (10.1.58),

$$\nabla^2\phi = \frac{n}{J}\frac{d}{d\phi}(nJ) = \frac{1}{J}\frac{d}{d\sigma}(nJ). \tag{10.1.62}$$

We are now ready to solve the transport equations (10.1.54)–(10.1.55). Inserting (10.1.62) into (10.1.54) yields

$$2n\frac{dv_0}{d\sigma} + \frac{1}{J}\frac{d}{d\sigma}(nJ)\,v_0 = 0. \tag{10.1.63}$$

This can be written as

$$\frac{2n}{\sqrt{nJ}}\frac{d(v_0\sqrt{nJ})}{d\sigma} = 0. \tag{10.1.64}$$

Integrating along the ray $\mathbf{x}(\sigma)$ from σ_0 to σ, with $v_0|_{\sigma_0} = V_0(\sigma_0)$ gives

$$v_0(\sigma) = V_0(\sigma_0)\left[\frac{n(\sigma_0)J(\sigma_0)}{n(\sigma)J(\sigma)}\right]^{1/2}, \tag{10.1.65}$$

where the dependence of the solution on variables other than σ has been suppressed. Similarly, the transport equations (10.1.55) can be written as

$$\frac{2n}{\sqrt{nJ}}\frac{d(v_j\sqrt{nJ})}{d\sigma} = -\nabla^2 v_{j-1}, \qquad j \geq 1. \tag{10.1.66}$$

Again integrating along a ray and putting $v_j|_{\sigma_0} = V_j(\sigma_0)$, we have

$$v_j(\sigma) = V_j(\sigma_0)\left[\frac{n(\sigma_0)J(\sigma_0)}{n(\sigma)J(\sigma)}\right]^{1/2} - \frac{1}{2\sqrt{n(\sigma)J(\sigma)}}\int_{\sigma_0}^{\sigma}\sqrt{\frac{J}{n}}\nabla^2 v_{j-1}\,d\sigma, \quad j \geq 1. \tag{10.1.67}$$

Although the derivation was carried out for the three-dimensional case, the results are valid in two dimensions as well, with J defined appropriately. On the reference wave front $\phi = 0$ we now introduce the single coordinate ξ and determine all the rays orthogonal to that wave front. In the resulting (ξ, ϕ) coordinate system we set $\mathbf{x} = \mathbf{r}(\xi, \phi)$ and find that the linear element ds on a wave front is given as $ds = J\,d\xi$, where $J = |\mathbf{r}_\xi|$. The area element dA then has the form $dA = ds\,d\sigma = (1/n)J\,d\xi\,d\phi$, since (10.1.44) is valid in two dimensions as well. In deriving (10.1.61) the integration is carried out over a plane region bounded by two wave-fronts and two rays. The expression for $\nabla^2\phi$ is the same as was found in the three dimensional problem, as is the case for the expressions for the phase function ϕ and the amplitude terms v_j.

In both cases J represents an *expansion* or *contraction factor* for the rays. It is proportional to an area or linear element on the wave fronts. It increases or decreases in magnitude according as the rays diverge or converge. Noting the form of v_0 in (10.1.65), we find that, to leading order in k, the amplitude of the solution increases as the rays converge and decreases as the rays diverge. This result is consistent with a basic principle of geometrical optics according to which the field is stronger the greater the concentration of the rays.

If the rays intersect at same point or curve, J vanishes and the amplitude terms v_j all blow up. The solution of the reduced wave equation of which we have found the asymptotic expansion either remains finite in that region or has a different type of singularity, as seen in our discussion of Green's functions above. The asymptotic result is not valid in regions where the rays intersect and modified (asymptotic) results must be found in such regions.

The dependence of v_0 on the divergence or convergence of the rays may be obtained by means of different method. Multiplying (10.1.12) by v_0 gives

$$v_0 [2\nabla\phi \cdot \nabla v_0 + v_0 \nabla^2 \phi] = \nabla \cdot (v_0^2 \nabla\phi) = 0. \tag{10.1.68}$$

Integrating (10.1.68) over the ray tube and using the divergence theorem gives

$$0 = \iint_R \nabla \cdot (v_0^2 \nabla\phi)\, dV = \int_{S_1} v_0^2 n\, da - \int_{S_0} v_0^2 n\, da, \tag{10.1.69}$$

as is easily seen on applying the results that lead from (10.1.56) to (10.1.57). In an approximate sense, (10.1.69) may be thought to represent a *conservation of energy* along the ray tubes of the solution of the reduced wave equation. If the ray tube shrinks down on a ray $\mathbf{x}(\sigma)$, we have

$$v_0^2 n\, da\big|_{\sigma_1} = v_0^2 n\, da\big|_{\sigma_0}, \tag{10.1.70}$$

where σ_1 and σ_0 correspond to two points on the ray. Noting relation $da = J\, d\xi\, d\eta$ between da and J, we find that v_0 at an arbitrary point on a ray is given in terms of $v_0|_{\sigma_0} = V(\sigma_0)$ precisely as in (10.1.65).

Specific Ray Systems and Wave Fronts

If the index of refraction n is a constant, some simple explicit forms for the function J and thereby for the amplitude term v_0 can be found. The rays are straight lines and we may assume without loss of generality that $n = 1$. Then, as shown by (10.1.44), both σ and ϕ measure arc length on the rays.

The case of *plane wave fronts* arises if the rays are *parallel lines*. Then the area element da remains constant on all wave fronts, so that J = constant as well. This implies that v_0 is constant along the rays. If v_0 retains the same constant value on all rays, the amplitude terms v_j, with $j \geq 1$, may be chosen to vanish, and an exact *plane wave* solution of the reduced wave equation results. (Since the wave fronts are orthogonal to the rays, it is clear that they must be planes.) If v_0 assumes different values on different rays, we obtain what is called a *general plane wave*. The wave fronts ϕ = constant are again planes, but the amplitude term v_0, although it is constant on each ray, may vary in magnitude from point to point on the wave fronts. A similar result applies in the two-dimensional problem.

To obtain *spherical wave fronts*, we introduce spherical coordinates

$$x = r\sin(\theta)\cos(\omega), \qquad y = r\sin(\theta)\sin(\omega), \qquad z = r\cos(\theta), \tag{10.1.71}$$

where $r \geq 0$, $0 \leq \theta \leq \pi$, and $0 \leq \omega \leq 2\pi$. With $n = 1$, we can put $\phi = r - r_1$, where r_1 is a constant and the reference wave front $\phi = 0$ is the sphere $r = r_1$. The ray equation (10.1.42) becomes $d^2\mathbf{x}/d\sigma^2 = \mathbf{0}$ and has the solution $\mathbf{x} = \mathbf{a} + \mathbf{b}\sigma$, where $\mathbf{a}$ and $\mathbf{b}$ are constant vectors and $|\mathbf{b}| = 1$ in view of (10.1.43). Replacing the variable σ by r and introducing the initial data

$$\mathbf{x}(r_1) = \mathbf{a} + \mathbf{b}\, r_1 = [r_1 \sin(\theta)\cos(\omega),\ r_1 \sin(\theta)\sin(\omega),\ r_1 \cos(\theta)], \quad (10.1.72)$$

$$\frac{d\mathbf{x}(r_1)}{dr} = \mathbf{b} = [\sin(\theta)\cos(\omega),\ \sin(\theta)\sin(\omega),\ \cos(\theta)], \quad (10.1.73)$$

we find that

$$\mathbf{x}(r) = \mathbf{b}\, r = [r \sin(\theta)\cos(\omega),\ r \sin(\theta)\sin(\omega),\ r \cos(\theta)]. \quad (10.1.74)$$

The data (10.1.72)–(10.1.73) imply that the rays $\mathbf{x}(r)$ pass through and are orthogonal to the sphere $r = r_1$, and (10.1.74) shows the rays to be radial lines issuing from the origin. The radial parameter r clearly measures arc length on the rays, and the spheres ϕ = constant are orthogonal to the rays.

To determine the function J for this case, we identify the wave front coordinates ξ and η on $\phi = 0$, with the angular variables θ and ω. Since $n = 1$, both ϕ and r measure arc length on the rays. The transformation from Cartesian to ray coordinates is $\mathbf{x} = \mathbf{r}(r, \theta, \omega) = [r \sin(\theta)\cos(\omega),\ r \sin(\theta)\sin(\omega),\ r \cos(\theta)]$, as seen from (10.1.71). Then da is an element of area on a sphere, and we have $da = r^2 \sin(\theta)\, d\theta\, d\omega = J\, d\theta\, d\omega$, where $J = |\mathbf{r}_\theta \times \mathbf{r}_\omega| = r^2 \sin(\theta)$. Thus

$$v_0 = V_0(r_0)\left[\frac{J(r_0))}{J(r)}\right]^{1/2} = V_0(r_0)\,\frac{r_0}{r} \quad (10.1.75)$$

on each ray. Assuming that $V_0(r_0)$ is independent of θ and ω and has the same value on each ray, the v_j $(j \geq 1)$ may be chosen to vanish and the solution of the reduced wave equation is the *spherical wave* $u = (c/r)e^{ikr}$, where c is an arbitrary constant. The asymptotic approach yields an exact solution of the reduced wave equation valid up to the singular point $r = 0$. That point is a *focus* for the rays since they all intersect there. Although the (asymptotic) solution has a valid form up to $r = 0$, it does blow up there since $J = 0$.

More generally, the term $V_0(r_0)$ in (10.1.75) may vary from ray to ray; that is, it has a nonconstant dependence on θ and ω. Then the wave fronts are still spheres, and v_0 decays like $1/r$ along a given ray, but a more complicated expression results for the amplitude of the entire field that involves all the v_j.

If we replace the spherical coordinates (10.1.71) by the cylindrical coordinates $x = r\sin(\theta)$, $y = r\cos(\theta)$, $z = z$ and proceed as before, we obtain *cylindrical wave fronts*. The rays are the radial lines of the polar coordinate system $x = r\sin(\theta)$, $y = r\cos(\theta)$ and the wave fronts are circular cylinders. Then $J = r$ and v_0 is given as $V_0(r_0)[r_0/r]^{1/2}$ on setting $n = 1$. In contrast to the case of spherical waves, even if we assume that $V_0(r_0)$ is independent of θ, the leading term v_0 in the

asymptotic expansion of the solution does not yield an exact solution of the reduced wave equation. This was demonstrated when an asymptotic *cylindrical wave* solution of the reduced wave equation was found previously from the free-space Green's function. Also, the result is not valid up to the singular point, as was the case for the spherical wave. The exact and asymptotic results are both singular at the focal point of the rays, but they have different types of singularities there, as was shown.

Next we consider a two-dimensional example in which the rays form a (smooth) envelope. If $\mathbf{x} = \mathbf{R}(\xi)$ is the equation of this envelope or *caustic curve* with ξ as an arc length parameter, the rays are just the tangent lines of this curve. If σ represents arc length along the rays, an orthogonal coordinate system in ξ and σ is given by the equations of transformation

$$\mathbf{x} = \mathbf{r}(\xi, \sigma) = \mathbf{R}(\xi) + (\sigma - \xi)\mathbf{R}'(\xi). \tag{10.1.76}$$

Since ξ is arc length, we have $|\mathbf{R}'(\xi)| = 1$, $\mathbf{R}'(\xi) \cdot \mathbf{R}''(\xi) = 0$, and $|\mathbf{R}''(\xi)| = 1/\rho(\xi)$, where $\rho(\xi)$ is the radius of curvature of the caustic curve. Then $\mathbf{r}_\xi = (\sigma - \xi)\mathbf{R}''(\xi)$ and $\mathbf{r}_\sigma = \mathbf{R}'(\xi)$ imply that $\mathbf{r}_\xi \cdot \mathbf{r}_\sigma = 0$, so that the (ξ, σ) coordinate system is orthogonal. The lines ξ = constant are the rays, while the curves σ = constant are the (orthogonal) wave fronts. This may be verified directly by noting that with

$$h_\xi = |\mathbf{r}_\xi| = \frac{|\sigma - \xi|}{\rho}, \qquad h_\sigma = |\mathbf{r}_\sigma| = 1, \tag{10.1.77}$$

we have

$$(\nabla\phi)^2 = \frac{1}{h_\xi^2}\,\phi_\xi^2 + \frac{1}{h_\sigma^2}\,\phi_\sigma^2 = \frac{\rho^2}{(\sigma - \xi)^2}\,\phi_\xi^2 + \phi_\sigma^2. \tag{10.1.78}$$

Thus $\phi = \sigma$ is a solution of the eiconal equation $(\nabla\phi)^2 = 1$. The function J for this ray system is $J = |\mathbf{r}_\xi| = |\sigma - \xi|/\rho$, and the solution v_0 has the form

$$v_0(\sigma) = V_0(\sigma_0)\left[\frac{J(\sigma_0)}{J(\sigma)}\right]^{1/2} = V_0(\sigma_0)\left|\frac{\sigma_0 - \xi}{\sigma - \xi}\right|^{1/2}. \tag{10.1.79}$$

The amplitude term v_0 is singular when $\sigma = \xi$, and this occurs at points on the caustic $\mathbf{x} = \mathbf{R}(\xi)$, as follows from (10.1.76). The rays intersect at the caustic so that $J = 0$ on that curve. Each member of the family of rays (for which $-\infty < \sigma < \infty$) intersects the caustic at $\sigma = \xi$. By considering the set of points $\sigma \geq \xi$ and $\sigma \leq \xi$ on each ray, we determine two groups of rays, each of which begins or ends at the caustic. The asymptotic solutions given for ϕ and v_0, combined with the results that can be obtained for all the v_j, are valid for each group of rays away from the caustic curve. On the caustic curve these asymptotic expansions blow up, while the corresponding solutions of the reduced wave equation are bounded there. Rather than reject the asymptotic approach in a neighborhood of the caustic, we present a *boundary layer approach* below that yields a finite asymptotic result at the caustic.

Boundary Value Problems for the Reduced Wave Equation

A typical *boundary value problem* for the *reduced wave equation* (10.1.1), which is to be solved asymptotically for large k, is formulated in the manner indicated in our discussion of Green's functions above. An *incident wave* $u_I(\mathbf{x})$ given as an asymptotic series of the form (10.1.7) and (10.1.11) approaches a boundary region and is scattered by it. The resulting *scattered wave* $u_S(\mathbf{x})$ is again an asymptotic series of the type (10.1.7) and (10.1.11) and must be determined. The *total field* $u(\mathbf{x}) = u_I(\mathbf{x}) + u_S(\mathbf{x})$, an asymptotic solution of (10.1.1), satisfies a *homogeneous boundary condition* of the first, second, third, or mixed kind on the boundary region. We mostly assume that $u(\mathbf{x}) = 0$ on the boundary in our discussion. The problem is to be solved in an exterior unbounded region and we require that $u_S(\mathbf{x})$ satisfy a *radiation condition at infinity*. The asymptotic version of the problem replaces the radiation condition with an *outgoing condition* at the boundary. If $\phi(\mathbf{x})$ is the phase term of the asymptotic expansion of $u_S(\mathbf{x})$, then $u_S(\mathbf{x})$ is said to be outgoing at the boundary if the exterior normal derivative of $\phi(\mathbf{x})$ [i.e., $\partial\phi(\mathbf{x})/\partial n = \nabla\phi(\mathbf{x}) \cdot \mathbf{N}$] is positive for every exterior unit normal $\mathbf{N}$ to the boundary. This means that $u_S(\mathbf{x})$ travels away from the boundary and is radiating toward infinity.

The specification of the incident wave $u_I(\mathbf{x})$, the boundary condition, and the outgoing condition uniquely specify $u_S(\mathbf{x})$ and thereby the asymptotic solution $u(\mathbf{x})$, provided that the phase and the amplitude terms in the asymptotic expansion of $u_S(\mathbf{x})$ can be determined. We assume that $u_I(\mathbf{x})$ is given as

$$u_I(\mathbf{x}) = e^{ik\psi(\mathbf{x})} \sum_{j=0}^{\infty} w_j(\mathbf{x})(ik)^{-j}, \tag{10.1.80}$$

where the phase $\psi(\mathbf{x})$ and the amplitude terms $w_j(\mathbf{x})$ are known. The scattered wave is expanded as

$$u_S(\mathbf{x}) = e^{ik\phi(\mathbf{x})} \sum_{j=0}^{\infty} v_j(\mathbf{x})(ik)^{-j}, \tag{10.1.81}$$

with $\phi(\mathbf{x})$ and the $v_j(\mathbf{x})$ to be determined. We have

$$e^{ik\phi(\mathbf{x})} \sum_{j=0}^{\infty} v_j(\mathbf{x})(ik)^{-j} = -e^{ik\psi(\mathbf{x})} \sum_{j=0}^{\infty} w_j(\mathbf{x})(ik)^{-j}, \tag{10.1.82}$$

on the boundary, if $u(\mathbf{x}) = u_I(\mathbf{x}) + u_S(\mathbf{x}) = 0$ there. This yields

$$\phi(\mathbf{x}) = \psi(\mathbf{x}), \qquad v_j(\mathbf{x}) = -w_j(\mathbf{x}), \qquad j \geq 0, \tag{10.1.83}$$

on the boundary.

As we have seen, the (initial) condition (10.1.83) for $\phi(\mathbf{x})$ is by itself insufficient to specify the phase uniquely. In determining the rays associated with the phase $\phi(\mathbf{x})$, we must also specify a direction for these rays on the (boundary) region where $\phi(\mathbf{x})$ is given. In general, there are two possible ray directions at each point where $\phi(\mathbf{x})$

is specified. Now, $\nabla\phi(\mathbf{x})$ has the direction of the rays as shown by (10.1.31), and it has been shown that if the rays are not tangent to the surface where $\phi(\mathbf{x})$ is specified, their two possible directions lie on opposite sides of the surface. Thus the outgoing condition on $u_S(\mathbf{x})$ uniquely specifies a ray direction at each point, as is easily seen.

Once the rays and the phase are determined, the amplitude terms $v_j(\mathbf{x})$ can be obtained by integrating along the rays as in (10.1.54)–(10.1.55) and using the data (10.1.83) as initial conditions.

Reflection of a Cylindrical Wave by a Parabola

We now consider the problem of the reflection of a cylindrical wave by a parabolic cylinder and construct its asymptotic solution employing the foregoing method. The line source of the (incident) cylindrical wave is assumed to lie on the z-axis, and the generators of the parabolic cylinder are parallel to the z-axis. It is assumed that index of refraction $n = 1$. Thus the problem may be treated as two-dimensional and we look for a solution in the form $u = u(x, y)$.

The equation of the parabola is given as

$$x = \frac{1}{2a}\, y^2 - \frac{a}{2}, \tag{10.1.84}$$

where $a > 0$ is a given constant. A *cylindrical wave* is generated by a source located at the origin, which is also the focus of the parabola. Consequently, the total field $u(x, y)$ satisfies the equation

$$\nabla^2 u(x, y) + k^2 u(x, y) = -\,\delta(x)\delta(y), \tag{10.1.85}$$

where $\delta(x)\delta(y)$ is the two-dimensional Dirac delta function. We assume that $u(x, y)$ vanishes on the parabola [i.e., $u\left(y^2/2a - a/2, y\right) = 0$], and that $u(x, y)$ satisfies a radiation condition at infinity. An outwardly radiating wave generated by the source is $u_I(x, y) = (i/4)H_0^{(1)}(kr)$, where $r^2 = x^2 + y^2$, as was shown in the beginning of this section. We are concerned with the asymptotic problem for large k. Thus we replace the foregoing by its asymptotic expansion. The two leading terms are given as

$$u_I(x, y) \approx \frac{i}{4}\sqrt{\frac{2}{\pi kr}}\left(1 + \frac{1}{8ikr}\right)\exp\left(ikr - \frac{i\pi}{4}\right) \tag{10.1.86}$$

on using the large argument asymptotic expansion of $H_0^{(1)}(kr)$.

The wave $u_I(x, y)$ is a cylindrical wave incident on the parabola. The asymptotic problem then looks for a solution $u(x, y) \approx u_I(x, y) + u_S(x, y)$, where the incident wave $u_I(x, y)$ is given by (10.1.86) and the scattered wave $u_S(x, y)$, expanded as in (10.1.81), is to be determined. The boundary condition requires that $u_S\left(y^2/2a - a/2, y\right) = -u_I\left(y^2/2a - a/2, y\right)$. The radiation condition implies that $u_S(x, y)$ is an outgoing wave at the parabola.

Comparing with the notation used in our discussion of the general problem yields as the phase and leading amplitude terms for the incident wave $u_I(x, y)$,

$$\psi(x, y) = r, \quad w_0(x, y) = \frac{i}{4}\sqrt{\frac{2}{\pi k}}\,\frac{e^{-i\pi/4}}{r^{1/2}},\ w_1(x, y) = \frac{i}{32}\sqrt{\frac{2}{\pi k}}\,\frac{e^{-i\pi/4}}{r^{3/2}}. \tag{10.1.87}$$

The other terms in the asymptotic expansion of $u_I(x, y)$ have not been specified in (10.1.86). The appropriate data for the scattered or reflected wave $u_S(x, y)$ on the parabola are that its phase term $\phi(x, y)$ satisfy

$$\phi(x, y) = r \qquad \text{on} \qquad x = \frac{1}{2a}\,y^2 - \frac{a}{2}, \tag{10.1.88}$$

and that $\phi(x, y)$ be *outgoing* on the parabola. The first two amplitude terms $v_0(x, y)$ and $v_1(x, y)$ of $u_S(x, y)$ must satisfy

$$v_0(x, y) = -\frac{i}{4}\sqrt{\frac{2}{\pi k}}\,\frac{1}{r^{1/2}}e^{-i\pi/4},\ v_1(x, y) = -\frac{i}{32}\sqrt{\frac{2}{\pi k}}\,\frac{1}{r^{3/2}}e^{-i\pi/4} \tag{10.1.89}$$

on the parabola, in view of (10.1.87).

To solve for the phase $\phi(x, y)$ of the scattered wave $u_S(x, y)$, we first represent the parabola in parametric form as

$$y = \tau, \qquad x = \frac{1}{2a}\,(\tau^2 - a^2), \qquad -\infty < \tau < \infty. \tag{10.1.90}$$

(We note that τ is not an arc length parameter here.) With σ as arc length along the rays we must solve the characteristic equations for $x(\sigma, \tau)$, $y(\sigma, \tau)$, $\phi(\sigma, \tau)$, $p(\sigma, \tau)$, and $q(\sigma, \tau)$. The initial conditions are given at $\sigma = 0$. Thus

$$y(0, \tau) = \tau, \qquad x(0, \tau) = \frac{1}{2a}\,(\tau^2 - a^2), \tag{10.1.91}$$

as follows from (10.1.90). That is, $\sigma = 0$ corresponds to the parabola (10.1.90). The condition (10.1.88) for ϕ yields

$$\phi(0, \tau) = \frac{1}{2a}\,(\tau^2 + a^2), \tag{10.1.92}$$

on expressing (10.1.90) in terms of the variable τ.

The initial conditions for $p(\sigma, \tau)$ and $q(\sigma, \tau)$ are determined by proceeding as in Section 2.4. The eiconal equation requires that $p^2(0, \tau) + q^2(0, \tau) = 1$, and the *strip condition* states that $\partial\phi(0, \tau)/\partial\tau = p(0, \tau)\ \partial x(0, \tau)/\partial\tau + q(0, \tau)\ \partial y(0, \tau)/\partial\tau$ on the parabola. This yields $\tau/a = p(0, \tau)\ \tau/a + q(0, \tau)$ on using (10.1.91)–(10.1.92). Solving for $p(0, \tau)$ and $q(0, \tau)$ gives

$$p(0, \tau) = 1, \qquad q(0, \tau) = 0, \tag{10.1.93}$$

$$p(0, \tau) = \frac{\tau^2 - a^2}{\tau^2 + a^2}, \qquad q(0, \tau) = \frac{2a\tau}{\tau^2 + a^2}. \tag{10.1.94}$$

One of these two sets of initial conditions for $p(\sigma, \tau)$ and $q(\sigma, \tau)$ must be rejected on the basis of the outgoing condition satisfied by $u_S(x, y)$. For this problem, this means that if $\mathbf{N}$ is the unit normal of the parabola that points towards the interior region of the parabola (which contains the focus), we must have $\nabla\phi \cdot \mathbf{N} > 0$ on the parabola. We easily find that the unit normal $\mathbf{N}$ is $\mathbf{N}(\tau) = [a, -\tau]/\sqrt{\tau^2 + a^2}$. Since $\nabla\phi = [p(\sigma, \tau),\ q(\sigma, \tau)]$, we see that on the parabola $\nabla\phi \cdot \mathbf{N} = a/\sqrt{\tau^2 + a^2} > 0$ if (10.1.93) is used, whereas $\nabla\phi \cdot \mathbf{N} = -a/\sqrt{\tau^2 + a^2} < 0$ if (10.1.94) is used. Therefore, we conclude that (10.1.93) are the appropriate initial conditions for the problem.

The characteristic equations (10.1.28)–(10.1.30) for this problem are

$$\frac{dx}{d\sigma} = p, \qquad \frac{dy}{d\sigma} = q, \qquad \frac{d\phi}{d\sigma} = 1, \qquad \frac{dp}{d\sigma} = 0, \qquad \frac{dq}{d\sigma} = 0, \tag{10.1.95}$$

since $n = 1$ and $d\sigma = 2\,ds$. The last two equations in (10.1.95) imply that $p(\sigma, \tau)$ and $q(\sigma, \tau)$ are constant along the characteristics. Thus they are equal to their initial values (10.1.93) and $p(\sigma, \tau) = 1$, $q(\sigma, \tau) = 0$. Inserting these results into the equations for x and y in (10.1.95) and using (10.1.91) gives

$$x(\sigma, \tau) = \sigma + \frac{1}{2a}(\tau^2 - a^2), \qquad y(\sigma, \tau) = \tau. \tag{10.1.96}$$

Finally, we obtain for ϕ, in view of (10.1.95) and (10.1.92), the result

$$\phi(\sigma, \tau) = \sigma + \frac{1}{2a}(\tau^2 + a^2). \tag{10.1.97}$$

We invert the system (10.1.96) and express σ and τ as functions of x and y, to obtain $\sigma = x - (1/2a)(y^2 - a^2)$, $\tau = y$. Inserting this into (10.1.97) yields

$$\phi(x, y) = x + a. \tag{10.1.98}$$

Thus $u_S(x, y)$ has the phase of a plane wave. Since the amplitude terms of $u_S(x, y)$ are not constant, as we will show, $u_S(x, y)$ is a *general plane wave.*

To solve for $v_0(x, y)$ and $v_1(x, y)$ we first note that $\nabla^2\phi(x, y) = 0$ in view of (10.1.98). Further, $r = (1/2a)(\tau^2 + a^2)$ on the parabola. Thus $2\nabla\phi \cdot \nabla v_0 + \nabla^2\phi\, v_0 = 2\,\partial v_0/\partial\sigma = 0$. Thus, $v_0(\sigma, \tau) = v_0(0, \tau)$ and $v_0(\sigma, \tau) = -(i/4)\sqrt{2/\pi k}\ \sqrt{2a/(\tau^2 + a^2)}\, e^{-i\pi/4}$ on using (10.1.89). In terms of x and y,

$$v_0(x, y) = -\frac{i}{4}\sqrt{\frac{2}{\pi k}}\sqrt{\frac{2a}{y^2 + a^2}}\, e^{-i\pi/4} \tag{10.1.99}$$

since $\tau = y$. The equation for $v_1(\sigma, \tau)$ is $2\,\partial v_1(\sigma, \tau)/\partial\sigma = -\nabla^2 v_0(\sigma, \tau)$, where the right side is a function of τ only, in view of (10.1.99). Expressed in terms of the (x, y) variables, $v_1(x, y)$ takes the form $v_1(x, y) = V_1(y)x + V_2(y)$, with given $V_1(y)$ and $V_2(y)$, which we do not exhibit.

We have shown that $u_S(x,y)$ is given as

$$u_S(x,y) \approx \left\{ -\frac{i}{4}\sqrt{\frac{2}{\pi k}}\sqrt{\frac{2a}{y^2+a^2}}\, e^{-i\pi/4} + \frac{1}{ik}(V_1(y)x + V_2(y)) \right\} \exp[ik(x+a)], \tag{10.1.100}$$

and this has the form of a *general plane wave*. Two terms have been calculated to show that the asymptotic result (10.1.100) is not uniformly valid. As x increases to the point where $x = O(k)$, it follows from (10.1.100) that the term $(1/ik)v_1(x,y)$ is of the same order of magnitude as $v_0(x,y)$. [We note that $1/\sqrt{k}$ is a factor of the full expression for $u_S(x,y)$.] Therefore, the asymptotic expansion of $u_S(x,y)$ becomes disordered, in that $(1/ik)v_1(x,y)$ is not of lower order in k than $v_0(x,y)$. The asymptotic expansion we have obtained for $u_S(x,y)$ is, consequently, useful only if $x < O(k)$.

The growth of the amplitude term $v_1(x,y)$, and that of further terms in the expansion of $u_S(x,y)$, corresponds to the *secular behavior* encountered and discussed in a number of problems in the Chapter 9. To analyze this question more closely, we insert $u_S(x,y) = v(x,y)\exp[ik(x+a)]$ into the reduced wave equation (10.1.1) with $n = 1$ and find that

$$2ikv_x(x,y) + v_{xx}(x,y) + v_{yy}(x,y) = 0. \tag{10.1.101}$$

Inserting $v(x,y) = \sum_{j=0}^{\infty} v_j(x,y)(ik)^{-j}$ into (10.1.101) and solving for the $v_j(x,y)$ shows that $v_j(x,y) = O(x^j)$ as $x \to \infty$, as has been found for $v_1(x,y)$.

The secularity problems encountered may be removed by using a *stretching transformation* similar to that used in the *boundary layer theory* of Section 9.3. Since our asymptotic results must be modified only for $x \geq O(k)$, we set $x = k\xi$ in (10.1.101) to emphasize the region where x is large. This yields

$$2iv_\xi(\xi,y) + v_{yy}(\xi,y) + \frac{1}{k^2}v_{\xi\xi}(\xi,y) = 0. \tag{10.1.102}$$

To leading order in k [i.e., with $v(\xi,y) \approx \hat{v}(\xi,y)$] we obtain a parabolic *Schrödinger equation*

$$2i\hat{v}_\xi(\xi,y) + \hat{v}_{yy}(\xi,y) = 0. \tag{10.1.103}$$

We must find a solution of (10.1.103) that matches the asymptotic result given above for $v(x,y)$, to leading order terms, as $\xi = x/k \to 0$. Effectively, we are considering an initial value problem for (10.1.103) in the region $\xi > 0$ with $\hat{v}(\xi,y)$ specified at a positive value of ξ.

A general solution of (10.1.103) is

$$\hat{v}(\xi,y) = \int_{-\infty}^{\infty} \left\{ \frac{1}{\sqrt{\xi}} \exp\left[\frac{i(y-s)^2}{2\xi} \right] \right\} f(s)\, ds. \tag{10.1.104}$$

The bracketed term in (10.1.104) is, apart from a constant factor, a fundamental solution of (10.1.103). By direct substitution we can verify that it satisfies (10.1.103)

for $\xi > 0$. $f(s)$ in (10.1.104) is arbitrary and must be specified by matching (10.1.104) with the foregoing asymptotic results.

To carry out the *matching procedure*, we rewrite (10.1.104) in terms of $x = k\xi$ and y to obtain

$$v(x,y) = \sqrt{\frac{k}{x}} \int_{-\infty}^{\infty} \exp\left[\frac{ik(y-s)^2}{2x}\right] f(s)\, ds. \tag{10.1.105}$$

Now if $x \leq O(k)$, we see that the exponential in (10.1.105) is rapidly oscillating since k is large by assumption. Therefore, we can evaluate (10.1.105) by the *method of stationary phase*, as presented in Section 5.7. The phase term $\phi(s) = (y-s)^2/2x$ in (10.1.105) has a stationary point where $\phi'(s_0) = 0$, that is, at $s_0 = y$. Then the leading order result as given in (5.7.4) is

$$v(x,y) \approx \sqrt{2\pi} \exp\left(\frac{i\pi}{4}\right) f(y). \tag{10.1.106}$$

This result, is valid for small $\xi = x/k$, must be equated to the leading amplitude term $v_0(x,y)$ given in (10.1.99). Thus $f(y) = -(1/4\pi)\sqrt{2a/(y^2+a^2)}$ and $v(x,y)$ [i.e., (10.1.105)] is completely specified.

We might expect that it should be possible to generate the amplitude term $v_1(x,y)$ from the asymptotic expansion of the integral (10.1.105). This is indeed possible, but we do not carry out this calculation.

Asymptotic Expansion at a Caustic

The foregoing discussion has shown that the geometrical optics results, in which the scattered and incident fields are expanded in the form (10.1.7) and (10.1.11), are not always adequate. In regions where the rays intersect or in the far field, the geometrical optics solution is not valid and a modified expansion or result is needed. Next we show how to construct a valid asymptotic result near a two-dimensional caustic along which the rays intersect and the geometrical optics result fails. The boundary layer method is used.

We consider a two-dimensional problem for the reduced wave equation (10.1.1) with $n = 1$ and an asymptotic solution of the form (10.1.7) and (10.1.11). The rays for this solution are assumed to have a smooth envelope. The envelope is a *caustic curve* for this problem, and the asymptotic solution breaks down on the caustic, as we have seen. We retain the notation introduced in our discussion of this problem [see (10.1.76)-(10.1.79)].

The equation of the caustic is assumed to be $\mathbf{x} = \mathbf{R}(\xi)$, where ξ is arc length on the curve. The rays are the tangent lines to the caustic and are given as

$$\mathbf{x} = \mathbf{r}(\xi,\sigma) = \mathbf{R}(\xi) + (\sigma - \xi)\mathbf{R}'(\xi) \tag{10.1.107}$$

with σ as arc length on the rays, as shown following (10.1.76). The phase is $\phi = \sigma$, and in view of (10.1.79), the leading term in the asymptotic result is

$$u \approx V_0(\sigma_0)\left(\frac{\xi - \sigma_0}{\xi - \sigma}\right)^{1/2} e^{ik\sigma}. \tag{10.1.108}$$

From (10.1.107) we see that when $\sigma = \xi$, the rays intersect the caustic and (10.1.108) shows that the asymptotic result is infinite on the caustic. We assume that σ and σ_0 in (10.1.108) are less than ξ and that $V_0(\sigma_0)$ is a prescribed function. Thus (10.1.108) represents a wave approaching the caustic $\mathbf{x} = \mathbf{R}(\xi)$, whose amplitude becomes infinite there.

Although the asymptotic field is singular at the caustic, the actual field is finite there. After the *incoming rays* of the wave that approaches the caustic pass the caustic, they become *outgoing rays*. The resulting outgoing field is well behaved, and we must determine its geometrical optics representation. In addition, no rays penetrate to the concave side of the caustic as seen in Figure 10.3. As a result, the geometrical optics field must vanish in that region. The actual field is, in fact, nonzero there. We wish to study the field at and near the caustic and to determine the transition undergone by the geometrical optics field on its passage through the caustic region, that is, how its amplitude and phase terms change.

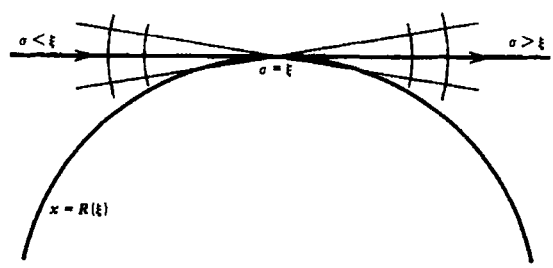

Figure 10.3 The caustic curve.

To study the field near the caustic $\mathbf{x} = \mathbf{R}(\xi)$, we express the reduced wave equation in (ξ, σ) coordinates as defined by (10.1.107). The coordinate system is orthogonal and we have from (10.1.77) $h_\xi = (\xi - \sigma)\rho(\xi)$, $h_\sigma = 1$, where $\rho = \rho(\xi)$ is the radius of curvature of the caustic. We are at present only considering values of σ for which $\sigma \leq \xi$, since we would otherwise have a double covering of the region above the caustic curve. That is, since the full tangent lines cover the region twice, we restrict ourselves to half-lines.

The reduced wave equation (10.1.1) with $n = 1$ takes the form

$$\frac{1}{\xi - \sigma}\frac{\partial}{\partial \sigma}\left[(\xi - \sigma)\frac{\partial u(\xi, \sigma)}{\partial \sigma}\right] + \frac{\rho(\xi)}{\xi - \sigma}\frac{\partial}{\partial \xi}\left[\frac{\rho(\xi)}{\xi - \sigma}\frac{\partial u(\xi, \sigma)}{\partial \xi}\right] + k^2 u(\xi, \sigma) = 0. \tag{10.1.109}$$

Let

$$u(\xi,\sigma) = w(\xi,\sigma)e^{ik\sigma} \tag{10.1.110}$$

and insert (10.1.110) into (10.1.109). We retain the phase term of the geometrical optics solution (10.1.108) in the expression (10.1.110) since it is the amplitude term of (10.1.108) that blows up at the caustic. $w(\xi,\sigma)$ satisfies

$$ik\left(2\frac{\partial w}{\partial\sigma} - \frac{1}{\xi-\sigma}w\right) + \frac{1}{\xi-\sigma}\frac{\partial}{\partial\sigma}\left[(\xi-\sigma)\frac{\partial w}{\partial\sigma}\right] + \frac{\rho(\xi)}{\xi-\sigma}\frac{\partial}{\partial\xi}\left[\frac{\rho(\xi)}{\xi-\sigma}\frac{\partial w}{\partial\xi}\right] = 0. \tag{10.1.111}$$

To emphasize the neighborhood of the caustic we set

$$\xi - \sigma = k^{-r}\eta, \qquad \xi = \lambda, \tag{10.1.112}$$

with the stretching exponent $r > 0$ to be determined. Since $k \gg 1$, the stretching transformation (10.1.112) implies that small values of $\xi - \sigma$ are to be considered. Now $\partial/\partial\xi = \partial/\partial\lambda + k^r\partial/\partial\eta$ and $\partial/\partial\sigma = -k^r\partial/\partial\eta$, so that (10.1.111) becomes

$$-ik^{1+r}\left[2\frac{\partial w(\eta,\lambda)}{\partial\eta} + \frac{1}{\eta}w(\eta,\lambda)\right] + \frac{k^{2r}}{\eta}\frac{\partial}{\partial\eta}\left[\eta\frac{\partial w(\eta,\lambda)}{\partial\eta}\right]$$
$$+\frac{k^r\rho(\lambda)}{\eta}\left(\frac{\partial}{\partial\lambda} + k^r\frac{\partial}{\partial\eta}\right)\left[\frac{k^r\rho(\lambda)}{\eta}\left(\frac{\partial w(\eta,\lambda)}{\partial\lambda} + k^r\frac{\partial w(\eta,\lambda)}{\partial\eta}\right)\right] = 0. \tag{10.1.113}$$

In assessing the significance of the terms in (10.1.113), we note that the coefficient of k^{1+r} should be retained, since in the absence of the stretching transformation it is this term that yields the geometrical optics amplitude results in (10.1.108). In addition, it is the absence of second derivative terms in the transport equations that invalidates the geometrical optics result at the caustic. Therefore, we must retain at least one second derivative term in (10.1.113). In effect, we have a *singular perturbation problem* in (10.1.111) and it is the neglect of the higher derivative terms that causes difficulties in the *geometrical optics expansion*. The highest power of k in the second derivative terms in (10.1.113) comes from k^{4r}. Thus we choose r so that k^{1+r} and k^{4r} are equally significant. The yields $1 + r = 4r$, with the result that $r = \frac{1}{3}$.

Accordingly, we set $r = \frac{1}{3}$ in (10.1.113) and expand $w(\eta,\lambda)$ in a series in powers of $k^{-1/3}$. With $w_0(\eta,\lambda)$ as the leading term in the series, we find that $w_0(\eta,\lambda)$ satisfies the *boundary layer equation*

$$-i\left[2\frac{\partial w_0(\eta,\lambda)}{\partial\eta} + \frac{1}{\eta}w_0(\eta,\lambda)\right] + \frac{\rho^2(\lambda)}{\eta}\frac{\partial}{\partial\eta}\left[\frac{1}{\eta}\frac{\partial w_0(\eta,\lambda)}{\partial\eta}\right] = 0. \tag{10.1.114}$$

This is an ODE for $w_0(\eta,\lambda)$ as a function of η and can be simplified by setting $w_0(\eta,\lambda) = \exp(i\eta^3/3\rho^2(\lambda))W(z)$, where

$$z = \frac{\eta^2}{(4\rho^4(\lambda))^{1/3}}, \tag{10.1.115}$$

and $W(z)$ satisfies the equation

$$W''(z) + zW(z) = 0. \tag{10.1.116}$$

With z replaced by $-z$, (10.1.116) becomes the *Airy equation*, which has two linearly independent solutions $Ai(z)$ and $Bi(z)$, some of whose properties we now present. The Airy function $Ai(z)$ decays exponentially and $Bi(z)$ grows exponentially as $z \to +\infty$. Both functions are oscillatory for $z < 0$. The asymptotic behavior of the Airy function $Ai(z)$ for complex-valued z expressed as $z = |z|e^{i\theta}$, with $|z| \geq 0$ and $-\pi \leq \theta \leq \pi$, is given as

$$Ai(z) \approx \frac{1}{2\pi}(z)^{-1/4} \exp\left(-\frac{2}{3}z^{3/2}\right), \qquad |z| \to \infty, \quad |\theta| < \pi, \tag{10.1.117}$$

and as $z \to -\infty$ (this corresponds to $\theta = \pm\pi$), we have

$$Ai(z) \approx \frac{1}{\sqrt{\pi}}(-z)^{-1/4} \sin\left(\frac{2}{3}(-z)^{3/2} + \frac{\pi}{4}\right), \qquad z \to -\infty. \tag{10.1.118}$$

From the functional relations

$$Ai(z) = -\omega Ai(\omega z) - \omega^2 Ai(\omega^2 z), \; Bi(z) = i\omega Ai(\omega z) - i\omega^2 Ai(\omega^2 z), \tag{10.1.119}$$

where $\omega = \exp(-2\pi i/3)$, we can determine the asymptotic behavior of the solutions $Bi(z)$, $Ai(\omega z)$, and $Ai(\omega^2 z)$ of the Airy equation in terms of the results (10.1.117)–(10.1.118) for $Ai(z)$.

As a result, a general solution of the equation (10.1.116) for $W(z)$ is

$$W(z) = c_1(\xi)Ai(-z) + c_2(\xi)Bi(-z), \tag{10.1.120}$$

where the functions $c_1(\xi)$ and $c_2(\xi)$ must be determined by matching the boundary layer result $w_0 e^{ik\sigma}$ with the geometrical optics result (10.1.108). It is easily seen from (10.1.117) that $Ai(-\omega z)$ has the asymptotic behavior

$$Ai(-\omega z) = Ai(ze^{\pi i/3}) \approx \frac{1}{2\sqrt{\pi}}(z)^{-1/4} \exp\left[-i\left(\frac{2}{3}z^{3/2} + \frac{\pi}{12}\right)\right], \; z \to \infty, \tag{10.1.121}$$

and it follows from (10.1.119) that $Ai(-\omega z) = (1/2i\omega)\,[Bi(-z) - iAi(-z)]$. Putting $c_1 = -ic_2$ in (10.1.120), we express $W(z)$ as $W(z) = c(\xi)Ai(-\omega z)$, where $c(\xi)$ is to be specified. As we now demonstrate, this choice for $W(z)$ is appropriate for matching the boundary layer and geometrical optics solution.

The boundary layer solution $u_0(\xi, \sigma) = w_0(\xi, \sigma)e^{ik\sigma}$ takes the form

$$u_0(\xi, \sigma) = c(\xi)Ai\left[\left(\frac{k^2 e^{i\pi}}{4\rho^4(\xi)}\right)^{1/2}(\xi - \sigma)^2\right] \exp\left[ik\left(\sigma + \frac{(\xi - \sigma)^3}{3\rho^2(\xi)}\right)\right]. \tag{10.1.122}$$

To match (10.1.122) with the geometrical optics result (10.1.108), we expand the Airy function asymptotically for $\xi - \sigma > 0$ and k large. In view of (10.1.121), we obtain

$$u_0(\xi,\sigma) \approx c(\xi)\,\frac{\rho(\xi)^{1/3}}{\sqrt{\pi}k^{1/6}2^{5/6}(\xi-\sigma)^{1/2}}\exp\left(ik\sigma - \frac{i\pi}{12}\right). \tag{10.1.123}$$

Comparing with (10.1.108) shows that $c(\xi)$ must be chosen as

$$c(\xi) = k^{1/6}\sqrt{\pi}\,2^{5/6}\rho(\xi)^{-1/3}V_0(\sigma_0)(\xi-\sigma_0)^{1/2}\exp\left(\frac{i\pi}{12}\right), \tag{10.1.124}$$

and the boundary layer term (10.1.122) is completely specified.

We find that the field is not infinite at the caustic as predicted by the geometrical optics result. Instead, as seen from (10.1.122) and (10.1.124), it has a finite value at the caustic, that is, when $\xi = \sigma$. However, the amplitude of the geometrical optics term is $O(1)$ in k away from the caustic, while at and near the caustic, the boundary layer term has an amplitude that is $O(k^{1/6})$. Since $k \gg 1$, we do find a growth in the amplitude of the field at the caustic as is to be expected, but it does not become infinite there.

While the singularity at the caustic has been eliminated, the foregoing result is not completely satisfactory, for the following reasons. First, it predicts that the field below the caustic (i.e., in the region not penetrated by the rays) grows without bound as we show. The field below the caustic is expected to be weaker than the geometrical optics field, and it should decay rather than grow below the caustic. Second, it is not apparent from our asymptotic solution what role is played by the outgoing geometrical optics field that results after the incident wave passes through the caustic and how to determine that field. Consequently, we modify our result so as to remove these difficulties.

To begin, we show that the boundary layer solution (10.1.122) grows without bound below the caustic. To do so, an approximate expression for the distance along the *normal lines* to the caustic valid in the boundary layer region is found. The wave fronts associated with the system of rays (10.1.107) are the curves σ = constant. Since the wave fronts are orthogonal to the rays and the rays are tangent to the caustic, we see that the wave fronts are orthogonal to the caustic. Now, distance along the wave front σ = constant is given as $ds = |\xi - \sigma|/\rho(\xi)\,d\xi$. The boundary layer region is determined by $\xi = \sigma + O(k^{-1/3})$ in view of (10.1.112). On integrating the expression for ds on the wave front σ = constant in the boundary layer region, we obtain

$$s \approx \frac{1}{2\rho(\xi)}(\xi-\sigma)^2 = k^{-2/3}\,\frac{\eta^2}{2\rho(\xi)} = \left(\frac{\rho(\xi)}{2}\right)^{1/3} k^{-2/3}z, \tag{10.1.125}$$

where (10.1.112) and (10.1.115) are used. Within the boundary layer, the variable s yields a measure of distance along the normal line to the caustic. Positive values of s correspond to points above the caustic (i.e., in the ray region), and negative values of s correspond to points below the caustic. Using (10.1.125), we express the boundary

layer solution (10.1.122) in terms of the variables ξ and s. To leading order we have

$$u_0 \approx c(\xi) Ai\left[\exp\left(\frac{i\pi}{3}\right)\left(\frac{2}{\rho(\xi)}\right)^{1/3} k^{2/3} s\right] e^{ik\xi}, \tag{10.1.126}$$

where we have set $\sigma = \xi + O(k^{-1/3})$ in the boundary layer. If (10.1.126) is evaluated for negative values of s in the region below the caustic, we find that $k^{2/3}s$ becomes large and negative. Then (10.1.117) shows that $|u_0|$ grows exponentially as $k^{2/3}s \to -\infty$.

This difficulty may be resolved by considering the behavior of not only the incident field, but also of the outgoing field near the caustic. The rays for both the incident and outgoing fields are given by (10.1.107). On a ray of the incident field we have $\sigma < \xi$, whereas on a ray of the outgoing field we have $\sigma > \xi$. At the caustic both rays meet and $\sigma = \xi$.

The leading term of the geometrical optics form of the incident field u is given in (10.1.108). We denote the outgoing field by $\hat{u}$. Since the ray structure is the same for both fields, the leading term in the geometrical optics representation of $\hat{u}$ must have the form

$$\hat{u} \approx \hat{V}_0(\xi)(\sigma - \xi)^{-1/2} e^{ik\sigma}. \tag{10.1.127}$$

The difference between (10.1.127) and (10.1.108) lies in the fact that $\sigma > \xi$ in (10.1.127), while $\sigma < \xi$ in (10.1.108). Further, the term $\hat{V}_0(\xi)$ is as yet unspecified. Both (10.1.127) and (10.1.108) are singular on the caustic.

The total field at a point (x, y) above the caustic is a sum of the incident and outgoing fields and is given as $u + \hat{u}$. We must note that even though the incident and outgoing fields are expressed in terms of the variables ξ and σ, at each point (x, y) these fields are evaluated at two different sets of values of ξ and σ. The determination of these values is indicated in Figure 10.4.

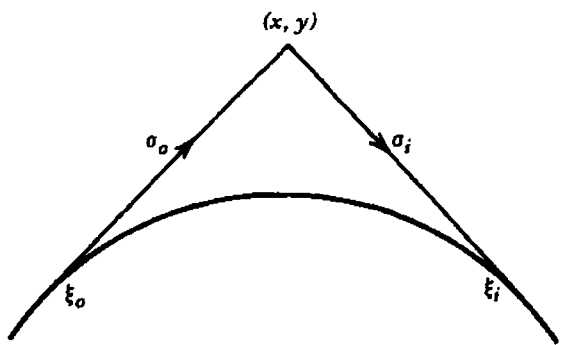

Figure 10.4 The incoming and outgoing rays.

To determine $\hat{V}_0(\xi)$ in (10.1.127) and to resolve the difficulty concerning the behavior of the field below the caustic, we consider the behavior of the total field $u + \hat{u}$

near the caustic. Now the boundary layer form of the incident field u was already given in (10.1.123)–(10.1.124). To study the boundary layer behavior of $\hat{u}$, we may carry over the preceding boundary layer results directly. The only difference is that the magnification element h_ξ has the form $h_\xi = (\sigma - \xi)/\rho(\xi)$ if $\sigma > \xi$, and this replaces the expression given previously.

However, (10.1.109) shows that the form of the reduced wave equation in the (ξ, σ) variables is unchanged if $\xi - \sigma$ is replaced by $\sigma - \xi$. Thus every result obtained up to (10.1.116) remains valid for the outgoing field. Therefore, if we define the leading order boundary layer term to be $\hat{w}_0$, we find that

$$\hat{w}_0 = \exp\left(\frac{i\eta^3}{3\rho^2(\xi)}\right) W(z), \tag{10.1.128}$$

where η, z and $W(z)$ are given as in (10.1.112), (10.1.115), and (10.1.116), respectively.

Noting the asymptotic result (10.1.117), we find that

$$Ai(-\omega^2 z) = Ai\left(ze^{-\pi i/3}\right) \approx \frac{1}{2\sqrt{\pi}}\,(z)^{-1/4}\,\exp\left[i\left(\frac{2}{3}z^{3/2} + \frac{\pi}{12}\right)\right],\ z \to \infty. \tag{10.1.129}$$

With $W(z) = \hat{c}(\xi)Ai(-\omega^2 z) = \hat{c}(\xi)Ai\left(ze^{-\pi i/3}\right)$, we obtain for $\hat{u}_0 = \hat{w}_0 e^{ik\sigma}$,

$$\hat{u}_0(\xi,\sigma) = \hat{c}(\xi)Ai\left[\left(\frac{k^2 e^{-\pi i}}{4\rho^4(\xi)}\right)^{1/2}(\xi-\sigma)^2\right]\exp\left[ik\left(\sigma + \frac{(\xi-\sigma)^3}{3\rho^2(\xi)}\right)\right]. \tag{10.1.130}$$

To specify $\hat{c}(\xi)$ we expand the Airy function in (10.1.130) asymptotically for $\sigma - \xi > 0$ and k large. Since $\xi - \sigma$ occurs in the Airy function in squared form we must use the asymptotic form (10.1.129), valid for large positive z. Thereby, the matching between the boundary layer result $\hat{u}_0$ and the geometrical optics result $\hat{u}$ is carried out.

The variable z that occurs in the above is defined in (10.1.115) and we have $(2/3)z^{3/2} = (k/3\rho^2(\xi))|\xi - \sigma|^3 = -(k/3\rho^2(\xi))(\xi - \sigma)^3$, since $\xi - \sigma < 0$, so that $|\xi - \sigma| = -(\xi - \sigma)$. This result shows why it is appropriate to choose the Airy function $Ai(-\omega^2 z)$ for this case (i.e., the outgoing field), for the asymptotic form of $\hat{u}_0(\xi, \sigma)$ is then found to be

$$\hat{u}_0(\xi,\sigma) \approx \hat{c}(\xi)\,\frac{\rho(\xi)^{1/3}}{\sqrt{\pi}k^{1/6}2^{5/6}(\sigma-\xi)^{1/2}}\exp\left(ik\sigma + \frac{i\pi}{12}\right). \tag{10.1.131}$$

Comparing (10.1.131) with the geometrical optics result (10.1.127) yields

$$\hat{c}(\xi) = k^{1/6}\sqrt{\pi}\,2^{5/6}\rho(\xi)^{-1/3}\hat{V}_0(\xi)\exp\left(-\frac{i\pi}{12}\right). \tag{10.1.132}$$

Although (10.1.132) expresses $\hat{c}(\xi)$ in terms of $\hat{V}_0(\xi)$, we still do not know the form of $\hat{V}_0(\xi)$. We expect that $\hat{V}_0(\xi)$ must be expressed in terms of the (given)

incident field. To establish this relationship, we examine the total field $u_0 + \hat{u}_0$ (to leading order) in the boundary layer region. To do so, it is not convenient to use the (ξ, σ) variables, since they have different values for u_0 and $\hat{u}_0$, as we have indicated. Instead, we use the variables ξ and s, where s measures distance along the normal from the caustic in the boundary layer region. To the level of approximation used, it follows from (10.1.125) and the discussion preceding it that the incident and outgoing fields are both evaluated at the same point (ξ, s). Thus (to leading order) the total boundary layer field $u_0 + \hat{u}_0$ is given as,

$$u_0 + \hat{u}_0 \approx k^{1/6}\sqrt{\pi}\, 2^{5/6}\rho(\xi)^{-1/3} e^{ik\xi} \tag{10.1.133}$$

$$\times \left\{ \frac{V_0(\sigma_0)e^{i\pi/12}}{(\xi-\sigma_0)^{-1/2}} Ai\left[\left(\frac{2k^2 e^{i\pi}}{\rho(\xi)}\right)^{1/3} s\right] + \hat{V}_0(\xi) e^{-i\pi/12} Ai\left[\left(\frac{2k^2 e^{-i\pi}}{\rho(\xi)}\right)^{1/3} s\right] \right\}.$$

We require that $u_0 + \hat{u}_0$ decay in the region below the caustic, that is, when $s < 0$ in (10.1.133). In view of the asymptotic result (10.1.117), this implies that the sum of the Airy functions in the bracketed term of (10.1.133) adds up to a multiple of $Ai[-(2/\rho(\xi))^{1/3}k^{2/3}s]$, for $Ai(-z)$ decays exponentially as $z \to -\infty$, whereas $Ai(-\omega z)$, $Ai(-\omega^2 z)$, and $Bi(-z)$ all grow exponentially as $z \to -\infty$. Using (10.1.119) with z replaced by $-z$ and noting that $Ai(-\omega z) = Ai[\exp(i\pi/3)z]$ and $Ai(-\omega^2 z) = Ai[\exp(-i\pi/3)z]$, we easily conclude that $\hat{V}_0(\xi)$ must be chosen such that $\hat{V}_0(\xi)\exp(-i\pi/12) = \omega V_0(\sigma_0)(\xi-\sigma_0)^{1/2}\exp(i\pi/12)$. Then (10.1.133) becomes

$$u_0+\hat{u}_0 \approx k^{1/6}\sqrt{\pi}\, e^{-i\pi/4} 2^{5/6}\rho(\xi)^{-1/3} V_0(\sigma_0)(\xi-\sigma_0)^{1/2} Ai\left[-\left(\frac{2k^2}{\rho(\xi)}\right)^{1/3} s\right] e^{ik\xi}. \tag{10.1.134}$$

and the field decays below the caustic when $k^{2/3}s$ gets large and negative.

Requiring that the field must decay below the caustic has enabled us to determine $\hat{V}_0(\xi)$. Inserting the result for $\hat{V}_0(\xi)$ into (10.1.127) yields the *outgoing geometrical optics field*

$$\hat{u} \approx V_0(\sigma_0)\left(\frac{\xi-\sigma_0}{\sigma-\xi}\right)^{1/2} e^{ik\sigma - i\pi/2} \tag{10.1.135}$$

with $\sigma > \xi$. To leading order, the incident and outgoing geometrical optics field have an amplitude term $|\sigma-\xi|^{-1/2}$, which is a consequence of the convergence and divergence of the rays as the wave approaches and leaves the caustic. However, on comparing the phase terms in (10.1.108) and (10.1.135), we find that the incident wave undergoes a *phase shift* by an amount of $\pi/2$ as it passes through the caustic and becomes the outgoing wave.

We have seen that according to geometrical optics, the field in the region below the caustic must be zero since no rays penetrate into that region. However, the boundary layer result showed that there is a nonzero field below the caustic that decays exponentially. In the theory of optics when there are nonzero fields in regions where geometrical optics predicts that the field must be zero, we attribute these effects

to the process of *diffraction of light.* In the following we consider scattering problems in which diffraction effects play a significant role.

Scattering by a Half-Plane

We assume that a *plane wave* $u_I(x,y,z) = e^{ikz}$ is incident on the *half-plane* $z = 0,\ y \le 0$. The resulting field $u(x,y,z) = u_I(x,y,z)+u_S(x,y,z)$, where $u_S(x,y,z)$ is the *scattered wave*, satisfies the *reduced wave equation* (10.1.1) with $n = 1$. The *total field* $u(x,y,z)$ is required to vanish on the half-plane. Thus we have

$$\nabla^2 u(x,y,z) + k^2 u(x,y,z) = 0,\ u(x,y,0) = 0,\ y \le 0. \tag{10.1.136}$$

Proceeding as before, we set $u_S(x,y,z) = e^{ik\phi(x,y,z)} \sum_{j=0}^{\infty} v_j(x,y,z)(ik)^{-j}$. On the boundary surface $z = 0,\ y \le 0$, the initial condition for $\phi(x,y,z)$ is

$$\phi(x,y,0) = 0, \qquad y \le 0, \tag{10.1.137}$$

and the initial conditions for the $v_j(x,y,z)$ are

$$v_0(x,y,0) = -1,\ y \le 0, \quad v_j(x,y,0) = 0,\ y \le 0,\ j \ge 1. \tag{10.1.138}$$

These conditions follow from the fact that if we express the incident field in the form $u_I(x,y,z) = e^{ik\psi(x,y,z)} \sum_{j=0}^{\infty} w_j(x,y,z)(ik)^{-j}$, as given in (10.1.80), the incident phase term $\psi(x,y,z) = z$ vanishes at $z = 0$, and the leading order amplitude term $w_0(x,y,z) = 1$ is the only nonzero term in the (asymptotic) expansion of the incident field $u_I(x,y,z) = e^{ilkz}$. We remark that this problem could be treated as being strictly two-dimensional in terms of the y,z variables, but we prefer to deal with it in three-dimensional form.

To determine the phase term $\phi(x,y,z)$ of $u_S(x,y,z)$, $\phi(x,y,z)$ must satisfy the outgoing condition at the boundary $z = 0$ in addition to (10.1.137). Clearly, the two solutions of the eiconal equation $(\nabla\phi(x,y,z))^2 = 1$ that vanish at $z = 0$ are $\phi_+(x,y,z) = z$ and $\phi_-(x,y,z) = -z$. Requiring that $u_S(x,y,z)$ be an outgoing wave at the boundary determines the appropriate choice for $\phi(x,y,z)$. To apply this condition we must distinguish between the two sides of the half-plane $z = 0,\ y \le 0$. Accordingly, we express the scattered wave $u_S(x,y,z)$ in the form $u_S(x,y,z) = v_\pm(x,y,z)e^{ik\phi_\pm(x,y,z)}$. On the side of the half-plane facing the half-space $z < 0$, we have $\phi = \phi_-(x,y,z) = -z$, since the wave $u_S(x,y,z) = v_-(x,y,z)e^{-ikz}$ travels away from the half-plane into the half-space $z < 0$. On the opposite side of the half-plane facing the region $z > 0$, we must set $\phi = \phi_+(x,y,z) = z$, since $u_S(x,y,z) = v_+(x,y,z)e^{ikz}$ travels toward the half-space $z > 0$ away from the half-plane $z = 0$.

Since the scattered field in the region $z < 0,\ y < 0$ is a plane wave, we immediately conclude from the initial conditions (10.1.138) for the amplitude terms that $v_{0-}(x,y,z) = v_{0-}(x,y,0) = -1$ and $v_{j-}(x,y,z) = v_{j-}(x,y,0) = 0$ for all $j \ge 1$. Consequently, the scattered wave in this region is given as

$$u_S(x,y,z) = -e^{-ikz}, \qquad z < 0,\ y \le 0. \tag{10.1.139}$$

In the region $z > 0,\ y \leq 0$, $u_S(x, y, z)$ is also a plane wave and (10.1.138) implies that $v_{0+}(x, y, z) = -1$ and $v_{j+}(x, y, z) = 0$ for $j \geq 1$. Thus the scattered wave in this region is given as

$$u_S(x, y, z) = -e^{ikz}, \qquad z > 0,\ y \leq 0. \tag{10.1.140}$$

There is no boundary surface in the half-space $y > 0$, and the scattered waves (10.1.139)–(10.1.140) do not penetrate this region since the rays are orthogonal to the half-plane $z = 0,\ y \leq 0$. Thus the total field $u(x, y, z) = u_I(x, y, z) + u_S(x, y, z)$ predicted by the *geometrical optics* approach is

$$u(x, y, z) = \begin{cases} e^{ikz}, & |z| < \infty,\ y > 0, \\ e^{ikz} - e^{-ikz}, & z < 0, \quad y < 0, \\ 0, & z > 0, \quad y < 0. \end{cases} \tag{10.1.141}$$

The plane $y = 0$ is a surface of discontinuity for the geometrical optics solution (10.1.141). Even though each of the representations of $u(x, y, z)$ given in (10.1.141) is an exact solution of the reduced wave equation (10.1.136) in the interior of the three indicated regions, the full expression for $u(x, y, z)$ is not a regular solution of (10.1.136). The half-plane $y = 0,\ z > 0$ is called a *shadow boundary* since it separates the (shadow) region $z > 0,\ y < 0$, where no (incident) rays penetrate, from the region $z > 0,\ y > 0$, where the incident field acts. In a similar fashion we refer to the half-plane $y = 0,\ z < 0$ as a *reflection boundary*. It separates the region where the incident and reflected field are observed from the region where there is only the incident field.

To obtain a *smooth transition* for the asymptotic solution of the scattering problem across the discontinuity boundaries, we use *boundary layer theory*. We consider the boundaries at $y = 0$ with $z < 0$ and $z > 0$ separately.

At the *shadow boundary* $y = 0,\ z > 0$, the field on one side is $u(x, y, z) = 0$, and on the other side it is $u(x, y, z) = e^{ikz}$. Thus we set $u(x, y, z) = w(x, y, z)e^{ikz}$ in (10.1.136) and obtain

$$2ikw_z(x, y, z) + w_{xx}(x, y, z) + w_{yy}(x, y, z) + w_{zz}(x, y, z) = 0. \tag{10.1.142}$$

To emphasize the neighborhood of $y = 0$, we introduce the stretching transformation $y = k^{-r}\eta$, with $r > 0$ to be determined. In the x, η, z variables, (10.1.142) becomes [in terms of the transformed dependent variable $\hat{w}(x, \eta, z)$]

$$2ik\hat{w}_z(x, \eta, z) + k^{2r}\hat{w}_{\eta\eta}(x, \eta, z) + \hat{w}_{xx}(x, \eta, z) + \hat{w}_{zz}(x, \eta, z) = 0. \tag{10.1.143}$$

To ensure that the original leading term $2ik\hat{w}_z(x, \eta, z)$ remains as significant as a second derivative term in (10.1.143), we must set $r = 1/2$. This yields $k(2i\hat{w}_z(x, \eta, z) + \hat{w}_{\eta\eta}(x, \eta, z)) + \hat{w}_{xx}(x, \eta, z) + \hat{w}_{zz}(x, \eta, z) = 0$, and the equation for the leading-order boundary layer term $\hat{w}_0(x, \eta, z)$ is

$$2i\frac{\partial \hat{w}_0(x, \eta, z)}{\partial z} + \frac{\partial^2 \hat{w}_0(x, \eta, z)}{\partial \eta^2} = 0. \tag{10.1.144}$$

This parabolic *Schrödinger equation* has already been encountered in our discussion of the reflection of a cylindrical wave by a parabola [see (10.1.103)]. We require that for large values of $\eta = k^{1/2}y$, the solution $\hat{w}_0(x, \eta, z)$ of (10.1.144) matches with the geometrical optics result (10.1.141).

We rewrite (10.1.144) in terms of the original y and z variables and obtain

$$2ik\frac{\partial w_0(x,y,z)}{\partial z} + \frac{\partial^2 w_0(x,y,z)}{\partial y^2} = 0. \tag{10.1.145}$$

Noting the results (10.1.104)-(10.1.105) in our earlier discussion of Schrödinger's equation, we consider a general solution of (10.1.145) in the form

$$w_0(x,y,z) = \sqrt{\frac{k}{z}}\int_{-\infty}^{\infty} f(s)\exp\left[\frac{ik(y-s)^2}{2z}\right] f(s)\,ds, \tag{10.1.146}$$

where $f(s)$ is to be specified by means of the matching process. Since k is large, we evaluate (10.1.146) asymptotically using the *method of stationary phase* as before. There is exactly one stationary point at $s = y$, and in view of (10.1.106) we have the asymptotic result $w_0(x,y,z) \approx \sqrt{2\pi}\, e^{i\pi/4} f(y)$, so that

$$w_0(x,y,z)e^{ikz} \approx \sqrt{2\pi}\, f(y)\exp\left(ikz + \frac{i\pi}{4}\right). \tag{10.1.147}$$

We must match (10.1.147) with the geometrical optics field (10.1.141) in the regions $z > 0, y > 0$ and $z > 0, y < 0$. The geometrical optics field vanishes in the (shadow) region $z > 0, y < 0$. so we must set $f(y) = 0$ for $y < 0$. For $z > 0, y > 0$, the geometrical optics field is $u_I(x,y,z) = e^{ikz}$, so that $f(y) = (1/\sqrt{2\pi})\exp(-i\pi/4)$ for $y > 0$. Thus, $f(y) = (1/\sqrt{2\pi})\exp(-i\pi/4)$, $y > 0$, $f(y) = 0$, $y < 0$, and $w_0(x,y,z)$ becomes

$$w_0(x,y,z) = \left(\frac{k}{2\pi z}\right)^{1/2}\exp\left(-\frac{i\pi}{4}\right)\int_0^{\infty}\exp\left[\frac{ik(y-s)^2}{2z}\right] ds. \tag{10.1.148}$$

With the change of variables $\sigma = (s-y)\,(k/2z)^{1/2}$, (10.1.148) takes the form

$$w_0(x,y,z) = \pi^{-1/2}\exp\left(-\frac{i\pi}{4}\right) F\left(-y\sqrt{\frac{k}{2z}}\right), \tag{10.1.149}$$

where the *Fresnel integral* $F(t)$ is defined as $F(t) = \int_t^{\infty} e^{i\sigma^2}\, d\sigma$. The field near the shadow boundary $y = 0$, $z > 0$ is thus given as, to leading order,

$$u(x,y,z) \approx \pi^{-1/2} F\left(-y\sqrt{\frac{k}{2z}}\right)\exp\left(ikz - \frac{i\pi}{4}\right). \tag{10.1.150}$$

Since $F(-\infty) = \sqrt{\pi}\exp(i\pi/4)$ and $F(+\infty) = 0$, we find that $w_0(x,y,z) \to 0$ as $z \to 0$ with $y < 0$, while $w_0(x,y,z) \to 1$ as $z \to 0$ with $y > 0$. Thus $w_0(x,y,z)$

in (10.1.149) is a solution of the parabolic equation (10.1.145) with the discontinuous initial data $w_0(x, y, 0) = 0$ for $y < 0$ and $w_0(x, y, 0) = 1$ for $y > 0$. The *parabolic equation method* of Section 9.3 applied to this problem would have led us to consider (10.1.145) with the aforementioned data. Further, $F(0) = \frac{1}{2}\sqrt{\pi}\exp(i\pi/4)$, so that $w_0(x, y, z)$ equals $\frac{1}{2}$ on the shadow boundary $y = 0$, $z > 0$. This value is the average of the amplitudes 0 and 1 of the field in the shadow and illuminated regions, respectively.

In the matching process we considered only the leading term in the asymptotic expansion of the boundary layer term $w_0(x, y, z)$. On evaluating the representation (10.1.148) of $w_0(x, y, z)$ by the *method of stationary phase*, we find that there is only one stationary point located at $s = y$. Since the domain of integration is $0 < s < \infty$, there is no stationary point in the region $y < 0$, $z > 0$. In the region $y \geq 0$, $z > 0$, the stationary point coincides with the endpoint $s = 0$ of the integration interval if $y = 0$. Otherwise, the stationary point $s = y$ and the endpoint $s = 0$ are distinct.

We now apply the stationary phase results of Section 5.7 to the integral (10.1.148) for $w_0(x, y, z)$. If $y < 0$, there is no stationary point in the given interval of integration. Thus the main asymptotic contribution to (10.1.148) comes from the endpoint $s = 0$, and we obtain

$$w_0(x, y, z) \approx -\frac{z}{y}\left(\frac{1}{2\pi kz}\right)^{1/2} \exp\left(\frac{iky^2}{2z} + \frac{i\pi}{4}\right), \qquad y < 0. \tag{10.1.151}$$

If $y > 0$, the integral (10.1.148) has a stationary point contribution from the point $s = y$ as well as an endpoint contribution from $s = 0$. The stationary point result was found to be $w_0(x, y, z) \approx 1$, while the endpoint result is identical to (10.1.151). Thus

$$w_0(x, y, z) \approx 1 - \frac{z}{y}\left(\frac{1}{2\pi kz}\right)^{1/2} \exp\left(\frac{iky^2}{2z} + \frac{i\pi}{4}\right), \qquad y > 0. \tag{10.1.152}$$

Both (10.1.151) and (10.1.152) are singular at $y = 0$ as well as at $z = 0$

Since $w_0(x, y, z)$ is valid only in the region near the shadow boundary, we conclude from (10.1.151)–(10.1.152) that the total field $u(x, y, z)$ has the form

$$u(x, y, z) \approx \begin{cases} -\frac{z}{y}\left(\frac{1}{2\pi kz}\right)^{1/2} \exp\left[ik\left(z + \frac{y^2}{2z}\right) + \frac{i\pi}{4}\right], & z > 0,\ y < 0, \\ e^{ikz} - \frac{z}{y}\left(\frac{1}{2\pi kz}\right)^{1/2} \exp\left[ik\left(z + \frac{y^2}{2z}\right) + \frac{i\pi}{4}\right], & z > 0,\ y > 0, \end{cases} \tag{10.1.153}$$

near the edge of the shadow boundary. On comparing (10.1.153) with the geometrical optics result (10.1.141), we observe that there is no counterpart to the $O(k^{-1/2})$ terms of (10.1.153) in (10.1.141). In fact, the geometrical optics result was expressed as an (asymptotic) series in integral powers of k^{-1}. Thus terms of order $k^{-1/2}$ could not arise in that series. These $O(k^{-1/2})$ terms are due to *diffraction effects*. In the preceding problem, the field due to diffraction below the caustic was found to decay exponentially. In this problem, the diffraction field is smaller than the geometrical

optics field by a multiplicative factor of $O(k^{-1/2})$. It is not exponentially small, so it must be accounted for in our asymptotic solution of the scattering problem.

An asymptotic method closely related to the geometrical optics approach that accounts for the effects of diffraction is J. B. Keller's *geometrical theory of diffraction*. In the context of the present problem this theory introduces a set of *diffracted rays* that emanate from the edge of the boundary surface (i.e., the half-plane $z = 0$, $y \leq 0$). The incident rays, in addition to giving rise to the reflected rays in the region $z < 0$, $y < 0$, generate a set of *diffracted rays*. Each incident ray that hits the edge yields a family of *edge-diffracted rays* that emanate from the edge as shown in Figure 10.5. These rays correspond to those of a cylindrical wave with a focal point at the edge. If we set $r = \sqrt{y^2 + z^2}$, the diffracted field $u_D(x, y, z)$—which supplements the geometrical optics field—is shown in the geometrical theory of diffraction to have the form

$$u_D(x, y, z) \approx \frac{\tilde{D}}{\sqrt{kr}} u_I(x, 0, 0)\, e^{ikr}, \tag{10.1.154}$$

where $\tilde{D}$ is referred to as a *diffraction coefficient* and $u_I(x, 0, 0)$ is the value of the incident field at the point of diffraction $(x, 0, 0)$ on the edge of the half-plane.

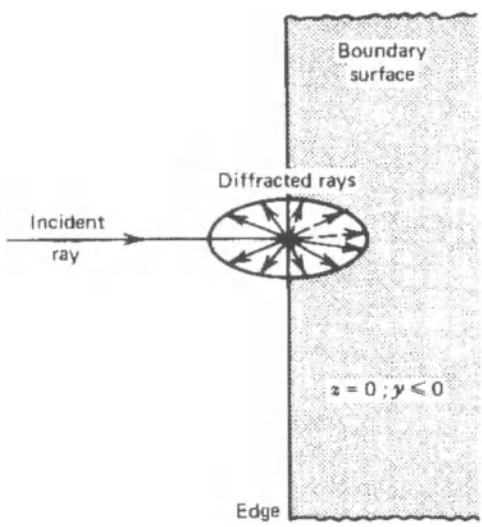

Figure 10.5 Edge-diffracted rays.

The diffraction coefficient $\tilde{D}$ must be determined either by solving a *canonical problem* or by using *boundary layer theory* near the edge. In general, a canonical problem is one in which the local geometry and other properties are the same as those for the given problem near the point of diffraction. Using either of these methods leads one to consider the problem of the scattering of a plane wave by a half-plane with the same boundary condition as for our original given problem. That is, the canonical

problem is identical to the given problem in the present case. (In general, of course, the canonical problem is much simpler than the given one.) The exact solution of our problem was found by *Sommerfeld*. On expanding his result asymptotically for large kr, it is found that the diffraction coefficient $\tilde{D}$ must be chosen as

$$\tilde{D} = -\frac{e^{i\pi/4}}{2(2\pi)^{1/2}}\left[\sec\left(\frac{\tilde{\theta}}{2}\right) + \csc\left(\frac{\tilde{\theta}}{2}\right)\right], \tag{10.1.155}$$

where $\tilde{\theta}$ is the angle the diffracted ray, when projected onto the (y, z)-plane, makes with the negative z-axis. Also, $u_I(x, 0, 0) = 1$ for all values of x.

The terms of order $k^{-1/2}$ in (10.1.153) must be identified with the diffracted field $u_D(x, y, z)$ in the region near the shadow boundary. To effect this comparison more easily, we introduce the cylindrical coordinate system $x = x$, $y = r\cos(\theta)$, $z = r\sin(\theta)$. The angles $\tilde{\theta}$ of (10.1.155) and θ given above are related by $\tilde{\theta} = \theta + \pi/2$. For small y with $z > 0$, we have $r = \sqrt{y^2 + z^2} \approx z + y^2/2z$. Then we easily conclude that $\sin(\theta) \approx 1$, $\cos(\theta) \approx 0$. Also, $(1/2)\left[\sec\left(\tilde{\theta}/2\right) + \csc\left(\tilde{\theta}/2\right)\right] = \left[\cos(\tilde{\theta}/2) + \sin(\tilde{\theta}/2)\right]/\sin(\tilde{\theta})$.

Noting the above gives $-(z/y)\,(1/2\pi kz)^{1/2}\exp\left[ik\left(z + y^2/2z\right) + i\pi/4\right] \approx -[\sin(\theta)/\cos(\theta)](1/2\pi kr)^{1/2}e^{ikr+i\pi/4} \approx -\left[(1/2\pi kr)^{1/2}/\sin(\tilde{\theta})\right]e^{ikr+i\pi/4}$ since $\cos(\theta) = \sin(\tilde{\theta})$ and $\sin(\theta) \approx 1$. Further, we find that $\theta \approx \pi/2$, so that $\tilde{\theta} \approx \pi$. Consequently, $\cos(\tilde{\theta}/2) \approx 0$ and $\sin(\tilde{\theta}/2) \approx 1$. Then we conclude that $\tilde{D}$ has the approximate form $\tilde{D} \approx -e^{i\pi/4}/\left[\sqrt{2\pi}\sin(\tilde{\theta})\right]$. Inserting this into (10.1.154) we find that the diffracted field $u_D(x, y, z)$ and the foregoing result agree in the neighborhood of the shadow boundary since $u_I(x, 0, 0) = 1$.

The field near the *reflection boundary* is considered in the exercises.

Scattering by a Circular Cylinder

A *plane wave* $u_I(x, y, z) = e^{ikx}$ is incident upon a *circular cylinder* of radius a, whose axis coincides with the z-axis. The total field $u(x, y, z) = u_I(x, y, z) + u_S(x, y, z)$, where u_S is the *scattered wave*, satisfies the reduced wave equation (10.1.1)—with $n = 1$—and vanishes on the surface of the cylinder. Clearly, this can be formulated as a problem in two dimensions.

Thus we consider $u = u(x, y)$ that satisfies the reduced wave equation

$$u_{xx}(x, y) + u_{yy}(x, y) + k^2u(x, y) = 0 \tag{10.1.156}$$

in the exterior of the circle $x^2 + y^2 = a^2$. On the circle we have $u(x, y) = 0$, $x^2 + y^2 = a^2$. With $u(x, y) = u_I(x, y) + u_S(x, y)$ where $u_I(x, y) = e^{ikx}$, we require that $u_S(x, y)$ satisfies the radiation condition at infinity.

Assuming that $k \gg 1$, we determine an asymptotic solution for $u_S(x,y)$ in the manner described. The scattered field $u_S(x,y)$ is expanded as $u_S(x,y) = e^{ik\phi(x,y)} \sum_{j=0}^{\infty} v_j(x,y)(ik)^{-j}$. On the circle, the phase term $\phi(x,y)$ satisfies the (initial) condition

$$\phi(x,y) = x, \qquad x^2 + y^2 = a^2 \tag{10.1.157}$$

as well as an outgoing condition. The amplitude terms $v_j(x,y)$ satisfy the initial conditions

$$v_0(x,y) = -1,\ x^2 + y^2 = a^2, \quad v_j(x,y) = 0,\ x^2 + y^2 = a^2,\ j \geq 1. \tag{10.1.158}$$

To solve for $\phi(x,y)$ we represent the circle in parametric form as $x = a\cos(\tau)$, $y = a\sin(\tau)$, $0 \leq \tau \leq 2\pi$. We note that τ is not an arc length parameter on the circle unless $a = 1$. With σ as an arc length parameter along the rays we must solve the characteristic equations for $x(\sigma,\tau)$, $y(\sigma,\tau)$, $\phi(\sigma,\tau)$, $p(\sigma,\tau)$, and $q(\sigma,\tau)$ in order to determine $\phi(x,y)$. The initial conditions for $x(\sigma,\tau)$ and $y(\sigma,\tau)$, which are given at $\sigma = 0$, are

$$x(0,\tau) = a\cos(\tau), \qquad y(0,\tau) = a\sin(\tau). \tag{10.1.159}$$

From (10.1.157) we have $\phi(0,\tau) = x(0,\tau) = a\cos(\tau)$.

The initial conditions for $p(\sigma,\tau)$ and $q(\sigma,\tau)$ must be found from the eiconal equation and strip condition as in Section 2.4. From the eiconal equation we have $p^2(0,\tau) + q^2(0,\tau) = 1$. The strip condition yields $-a\sin(\tau) = -p(0,\tau)a\sin(\tau) + q(0,\tau)\, a\cos(\tau)$. On solving these equations we easily conclude that there are two possible sets of solutions. They are

$$p(0,\tau) = 1, \qquad q(0,\tau) = 0, \tag{10.1.160}$$

$$p(0,\tau) = -\cos(2\tau), \qquad q(0,\tau) = -\sin(2\tau). \tag{10.1.161}$$

The appropriate solutions must be determined from the outgoing condition. This means that if $\mathbf{N}$ is the exterior unit normal to the circle, $\nabla\phi\cdot\mathbf{N} = [p(0,\tau), q(0,\tau)]\cdot\mathbf{N} > 0$. The normal $\mathbf{N}$ for $x^2 + y^2 = a^2$ is $\mathbf{N} = [\cos(\tau), \sin(\tau)]$. Thus $\nabla\phi \cdot \mathbf{N} = [p(0,\tau), q(0,\tau)] \cdot \mathbf{N} = \cos(\tau)$ if (10.1.160) is used, and $\nabla\phi \cdot \mathbf{N} = [p(0,\tau), q(0,\tau)] \cdot \mathbf{N} = -\cos(\tau)$ if (10.1.161) is used. Therefore, since $\cos(\tau) > 0$ for $0 \leq \tau < \pi/2$ and $3\pi/2 < \tau \leq 2\pi$, and $\cos(\tau) < 0$ for $\pi/2 < \tau < 3\pi/2$, we see that $\nabla\phi \cdot \mathbf{N} > 0$ on the left semicircle (i.e., if $\pi/2 < \tau < 3\pi/2$) when (10.1.161) is chosen, while $\nabla\phi \cdot \mathbf{N} > 0$ on the right semicircle if (10.1.160) is selected. At the points $(0,a)$ and $(0,-a)$ where $\tau = \pi/2$ and $\tau = 3\pi/2$, respectively, the rays of the scattered wave are tangent to the circle.

It has been shown that on the right semicircle (i.e., $x^2 + y^2 = a^2$ with $x > 0$) we must set $p(0,\tau) = 1$ and $q(0,\tau) = 0$. Since the characteristic equations are

$$\frac{dx}{d\sigma} = p, \qquad \frac{dy}{d\sigma} = q, \qquad \frac{d\phi}{d\sigma} = 1, \qquad \frac{dp}{d\sigma} = 0, \qquad \frac{dq}{d\sigma} = 0, \tag{10.1.162}$$

we see that $p(\sigma,\tau) = 1$ and $q(\sigma,\tau) = 0$. Then, the initial conditions imply that $x(\sigma,\tau) = \sigma + a\cos(\tau)$, $y(\sigma,\tau) = a\sin(\tau)$, $\phi(\sigma,\tau) = \sigma + a\cos(\tau) = x$. Thus the rays of the scattered wave that issue from the right semicircle are identical to those of the incident wave. The phase term $\phi(x,y)$ equals x, the phase of the incident wave. Further, since $\nabla^2\phi = 0$, we conclude that $v_0(x,y) = -1$, $v_j(x,y) = 0$, $j \geq 1$, from the transport equations (10.1.54)–(10.1.55) and the initial conditions. Therefore, the scattered field is $u_S(x,y) = -e^{ikx}$, $x^2 + y^2 \geq a^2$, $x > 0$, $|y| < a$, in the region directly behind the circle where no incident rays penetrate. Thus, $u(x,y) = u_I(x,y) + u_S(x,y)$ is

$$u(x,y) = e^{ikx} - e^{ikx} = 0, \qquad x^2 + y^2 \geq a^2,\ x > 0,\ |y| < a. \tag{10.1.163}$$

The *geometrical optics* solution yields a zero total field or, equivalently, a *shadow* in the region behind the circle not illuminated by the incident rays.

On the left semicircle (i.e., $x^2 + y^2 = a^2$ with $x < 0$) we must have $p(0,\tau) = -\cos(2\tau)$ and $q(0,\tau) = -\sin(2\tau)$, as has been shown. Since $p(\sigma,\tau)$ and $q(\sigma,\tau)$ are constant along the characteristics, they retain their initial values along these curves. Consequently, the equations for $x(\sigma,\tau)$ and $y(\sigma,\tau)$ are

$$\frac{dx(\sigma,\tau)}{d\sigma} = p(\sigma,\tau) = -\cos(2\tau), \qquad \frac{dy}{d\sigma} = q(\sigma,\tau) = -\sin(2\tau), \tag{10.1.164}$$

with $\pi/2 \leq \tau \leq 3\pi/2$. Noting the initial conditions (10.1.159), we find that

$$x(\sigma,\tau) = a\cos(\tau) - \sigma\cos(2\tau), \qquad y(\sigma,\tau) = a\sin(\tau) - \sigma\sin(2\tau). \tag{10.1.165}$$

Also, $d\phi(\sigma,\tau)/d\sigma = 1$ and (10.236) imply that $\phi(\sigma,\tau) = \sigma + a\cos(\tau)$, $\pi/2 \leq \tau \leq 3\pi/2$. It is not possible to express σ and τ as functions of x and y in simple form. Therefore, we leave our result for the phase $\phi(\sigma,\tau)$ in parametric form.

The *rays* of the *scattered* or *reflected field* $u_S(\sigma,\tau)$ are determined from (10.1.165), and their direction is given by (10.1.164). The direction of the incident rays is that of the unit vector $\mathbf{i}$. The exterior unit normal vector to the circle is $\mathbf{N} = [\cos(\tau), \sin(\tau)]$. It is a simple matter to show that the *angle of incidence* α_I of the *incident ray* equals the *angle of reflection* α_R of the *reflected ray*. These angles are defined in Figure 10.6. In fact, we have $\alpha_I = \alpha_R = |\pi - \tau|$ with $\pi/2 \leq \tau \leq 3\pi/2$. When $\tau = \pi/2$ and $\tau = 3\pi/2$, the reflected rays coincide with the incident rays. That is, the incident rays are tangent to the circle at the points $(0,a)$ and $(0,-a)$ and their extension beyond these points into the region $x > 0$ coincides with the reflected rays. (We note that the equality of the angles of incidence and reflection is a general principle of geometrical optics.) The region covered by the reflected and the incident rays is given as $x^2 + y^2 \geq a^2$, with $-\infty < y < \infty$ for $x \leq 0$ and $|y| \geq a$ for $x > 0$. This region is referred to as the *illuminated region*.

We now proceed to determine the leading-order amplitude term $v_0(\sigma,\tau)$ of the reflected wave u_S. To do so we express the transport equation for $v_0(\sigma,\tau)$, which is given as $2\,\partial v_0(\sigma,\tau)/\partial\sigma + v_0(\sigma,\tau)\nabla^2\phi(\sigma,\tau) = 0$, in the ray coordinates σ and τ determined from (10.1.165). The Laplacian ∇^2 must be expressed in the

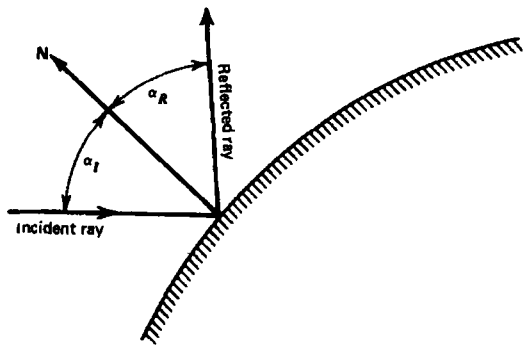

Figure 10.6 Angles of incidence and reflection.

ray coordinate system so that $\nabla^2\phi(\sigma,\tau)$ can be evaluated. It is readily found that $\nabla^2\phi(\sigma,\tau) = 1/[\sigma - (a/2)\cos(\tau)]$, $\pi/2 < \tau < 3\pi/2$. Inserting this result into the above and using the initial condition $v_0(0,\tau) = -1$, we obtain $v_0(\sigma,\tau) = -\sqrt{[(a/2)\cos(\tau)]/[(a/2)\cos(\tau) - \sigma]}$, $\pi/2 < \tau < 3\pi/2$, as is easily verified. Thus the scattered field is given, to leading order, as

$$u_S(\sigma,\tau) \approx -\sqrt{\frac{(a/2)\cos(\tau)}{(a/2)\cos(\tau) - \sigma}}\ \exp[ik(\sigma + a\cos(\tau))], \qquad \frac{\pi}{2} < \tau < \frac{3\pi}{2}. \tag{10.1.166}$$

The curve

$$\sigma = \frac{a}{2}\cos(\tau), \qquad \frac{\pi}{2} \le \tau \le \frac{3\pi}{2}, \tag{10.1.167}$$

is the envelope of the reflected ray system (10.1.165), as is verified by evaluating the Jacobian of the transformation (10.1.165) or, equivalently, by using the results in the Appendix of Chapter 2. When $\pi/2 < \tau < 3\pi/2$, we find that $\cos(\tau) < 0$, so that σ is negative in (10.1.167). Therefore, the envelope lies within the circle $x^2 + y^2 = a^2$ for those values of τ, since $\sigma \ge 0$ corresponds to points on or outside the circle. However, if $\tau = \pi/2$ or $\tau = 3\pi/2$, we have $\cos(\pi/2) = \cos(3\pi/2) = 0$, so that $\sigma = 0$ there. Consequently, the envelope or caustic curve (10.1.167) intersects the circle at the points $(0,a)$ and $(0,-a)$. It is for this reason that $u_S(\sigma,\tau)$ is not well defined at $\tau = \pi/2$ and $\tau = 3\pi/2$.

The asymptotic (geometrical optics) result we have obtained is the following. In the *illuminated region* the solution $u = u(\sigma,\tau)$ is

$$u \approx \exp[ik(a\cos(\tau) - \sigma\cos(2\tau))] - \sqrt{\frac{(a/2)\cos(\tau)}{(a/2)\cos(\tau) - \sigma}}\ \exp[ik(\sigma + a\cos(\tau))], \tag{10.1.168}$$

and this is valid for $\pi/2 < \tau < 3\pi/2$ and $\sigma \ge 0$. In the *shadow region* we have $u(x,y) \approx 0$, $x^2 + y^2 \ge a^2$, $x > 0$, $|y| < a$. The lines $y = \pm a$, $x \ge 0$ are the

shadow boundaries that separate the illuminated region from the shadow region. The asymptotic solution is not continuous across these lines.

The field near the shadow boundaries may be studied by the use of the boundary layer theory. The parabolic equation (10.1.144) may be derived. However, the solution must be matched not only with the incident and reflected waves in the illuminated region and the zero field in the shadow region. The field in the shadow boundary must also be matched with the field near the *points of diffraction* $(0, a)$ and $(0, -a)$. The field near these points may be analyzed by proceeding as in Section 9.3 using boundary layer methods. The problem is rather complicated and has undergone much investigation. The *geometrical theory of diffraction* introduces *surface diffracted rays*, which yield an additional exponentially small field in the shadow and illuminated regions. However, we do not pursue these matters here any further.

Propagation of a Gaussian Beam

We have seen above that the geometrical optics expansion associated with a *general plane wave* is not valid over large distances because of the occurrence of secular effects. In the following we consider the propagation of a narrow beam of light, as might be generated by a *laser*, and analyze this problem by means of a geometrical optics expansion.

We consider a *light beam*, in the two-dimensional case, that propagates in the direction of the positive x-axis with its amplitude concentrated near the x-axis. It is represented by a function $u = u(x, y)$ that satisfies the reduced wave equation (10.1.1), where we set $n = 1$. We put $u(x, y) = v(x, y)e^{ik\phi(x,y)}$ and expand $v(x, y)$ as in (10.1.11). At $x = 0$, the phase term $\phi(x, y)$ and the amplitude term $v(x, y)$ are given as

$$\phi(0, y) = 0, \qquad v(0, y) = A \exp\left(-\frac{y^2}{2a}\right), \tag{10.1.169}$$

where A and a are positive constants. The solution is determined in the half-plane $x > 0$, so that (10.1.169) represents a boundary condition for the beam field $u(x, y)$. The solution is required to be outgoing at $x = 0$.

We do not consider how the beam is formed in the half-plane $x < 0$, but assume that it has the phase and amplitude given in (10.1.169) when it reaches the y-axis. Since the amplitude is in the form of a Gaussian distribution (see Section 1.1) we refer to it as a *Gaussian beam*. The *beam half-width* is a measure of the distance from the x-axis within which the amplitude of the beam is concentrated and it equals $\sqrt{2a}$ on the y-axis. (The half-width generally varies with x.)

Now $\phi(x, y)$ is a solution of the eiconal equation (10.1.9) with $n = 1$. The initial condition (10.1.169) together with the outgoing condition yields the solution $\phi(x, y) = x$. Thus, $u(x, y) = v(x, y)\, e^{ikx}$ and $v(x, y)$ satisfies

$$2ikv_x(x, y) + v_{xx}(x, y) + v_{yy}(x, y) = 0. \tag{10.1.170}$$

On expanding $v(x, y)$ as in (10.1.11), we find that $v_0(x, y)$ satisfies the equation $\partial v_0(x, y)/\partial x = 0$ and the boundary condition (10.1.169). The remaining $v_j(x, y)$

satisfy a nonhomogeneous version of the equation for $v_0(x, y)$ and they all vanish at $x = 0$. On solving for $v_0(x, y)$ and $v_1(x, y)$, we obtain

$$v(x, y) \approx v_0(x, y) + \frac{1}{ik} v_1(x, y) = A \exp\left(-\frac{y^2}{2a}\right)\left(1 + \frac{ix}{2ka} - \frac{ixy^2}{2ka^2}\right). \tag{10.1.171}$$

Since the amplitude term is concentrated near $y = 0$, we determine that the asymptotic expansion of $v(x, y)$ begins to become disordered when $x \approx 2ka$. Then the second term in the expansion of $v(x, y)$ has the same magnitude as the leading term, when $y = 0$. The geometrical optics description of the beam is not valid as x reaches this value and diffraction effects begin to play a role. Consequently, the distance $x = 2ka$ is referred to as a *diffraction length.*

The geometrical optics representation of the beam is that of a *general plane wave.* To study the solution in the secular region where $x \geq O(ka)$ we proceed as before. The results leading from (10.1.101) to (10.1.105) apply to the present problem as well, and the expression in (10.1.105) is equated to the Gaussian amplitude term in (10.1.169). This yields

$$v(x, y) \approx A e^{-i\pi/4} \sqrt{\frac{k}{2\pi x}} \int_{-\infty}^{\infty} \exp\left[\frac{ik(y - s)^2}{2x} - \frac{s^2}{2a}\right] ds. \tag{10.1.172}$$

The integral can be evaluated explicitly by completing the square in the exponential term and introducing a complex-valued change of variables that reduces the integral to the form given in Exercise 1.1.2. (The transformation may be carried out formally, and the resulting expression can then be shown to satisfy the conditions of the problem.) The *Gaussian beam* representation is thus given as

$$u(x, y) \approx A \frac{\sqrt{ka}}{\sqrt{x - ika}} \exp\left[ikx + \frac{iky^2}{2(x - ika)} - \frac{i\pi}{4}\right]. \tag{10.1.173}$$

If we express (10.1.173) in the form $u(x, y) = v(x, y)e^{ik\phi(x,y)}$ with $\phi(x, y) = x$, we find that it satisfies (10.1.169) at $x = 0$. Further, $v(x, y)$ is a solution of the *paraxial wave equation*

$$2ikv_x(x, y) + v_{yy}(x, y) = 0, \tag{10.1.174}$$

which is, in fact, the parabolic equation approximation to (10.1.170). It has the form of a Schrödinger equation. The term *paraxial* signifies that it describes the field $u(x, y) = v(x, y)e^{ikx}$ in a narrow region parallel to the axis of propagation (in this case the x-axis).

If $x < O(ka)$, the beam representation (10.1.173) approximates the form of the general plane wave (10.1.171), since $\sqrt{ka}/\sqrt{x - ika}\, e^{-i\pi/4} \approx 1 - ix/2ka$ and $iky^2/2(x - ika) \approx -y^2/2a + ixy^2/2ka^2$. Then, with $y \approx 0$ we have $\exp[ixy^2/2ka^2] \approx 1 + ixy^2/2ka^2$ and the approximation (10.1.171) follows.

The (approximate) general plane wave representation of the beam (10.1.173) loses its validity for large x. The beam begins to spread out and assumes the form of

a general *cylindrical wave*. This effect is a result of *diffraction*. To see this we introduce polar coordinates $x = r\cos(\theta)$, $y = r\sin(\theta)$, and assume that $x \gg ka$ and that $y/x \leq 1$. Then the radial variable r can be approximated as $r \approx x + y^2/2x$ and (10.1.173) takes the form

$$u(x,y) \approx A\frac{\sqrt{ka}}{\sqrt{r}} \exp\left(-\frac{1}{2}k^2 a \tan^2(\theta)\right) \exp\left(ikr - \frac{i\pi}{4}\right). \tag{10.1.175}$$

This has the form of a cylindrical wave with source point at the origin. Each value of θ determines a ray and we see that as $|\theta|$ increases, the amplitude term decreases so that we have a *cylindrical beam.*

Nonlinear Reduced Wave Equation and Nonlinear Geometrical Optics

Our discussion has so far been limited to the linear reduced wave equation (10.1.1) and its asymptotic solutions. However, in the theory of *nonlinear optics* one is led to consider a nonlinear version of (10.1.1). If a strong *laser beam* propagates in a medium, it may happen as a result of various effects that the index of refraction is altered as the beam traverses the medium. Assuming the medium to be otherwise homogeneous, the index of refraction n is found to become a function of the *intensity* of the field. If $u(\mathbf{x})$ represents the (complex) field, we have $n = n(|u(\mathbf{x})|^2)$ and the field satisfies the *nonlinear reduced wave equation*

$$\nabla^2 u(\mathbf{x}) + k^2 n^2(|u(\mathbf{x})|^2)u(\mathbf{x}) = 0 \tag{10.1.176}$$

in two or three space dimensions. We again assume that the wave number k is large and look for asymptotic solutions of (10.1.176).

It is immediately possible to construct *plane wave* solutions of (10.1.176). For example,

$$u(\mathbf{x}) = A\exp[ikn(A^2)x], \tag{10.1.177}$$

where the amplitude term A is a real constant and x is a scalar variable, satisfies (10.1.176). We note that in contrast to the linear problem studied in the foregoing, the phase in (10.1.177) is a function of the amplitude.

Noting that there are solutions of (10.1.176) in the form (10.1.177), we proceed as in the linear case, and look for a solution of the nonlinear reduced wave equation in the form

$$u(\mathbf{x}) = A(\mathbf{x})e^{ik\phi(\mathbf{x})}, \tag{10.1.178}$$

where the amplitude term $A(\mathbf{x})$ and the phase term $\phi(\mathbf{x})$ are assumed to be real valued (we no longer require that the amplitude be a constant). Inserting (10.1.178) into (10.1.176) and equating real and imaginary parts to zero gives

$$[(\nabla\phi(\mathbf{x}))^2 - n^2(A^2(\mathbf{x}))]A(\mathbf{x}) - \frac{1}{k^2}\nabla^2 A(\mathbf{x}) = 0, \tag{10.1.179}$$

$$2\nabla\phi(\mathbf{x}) \cdot \nabla A(\mathbf{x}) + A(\mathbf{x})\nabla^2\phi(\mathbf{x}) = 0. \tag{10.1.180}$$

In the geometrical optics approximation we let $k \to \infty$ in (10.1.179) and obtain

$$(\nabla\phi(\mathbf{x}))^2 = n^2(A^2(\mathbf{x})), \quad \nabla \cdot (A^2(\mathbf{x})\nabla\phi(\mathbf{x})) = 0, \tag{10.1.181}$$

where (10.1.68) has been used in (10.1.180). The equations (10.1.181) are referred to as the equations of *nonlinear geometrical optics*. In contrast to the situation in linear optics, these equations are coupled and must be solved simultaneously. Nevertheless, the *eiconal equation* in (10.1.181) determines phase surfaces $\phi(\mathbf{x})$ = constant and rays orthogonal to these surfaces. The *transport equation* in (10.1.181) determines the amplitude $A(\mathbf{x})$ in terms of the divergence and convergence of the rays, as in the linear problem. However, we cannot determine the rays without knowing the amplitude, and vice versa.

Self-Focusing and Self-Trapping of Beams in a Nonlinear Medium

We do not consider how the system (10.1.181) may be solved in general. Instead, we examine how a *beam* propagates in the nonlinear medium on the basis of the geometrical optics equations. A beam is characterized as follows. Let $\phi(\mathbf{x})$ = constant represent a particular phase surface or wave front. The amplitude $A(\mathbf{x})$ is assumed to have a maximum at a point P on the wave front and to decrease sharply in value as we move away from P in all directions on the wave front. Thus the significant values of the field are concentrated near P on the wave front $\phi(\mathbf{x})$ = constant, and these values are propagated along the rays in the form of a beam. We now assume that the index of refraction $n(A^2(\mathbf{x}))$ decreases as the amplitude $A(\mathbf{x})$ decreases. Since the amplitude falls off as we move away from P on $\phi(\mathbf{x})$ = constant, so does $n(A^2(\mathbf{x}))$. On considering the two neighboring wave fronts $\phi(\mathbf{x})$ = constant and $\phi(\mathbf{x}) + d\phi$ = constant, the distance between them measured on the (orthogonal) rays is given as $d\sigma = (1/n(A^2(\mathbf{x})))\, d\phi$ in view of (10.1.58). Since $d\phi$ is fixed, we find that as $A(\mathbf{x})$ and, consequently, $n(A^2(\mathbf{x}))$ decrease [on the wave front $\phi(\mathbf{x})$ = constant] as we move away from the point P, the distance $d\sigma$ between the two wave fronts increases. That is, $\phi(\mathbf{x})$ = constant and $\phi(\mathbf{x}) + d\phi$ = constant are nearest each other at P and become increasingly separated away from P. The rays, however, always remain orthogonal to the wave fronts. On applying this argument to a succession of wave fronts, we obtain the situation pictured in Figure 10.7. The wave fronts become increasingly concave and the (orthogonal) rays begin to converge. If the field is symmetric with respect to the point P, the rays may eventually focus at some point. In any case, the rays begin to converge, and this property of nonlinear optics is known as the *self-focusing effect*. Even if the wave front $\phi(\mathbf{x})$ = constant is plane or initially convex, the dependence of the index of refraction n on the amplitude will bend the rays and force them to converge, possibly even to a focus. A medium for which the index of refraction has this property is called a *focusing medium*.

However, as the rays converge, the amplitude of the beam begins to increase in view of the relation between the amplitude and the ray structure. If the amplitude $A(\mathbf{x})$ becomes extremely large or possibly singular, the geometrical optics approximation that predicts the *self-focusing effect* no longer remains valid and *diffraction effects*

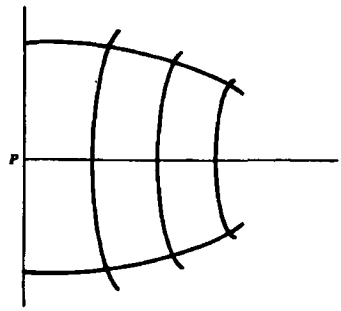

Figure 10.7 The self-focusing effect.

must be included. That is, we must retain second derivative terms in (10.1.179). In the linear problems discussed earlier the inclusion of diffraction effects was shown to remove the singularities predicted by the geometrical optics approach and to lead to a spreading of the rays in beam propagation. Here, because we have a nonlinear problem, it may happen that a beam that undergoes focusing in the geometrical optics approximation may actually focus and become singular even if diffraction effects are included.

To see that it is possible to obtain beams that do not undergo self-focusing, we construct an exact solution of (10.1.176) or, equivalently, of (10.1.179)–(10.1.180), which represents a beam. We consider the two-dimensional case with $u = u(x, y)$ and assume that $n = n(|u(x, y)|^2)$ is given as $n(|u(x, y)|^2) = n_0^2 + n_1|u(x, y)|^2$, that is, we have a quadratic nonlinearity. The constant term n_0 is the linear index of refraction that occurs when nonlinear effects can be neglected. The constant n_1 is positive, so that $n(A^2(x, y))$ decreases as $A^2(x, y)$ decreases. To solve (10.1.179)–(10.1.180) we set, with constant λ,

$$\phi(x, y) = \lambda x, \qquad A(x, y) = \tilde{A}(y). \tag{10.1.182}$$

Then (10.1.180) is satisfied identically and (10.1.179) becomes

$$\tilde{A}''(y) - k^2(\lambda^2 - n_0^2)\tilde{A}(y) + k^2 n_1 \tilde{A}^3(y) = 0. \tag{10.1.183}$$

It is easily verified that a solution of (10.1.183) that vanishes together with its derivatives as $|y| \to \infty$ is

$$\tilde{A}(y) = \left[\frac{2(\lambda^2 - n_0^2)}{n_1}\right]^{1/2} \operatorname{sech}[k(\lambda^2 - n_0^2)^{1/2} y] \tag{10.1.184}$$

if we assume that $|\lambda| > n_0 > 0$. The amplitude term $\tilde{A}(y)$ decays at an exponential

rate as $|y| \to \infty$ and is significantly different from zero only in the interval $|y| \leq O[k^{-1}(\lambda^2 - n_0^2)^{-1/2}]$. Thus the solution

$$u(x,y) = \left[\frac{2(\lambda^2 - n_0^2)}{n_1}\right]^{1/2} \operatorname{sech}[k(\lambda^2 - n_0^2)^{1/2} y]\, e^{ik\lambda x} \tag{10.1.185}$$

represents a *beam* that travels without change of amplitude along the x-axis. We have shown in our discussion of a Gaussian beam, that a plane wave solution of the linear reduced wave equation in the form of a beam is described by a parabolic equation in the far field. Although the geometric optics result in that problem predicted that the plane wave travels without change of amplitude to leading order, the diffraction effects, as characterized by the parabolic equation, diffuse and spread the field out at large distances. In the nonlinear problem considered here, the diffraction effects again tend to spread out the beam, while the nonlinear effects tend to focus the beam. Since these effects appear to be balanced in (10.1.185) and the beam travels without spreading, the solution (10.1.185) is referred to as a *self-trapped beam.*

It should be noted that if the index $n(A^2)$ is such that it increases as A decreases, our discussion shows that the rays diverge as a result of the nonlinearity. An initially thin beam would begin to spread out as it moved through the nonlinear medium. Therefore, we refer to such media as *defocusing media.* For example, if $n(|u(\mathbf{x})|^2) = n_0^2 + n_1|u(\mathbf{x})|^2$ with $n_1 < 0$, the index n has the defocusing property. Self-trapped solutions of the form (10.1.185) are not possible for such a medium since the amplitude A is real by assumption, and if $n_1 < 0$, the amplitude term (10.1.184) is purely imaginary. In fact, since both diffraction and nonlinear effects tend to spread out the beam, we would not expect that a self-trapped beam would occur in a defocusing medium.

To determine whether the self-trapped beam (10.1.185) can be expressed in some simpler form, we express the hyperbolic secant in terms of exponentials and easily conclude that

$$u(x,y) = 2\left[\frac{2(\lambda^2 - n_0^2)}{n_1}\right]^{1/2} \exp[ik\lambda x - k(\lambda^2 - n_0^2)^{1/2}|y|] \times \sum_{j=0}^{\infty}(-1)^j \exp[-2jk(\lambda^2 - n_0^2)^{1/2}|y|] \tag{10.1.186}$$

if $|y| > 0$. Since the amplitude term is exponentially small, we could not generate (10.1.186) by looking for an asymptotic solution of (10.1.176) in the form $u = Ae^{ik\phi}$, where A is expanded in inverse powers of k. However, since λ is arbitrary (apart from the requirement that $\lambda^2 > n_0^2$) we may consider solutions of the form (10.1.185), in which $\lambda^2 \approx n_0^2 + c^2/k^2$, where c is a constant. Then (10.1.185) becomes

$$u(x,y) \approx \frac{c}{k}\left(\frac{2}{n_1}\right)^{1/2} \operatorname{sech}(cy) \exp\left[i\left(kn_0 x + \frac{c^2 x}{2kn_0}\right)\right]. \tag{10.1.187}$$

This suggests that if we ask for solutions of (10.1.176) in the form $|u(\mathbf{x})| = O(k^{-r})$ with $r > 0$, an asymptotic approach similar to that used for the linear problem can

be applied to the nonlinear problem. We note that with the assumption above on $|u(\mathbf{x})| = A(\mathbf{x})$, we are essentially considering a weakly nonlinear problem since $k \gg 1$ by assumption.

With $|u(\mathbf{x})| = O(k^{-r})$, $r > 0$, we assume that $n(|u(\mathbf{x})|^2)$ can be expanded as

$$n(|u(\mathbf{x})|^2) = n_0^2 + n_1|u(\mathbf{x})|^2 + n_2|u(\mathbf{x})|^4 + \cdots, \tag{10.1.188}$$

where the coefficients n_j with $j = 1, 2, \ldots$ are all constants and $n_1 > 0$. As we have indicated, with n_1 positive, the index of refraction (10.1.188) corresponds to a focusing medium if $|u(\mathbf{x})|$ is small. We define the amplitude $|u(\mathbf{x})|$ as $|u(\mathbf{x})| = A(\mathbf{x}) = (1/k^r)a(\mathbf{x})$, $r > 0$. With $u(\mathbf{x}) = A(\mathbf{x})e^{ik\phi(\mathbf{x})}$ as in (10.1.178), we obtain for $\phi(\mathbf{x})$ and $a(\mathbf{x})$,

$$\left\{[(\nabla\phi(\mathbf{x}))^2 - n_0^2]a(\mathbf{x}) - \frac{1}{k^{2r}}n_1a^3(\mathbf{x}) - \frac{1}{k^{4r}}n_2a^3(\mathbf{x}) - \cdots\right\} - \frac{1}{k^2}\nabla^2a(\mathbf{x}) = 0, \tag{10.1.189}$$

$$2\nabla\phi(\mathbf{x}) \cdot \nabla a(\mathbf{x}) + a(\mathbf{x})\nabla^2\phi(\mathbf{x}) = 0. \tag{10.1.190}$$

To proceed with the solution of (10.1.189)–(10.1.190) we must specify the exponent r and then expand $\phi(\mathbf{x})$ and $a(\mathbf{x})$ in inverse powers of k. The most meaningful and interesting choices for r are $r = \frac{1}{2}$ and $r = 1$. With $r = \frac{1}{2}$ we find that $\phi(\mathbf{x})$ and $a(\mathbf{x})$ should be expanded in powers of k^{-1}, whereas if $r = 1$, the expansion of $\phi(\mathbf{x})$ and $a(\mathbf{x})$ should contain powers of k^{-2}. Now if $r = 1$, we see from (10.1.189) that the cubic term $a^3(\mathbf{x})$ and the term $\nabla^2a(\mathbf{x})$ are of the same order in k. Thus at the second level of approximation the (leading) nonlinear term and the diffraction term $\nabla^2a(\mathbf{x})$ are balanced. A particular case of this occurs in the result (10.1.187) where $A(\mathbf{x}) = O(1/k)$ and we have a self-trapped wave. However, if $r = \frac{1}{2}$, the nonlinear effect due to $a^3(\mathbf{x})$ enters at the second level of approximation. The diffraction effect due to $\nabla^2a(\mathbf{x})$ does not occur until the third level of approximation as we shall see. We begin with the case where $r = \frac{1}{2}$. The case where $r = 1$ plays a role in the analysis of a propagating beam is given below.

With $r = \frac{1}{2}$ in (10.1.189), we expand $\phi(\mathbf{x})$ and $a(\mathbf{x})$ as

$$\phi(\mathbf{x}) = \sum_{j=0}^{\infty}\phi_j(\mathbf{x})k^{-j}, \quad a(\mathbf{x}) = \sum_{j=0}^{\infty}a_j(\mathbf{x})k^{-j}, \tag{10.1.191}$$

and insert these expansions into (10.1.189)–(10.1.190). Collecting like powers of $1/k$ and equating their coefficients to zero gives

$$(\nabla\phi_0(\mathbf{x}))^2 = n_0^2, \tag{10.1.192}$$

$$2\nabla\phi_0(\mathbf{x}) \cdot \nabla a_0(\mathbf{x}) + a_0(\mathbf{x})\nabla^2\phi_0(\mathbf{x}) = 0, \tag{10.1.193}$$

$$2\nabla\phi_0(\mathbf{x}) \cdot \nabla\phi_1(\mathbf{x}) = n_1a_0^2(\mathbf{x}), \tag{10.1.194}$$

$$2\nabla\phi_0(\mathbf{x})\cdot\nabla a_1(\mathbf{x}) + a_1(\mathbf{x})\nabla^2\phi_0(\mathbf{x}) = -2\nabla\phi_1(\mathbf{x})\cdot\nabla a_0(\mathbf{x}) - a_0(\mathbf{x})\nabla^2\phi_1(\mathbf{x}), \tag{10.1.195}$$

$$2\nabla\phi_0(\mathbf{x})\cdot\nabla\phi_2(\mathbf{x}) = n_2 a_0^4(\mathbf{x}) + 3n_1 a_0(\mathbf{x})a_1(\mathbf{x}) + \frac{1}{a_0(\mathbf{x})}\nabla^2 a_0(\mathbf{x}) - \frac{2a_1(\mathbf{x})}{a_0(\mathbf{x})}\nabla\phi_0(\mathbf{x})\cdot\nabla\phi_1(\mathbf{x}) - (\nabla\phi_1(\mathbf{x}))^2, \tag{10.1.196}$$

$$2\nabla\phi_0(\mathbf{x})\cdot\nabla a_2(\mathbf{x}) + a_2(\mathbf{x})\nabla^2\phi_0(\mathbf{x}) = -2\nabla\phi_2(\mathbf{x})\cdot\nabla a_0(\mathbf{x}) - a_0(\mathbf{x})\nabla^2\phi_2(\mathbf{x}) - 2\nabla\phi_1(\mathbf{x})\cdot\nabla a_1(\mathbf{x}) - a_1(\mathbf{x})\nabla^2\phi_1(\mathbf{x}) \tag{10.1.197}$$

as the leading equations. We note that the nonlinear term $n_1 a_0^2(\mathbf{x})$ enters in the equation for $\phi_1(\mathbf{x})$, whereas the diffraction term $\nabla^2 a_0(\mathbf{x})$ enters in the equation for $\phi_2(\mathbf{x})$. The preceding equations and all further equations for the $\phi_j(\mathbf{x})$ and $a_j(\mathbf{x})$ are to be solved recursively. The equations (10.1.192) and (10.1.193) for $\phi_0(\mathbf{x})$ and $a_0(\mathbf{x})$ are just the eiconal and transport equations of linear geometrical optics. Each of these equations reduces to an ODE along the rays associated with the phase term $\phi_0(\mathbf{x})$.

Rather than discuss the solution of these equations for a general class of boundary value problems, we consider a specific two-dimensional problem that exhibits some of the basic features of beam propagation in a nonlinear medium.

Propagation of a Beam in a Nonlinear Medium

The (x, y)-plane is assumed to be divided into a *linear region* and a *nonlinear region*. The *interface* between the two regions is the line $x = 0$. A *Gaussian beam* is incident on the nonlinear region $x > 0$ from the linear region $x < 0$. We do not consider the field in the linear region, only in the nonlinear region.

The field $u(0, y)$ equals the incident field $u_I(0, y)$, so that

$$u(0,y) = u_I(0,y) = \frac{1}{\sqrt{k}}\, E\, \exp\left(-\frac{y^2}{\alpha^2}\right), \tag{10.1.198}$$

where E and α are real positive constants (see the foregoing discussion of the propagation of a beam in a linear medium). We take $\mathbf{x}$ to represent x, y and express the nonlinear field $u(x, y)$ in the half-plane $x > 0$ as

$$u(x,y) = \frac{1}{\sqrt{k}}\, a(x,y)\, e^{ik\phi(x,y)}, \tag{10.1.199}$$

where $a(x, y)$ and $\phi(x, y)$ are real valued and are to be expanded in inverse powers of k as in (10.1.191). $u(x, y)$ satisfies the nonlinear reduced wave equation (10.1.176) with n expanded as in (10.1.188). The terms $\phi_j(x, y)$ and $a_j(x, y)$ in the expansion of $\phi(x, y)$ and $a(x, y)$ satisfy (10.1.192)–(10.1.197) for $j = 0$, 1, and 2, and their initial values are found from (10.1.198) to be

$$\phi_j(0,y) = 0,\ j \geq 0, \quad a_0(0,y) = E\exp\left(-\frac{y^2}{\alpha^2}\right),\ a_j(0,y) = 0,\ j \geq 1. \tag{10.1.200}$$

We also require that (10.1.199) be outgoing from the interface $x = 0$ toward $x > 0$, so that $\phi_0(x, y)$ must satisfy the outgoing condition at $x = 0$.

Now $\phi_0(x, y)$ satisfies the eiconal equation (10.1.192), with the initial condition $\phi_0(0, y) = 0$ and the outgoing condition at $x = 0$. Thus, $\phi_0(x, y) = n_0 x$. The equation (10.1.193) for $a_0(x, y)$ then becomes $2n_0\, \partial a_0(x, y)/\partial x = 0$. In view of (10.1.200) we find that $a_0(x, y) = E \exp\left(-y^2/\alpha^2\right)$. To leading order we have

$$u(x, y) \approx \frac{E}{\sqrt{k}} \exp\left(-\frac{y^2}{\alpha^2}\right) e^{ikn_0 x}. \tag{10.1.201}$$

This represents a beam propagating in the x-direction. The *beam half-width* is given as α since for $|y| > \alpha$ the amplitude term of $u(x, y)$ is exponentially small (see the discussion of the propagation of a beam in a linear medium). At this level of approximation neither the *nonlinear* nor the *diffraction effects* have been accounted for and we consider those effects next.

The equation (10.1.194) for $\phi_1(x, y)$, has the form $2n_0(x, y)\partial\phi_1(x, y)/\partial x = n_1 a_0^2(x, y) = n_1 E^2 \exp\left(-2y^2/\alpha^2\right)$. Noting (10.1.200) gives

$$\phi_1(x, y) = \frac{1}{2}\frac{n_1}{n_0} E^2 \exp\left(-\frac{2y^2}{\alpha^2}\right) x. \tag{10.1.202}$$

Thus the phase term $\phi(x, y)$ is given as

$$\phi(x, y) \approx \phi_0(x, y) + \frac{1}{k}\phi_1(x, y) = n_0 x + \frac{1}{2k}\frac{n_1}{n_0} E^2 \exp\left(-\frac{2y^2}{\alpha^2}\right) x. \tag{10.1.203}$$

The rays, which are orthogonal to the wave fronts, have the direction of $\nabla\phi(x, y)$, and this is

$$\nabla\phi(x, y) \approx \left[n_0 + \frac{n_1}{2n_0 k} a_0^2(x, y)\right] \mathbf{i} + \left[-\frac{2n_1 y}{n_0 k\alpha^2} a_0^2(x, y) x\right] \mathbf{j}. \tag{10.1.204}$$

At $y = 0$, $\nabla\phi(x, y)$ has the direction of the vector $\mathbf{i}$. For $y \neq 0$, we see from (10.1.204) that for fixed y as x increases, the $\mathbf{j}$ component of $\nabla\phi(x, y)$ becomes increasingly negative if $y > 0$ and increasingly positive if $y < 0$. At $x = 0$, $\nabla\phi(x, y)$ has no $\mathbf{j}$ component. This indicates the presence of a *self-focusing effect*, in that the rays start bending toward the x-axis as x increases, even though they were originally parallel at $x = 0$. We note that if $n_1 < 0$, the reverse is true and defocusing takes place. We also observe the dependence of the phase on the amplitude in (10.1.203). This is a strictly *nonlinear effect*.

The equation (10.1.195) for $a_1(x, y)$ becomes

$$2n_0 \frac{\partial a_1(x, y)}{\partial x} = \left(\frac{2n_1 E^3 x}{n_0\alpha^2} - \frac{16n_1 E^3 y^2 x}{n_0\alpha^4}\right) \exp\left(-\frac{3y^2}{\alpha^2}\right), \tag{10.1.205}$$

so that

$$a_1(x, y) = \left(\frac{n_1 E^3 x^2}{2n_0^2\alpha^2} - \frac{4n_1 E^3 y^2 x^2}{n_0^2\alpha^4}\right) \exp\left(-\frac{3y^2}{\alpha^2}\right), \tag{10.1.206}$$

since $a_1(0, y) = 0$. This can be written as

$$a_1(x, y) = \left(\frac{n_1 E^2 x^2}{2n_0^2 \alpha^2} - \frac{4n_1 E^2 y^2 x^2}{n_0^2 \alpha^4} \right) \exp\left(-\frac{2y^2}{\alpha^2} \right) a_0(x, y), \qquad (10.1.207)$$

and we obtain

$$a(x, y) \approx \left[1 + \left(\frac{n_1 E^2 x^2}{2kn_0^2 \alpha^2} - \frac{4n_1 E^2 y^2 x^2}{kn_0^2 \alpha^4} \right) \exp\left(-\frac{2y^2}{\alpha^2} \right) \right] E \exp\left(-\frac{y^2}{\alpha^2} \right). \qquad (10.1.208)$$

The presence of terms of order x^2 in (10.1.207) and, consequently, in (10.1.208) shows that $a_1(x, y)$ undergoes *secular growth* as $x \to \infty$. Since the exponential term in (10.1.207) is small for large $|y|$ and the term of order y^2 is small for small $|y|$, we conclude that the first term in parentheses in (10.1.207) is the most significant for large x. With x fixed, the term that corresponds to it in (10.1.208) attains its maximum at $y = 0$. Thus it can be said that when $n_1 E^2 x^2 / 2kn_0^2 \alpha^2 \approx 1$, we have the first value of x at which $a_0(x, y)$ and $(1/k)a_1(x, y)$ are essentially of the same order of magnitude and the asymptotic series for $a(x, y)$ begins to become disordered. We denote the value of x determined from the above by x_f and we see that

$$x_f \approx \frac{\alpha n_0 \sqrt{2k}}{E\sqrt{n_1}}. \qquad (10.1.209)$$

Apart from a (dimensionless) constant factor, this determines the value of x at which the initially plane wave is expected to focus, and we refer to it as a *focal length.* We associate the region where secular growth becomes significant with the onset of the self-focusing region for the nonlinear medium. We now demonstrate that such a relationship does exist.

Since our geometrical optics result for $\phi(x, y)$ and $a(x, y)$ breaks down when $x = O(\sqrt{k})$, as follows from (10.1.209), we study that region using a boundary layer approach. It is easily seen from (10.1.196) that $\phi_2(x, y)$ contains a term of the order x^3. Thus $(1/k)\phi_1(x, y) + (1/k^2)\phi_2(x, y)$ are of the same order of magnitude when $x = O(\sqrt{k})$ since $\phi_1(x, y) = O(x)$. However, the leading term $\phi_0(x, y) = n_0 x$ is never of the same order of magnitude as the other terms in the expansion of $\phi(x, y)$. Thus secular behavior occurs in the expansions of both $\phi(x, y)$ and $a(x, y)$, but $\phi_0(x, y)$ is not involved in it. As a result we shall modify only the series expansion of $\phi(x, y)$ beginning with the $\phi_1(x, y)$ term. Since $(1/k)\phi_1(x, y) = O(1/\sqrt{k})$ in the secular region $x = O(\sqrt{k})$, we set $\phi(x, y) = n_0 x + 1/\sqrt{k}\ \hat{\phi}(x, y)$ in (10.1.189)–(10.1.190), where we set $r = \frac{1}{2}$. Then

$$\left\{ \frac{2n_0}{\sqrt{k}} \frac{\partial \hat{\phi}(x, y)}{\partial x} a(x, y) + \frac{1}{k} (\nabla \hat{\phi}(x, y))^2 a(x, y) - \frac{1}{k} n_1 a^3(x, y) + \cdots \right\}$$

$$- \frac{1}{k^2} \nabla^2 a(x, y) = 0, \qquad (10.1.210)$$

$$2n_0\frac{\partial a(x,y)}{\partial x}+\frac{2}{\sqrt{k}}\nabla\hat{\phi}(x,y)\cdot\nabla a(x,y)+\frac{1}{\sqrt{k}}\,a(x,y)\nabla^2\hat{\phi}(x,y)=0. \quad (10.1.211)$$

Since $x=O(\sqrt{k})$ in the secular region we set $x=\sqrt{k}\,\xi$ in (10.1.210)–(10.1.211). Then $\partial/\partial x=(1/\sqrt{k})\partial/\partial\xi$, so that to leading order

$$\frac{1}{k}\left\{2n_0\,\frac{\partial\hat{\phi}(\xi,y)}{\partial\xi}+\left(\frac{\partial\hat{\phi}(\xi,y)}{\partial y}\right)^2-n_1\hat{a}^2(\xi,y)\right\}\hat{a}(\xi,y)+O\left(\frac{1}{k^2}\right)=0, \quad (10.1.212)$$

$$\frac{1}{\sqrt{k}}\left\{2n_0\,\frac{\partial\hat{a}(\xi,y)}{\partial\xi}+2\,\frac{\partial\hat{\phi}(\xi,y)}{\partial y}\frac{\partial\hat{a}(\xi,y)}{\partial y}+\hat{a}(\xi,y)\,\frac{\partial^2\hat{\phi}(\xi,y)}{\partial y^2}\right\}+O\left(\frac{1}{k}\right)=0. \quad (10.1.213)$$

To solve, we expand $\hat{a}(\xi,y)$ and $\hat{\phi}(\xi,y)$ in powers of $1/\sqrt{k}$ and obtain to leading order for $\hat{a}_0(\xi,y)$ and $\hat{\phi}_0(\xi,y)$,

$$2n_0\,\frac{\partial\hat{\phi}_0(\xi,y)}{\partial\xi}+\left(\frac{\partial\hat{\phi}_0(\xi,y)}{\partial y}\right)^2-n_1\hat{a}_0^2(\xi,y)=0, \quad (10.1.214)$$

$$2n_0\,\frac{\partial\hat{a}_0(\xi,y)}{\partial\xi}+2\,\frac{\partial\hat{\phi}_0(\xi,y)}{\partial y}\frac{\partial\hat{a}_0(\xi,y)}{\partial y}+\hat{a}_0(\xi,y)\frac{\partial^2\hat{\phi}_0(\xi,y)}{\partial y^2}=0. \quad (10.1.215)$$

Although these equations are coupled as is the case for the nonlinear geometrical optics equations (10.1.181), they are, nevertheless, easier to solve.

A number of solutions of (10.1.214)–(10.1.215) have been obtained by *Akhmanov* and others. We shall find one of these solutions and match it with our geometrical optics results for $a_0(x,y)$ and $\phi_1(x,y)$. To carry this out, we require that the field be considered only near the center of the beam (i.e., near $y=0$). This yields the approximations

$$\phi_1(x,y)=\frac{1}{2}\,\frac{n_1}{n_0}\,E^2\exp\left(-\frac{2y^2}{\alpha^2}\right)x\approx\frac{1}{2}\,\frac{n_1}{n_0}\,E^2x-\frac{n_1E^2}{n_0\alpha^2}\,y^2x \quad (10.1.216)$$

$$a_0^2(x,y)=E^2\exp\left(-\frac{2y^2}{\alpha^2}\right)\approx E^2\left(1-\frac{2y^2}{\alpha^2}\right). \quad (10.1.217)$$

An exact solution of (10.1.214)–(10.1.215) can be found that has the form

$$\hat{\phi}_0(\xi,y)=\gamma(\xi)+\frac{1}{2}y^2\beta(\xi),\quad \hat{a}_0^2(\xi,y)=\frac{E^2}{f(\xi)}\left(1-\frac{2y^2}{\alpha^2f^2(\xi)}\right), \quad (10.1.218)$$

where $\gamma(\xi)$, $\beta(\xi)$, and $f(\xi)$ are to be specified on substituting (10.1.218) into (10.1.214)–(10.1.215). We match $\hat{\phi}_0(\xi,y)$ and $\hat{a}_0^2(\xi,y)$ with $\phi_1(x,y)$ and $a_0^2(x,y)$ by noting that $x=\sqrt{k}\,\xi$, so that for small ξ we have moderate or large values of x. On comparing both sets of equations, we find that $\gamma(\xi)$ and $\beta(\xi)$ must tend to zero as

$\xi \to 0$, while $f(\xi) \to 1$ as $\xi \to 0$. Thus we set $\gamma(0) = \beta(0) = 0$, $f(0) = 1$. Before inserting (10.1.218) into (10.1.214)–(10.1.215) we multiply (10.1.215) by $\hat{a}_0(\xi, y)$ to simplify the calculations. We obtain from (10.1.214) and from the modified form of (10.1.215), respectively,

$$n_0 y^2 \beta'(\xi) + 2n_0 \gamma'(\xi) + y^2 \beta^2(\xi) - \frac{n_1 E^2}{f(\xi)} + \frac{2n_1 E^2 y^2}{\alpha^2 f^3(\xi)} = 0, \tag{10.1.219}$$

$$\frac{E^2 \beta(\xi)}{f(\xi)} - \frac{6E^2 y^2 \beta(\xi)}{\alpha^2 f^3(\xi)} - \frac{n_0 E^2 f'(\xi)}{f^2(\xi)} + \frac{6n_0 E^2 y^2 f'(\xi)}{\alpha^2 f^4(\xi)} = 0. \tag{10.1.220}$$

Now if we set $\gamma'(\xi) = n_1 E^2 / 2n_0 f(\xi)$ in (10.1.219), the two terms in that equation that do not contain a factor y^2 cancel out. We are left with

$$n_0 \beta'(\xi) + \beta^2(\xi) + \frac{2n_1 E^2}{\alpha^2 f^3(\xi)} = 0. \tag{10.1.221}$$

Similarly, if we set

$$\beta(\xi) = \frac{n_0 f'(\xi)}{f(\xi)}, \tag{10.1.222}$$

the two terms in (10.1.220) that do not contain a factor y^2 drop out. With this choice of $\beta(\xi)$ the remaining terms in (10.1.220) cancel out, so that this equation is satisfied regardless of the form of $f(\xi)$. Therefore, $f(\xi)$ must be specified from (10.1.221). Inserting (10.1.222) into (10.1.221) yields

$$f''(\xi) = -\frac{2n_1 E^2}{n_0^2 \alpha^2 f^2(\xi)}. \tag{10.1.223}$$

An implicit solution of (10.1.223) with $f(0) = 1$, is

$$f(\xi) = \sin^2 \left[\frac{\pi}{2} + \sqrt{f'(\xi)[1 - f(\xi)]} + \frac{E\sqrt{2n_1}\xi}{\alpha n_0} \right]. \tag{10.1.224}$$

Then $\gamma(\xi)$ and $\beta(\xi)$ can be specified, but we do not carry this out.

Since $f(\xi)$ appears in the denominator in the expression (10.1.218) for $\hat{a}_0^2(\xi, y)$, we see that when $f(\xi) = 0$, the amplitude blows up. This corresponds to a *focal point* for the field. To determine the location of this focal point, we set $f(\xi) = 0$ in (10.1.224) and obtain as the (first) focal length

$$\hat{x}_f = \sqrt{k}\, \xi_f = \frac{\pi}{2} \frac{\alpha n_0 \sqrt{k}}{E\sqrt{2n_1}} \tag{10.1.225}$$

since $\sin(\pi) = 0$. There are, in fact, infinitely many focal points predicted by (10.1.224), but the others are of no interest to us.

We want to match the preceding *boundary layer results* with the previously given *geometrical optics results*. This requires that we consider $\gamma(\xi)$, $\beta(\xi)$, and $f(\xi)$ for small ξ. Since $f(0) = 1$, $f'(0) = 0$, and $f''(0) = -2n_1 E^2 / n_0^2 \alpha^2$ in view

of (10.1.223), we have $f(\xi) = f\left(x/\sqrt{k}\right) \approx 1 - n_1 E^2 x^2/n_0^2\alpha^2 k$. As a result we find that $\beta(\xi) = \beta\left(x/\sqrt{k}\right) \approx -2n_1 E^2 x/n_0\alpha^2\sqrt{k}$, $\gamma(\xi) = \gamma\left(x/\sqrt{k}\right) \approx n_1 E^2 x/2n_0\sqrt{k}$. Inserting these expressions into (10.1.216) yields

$$\phi(x,y) \approx n_0 x + \frac{1}{k}\left(\frac{n_1 E^2 x}{2n_0} - \frac{n_1 E^2 y^2 x}{n_0\alpha^2}\right), \tag{10.1.226}$$

which agrees with the geometrical optics result (10.1.203) when that is evaluated near the axis of the beam as in (10.1.217).

For $\hat{a}_0(\xi, y)$ we have

$$\hat{a}_0(\xi,y) \approx E\left(1 - \frac{n_1 E^2 x^2}{n_0^2\alpha^2 k}\right)^{-1/2}\left(1 - \frac{2y^2}{\alpha^2}\right) \approx E\left(1 + \frac{n_1 E^2 x^2}{2n_0^2\alpha^2 k}\right) \tag{10.1.227}$$

for small y. This agrees with the expression (10.1.208) for $a(x, y)$ near the axis of the beam. Furthermore, if we determine the focal point for $\hat{a}_0(\xi, y)$ to the level of approximation used in (10.1.227), we find that the *focal length* is $\hat{x}_f \approx \alpha n_0\sqrt{k}/E\sqrt{n_1}$, since $f(\xi)$ vanishes at this point if it is approximated as above. Comparing this value with (10.1.225), we find that they differ by a factor of $\pi/2\sqrt{2} \approx 1.1$. The difference is not substantial. The result given in (10.1.209) on the basis of the geometrical optics approach differs from the above by a factor of $\sqrt{2}$. We see from (10.1.227) that this difference is due to the factor $\frac{1}{2}$, which occurs when the binomial expansion is used in (10.1.227). Nevertheless, all expressions for the focal length contain the same dependence on the parameters α, n_0, n_1, E, and k.

Even though the boundary layer equations (10.1.214)–(10.1.215) predict that focusing occurs, as we have seen, it may yet be that focusing is prevented if diffraction effects, which have so far been neglected, are taken into account. That is, we must consider equations for the field that include second derivative terms in the amplitude. The self-trapped solution (10.1.187) shows that plane wave solutions (in an approximate form) do exist that do not undergo self-focusing. To analyze this problem, we now obtain the terms $\phi_2(x, y)$ and $a_2(x, y)$ in the expansion of the geometrical optics solution (10.1.199).

The phase term $\phi_2(x, y)$ satisfies (10.1.196). Using the expressions obtained above for $\phi_0(x, y)$, $\phi_1(x, y)$, $a_0(x, y)$, and $a_1(x, y)$ and the initial condition $\phi_2(0, y) = 0$, we easily obtain

$$\phi_2(x,y) = \frac{2xy^2 - x\alpha^2}{n_0\alpha^4} + \left[\frac{(4n_2n_0^2 - n_1^2)x}{8n_0^3} + \frac{(\alpha^2 - 12y^2)n_1^2x^3}{6n_0^3\alpha^4}\right]E^4\exp\left(-\frac{4y^2}{\alpha^2}\right). \tag{10.1.228}$$

We insert (10.1.228) into the equation (10.1.197) for $a_2(x, y)$ and integrate, retaining only the most significant terms near $y = 0$ and for x large. Then

$$a_2(x,y) \approx -\frac{x^2E}{n_0^2\alpha^4}\exp\left(-\frac{y^2}{\alpha^2}\right) + \left(\frac{n_1^2x^4E^5}{2n_0^4\alpha^4} + \frac{n_2x^2E^5}{n_0^2\alpha^2}\right)\exp\left(-\frac{5y^2}{\alpha^2}\right). \tag{10.1.229}$$

In the absence of nonlinear effects (i.e., with $n_1 = n_2 = 0$) the first term on the right in (10.1.229) is the only contribution to $a_2(x, y)$. It represents a *diffraction effect*. Since $a_1(x, y) = 0$ if $n_1 = 0$, we have $a(x, y) \approx a_0(x, y) + (1/k^2)a_2(x, y) = \left(1 - x^2/k^2 n_0^2 \alpha^4\right) E \exp\left(-y^2/\alpha^2\right)$. Thus for a linear medium, secular effects occur when

$$x_d \approx k n_0 \alpha^2, \tag{10.1.230}$$

where x_d is the diffraction length [it agrees with (10.1.172) if we put $\alpha^2 = 2a$]. As shown in our discussion of the propagation of a Gaussian beam in a linear medium when $x = O(k)$, the Gaussian beam field must be described by means of a *Schrödinger equation*. However, since diffraction effects occur when $x = 0(k)$ and self-focusing occurs when $x = O(\sqrt{k})$ in view of (10.1.209), it appears that diffraction cannot, in general, influence and counteract self-focusing.

When n_1 and n_2 are not equal to zero, we obtain near $y = 0$,

$$a(x, y) \approx a_0(x, y) + \frac{1}{k} a_1(x, y) + \frac{1}{k^2} a_2(x, y) = E + \frac{n_1 E^3 x^2}{2 n_0^2 \alpha^2 k} - \frac{E x^2}{n_0^2 \alpha^4 k^2} + \frac{n_2 E^5 x^2}{n_0^2 \alpha^2 k^2}. \tag{10.1.231}$$

Secular effects occur when $(1/k)a_1(x, y) \approx a_0(x, y)$ in the nonlinear problem and when $(1/k^2)a_2(x, y) \approx a_0(x, y)$ in the linear problem. The combined effects in (10.1.231) for which $a_0(x, y) \approx (1/k)a_1(x, y) \approx (1/k^2)a_2(x, y)$ yield

$$\left(\frac{n_1 E^2}{2 n_0^2 \alpha^2 k} + \frac{n_2 E^4}{n_0^2 \alpha^2 k^2} - \frac{1}{n_0^2 \alpha^4 k^2} \right) x^2 \approx 1. \tag{10.1.232}$$

Noting (10.1.209) and (10.1.230), we have

$$\frac{1}{x^2} \approx \frac{1}{x_f^2} - \frac{1}{x_d^2} + \left(\frac{E^2 \sqrt{n_2}}{n_0 \alpha k} \right)^2, \tag{10.1.233}$$

and we assume that $n_2 > 0$. Arguing as above, we could state that (10.1.233) determines a first value of x at which self-focusing becomes significant if we include diffraction and higher-order nonlinear effects. We have established a relationship between the secularity and self-focusing regions previously. Thus (10.1.233) shows that diffraction need not prevent self-focusing, but it does delay its onset. The higher-order nonlinear effect enhances the self-focusing property and brings the focal region closer to the origin. This result is not unexpected in view of our general discussion of the self-focusing effect. If k is large and the other terms in (10.1.232) are of moderate size, the diffraction and higher order nonlinear effects do not appear to be significant.

An interesting situation arises if $x_f = x_d$, so that the focal and diffraction lengths are equal. This implies that

$$E^2 = \frac{2}{n_1 \alpha^2 k}, \tag{10.1.234}$$

so that E is small for large k. With this choice for E, the last term in (10.1.231) is of order $k^{-4}E$, so that

$$a(x,y) \approx a_0(x,y) + \frac{1}{k}\, a_1(x,y) + \frac{1}{k^2}\, a_2(x,y) = E + O(k^{-4}E). \quad (10.1.235)$$

Also, in (10.1.232) we may eliminate the n_2 term since it is of order k^{-4} compared to the other two bracketed terms, which are of order k^{-2}. Consequently, (10.1.233) becomes $1/x^2 \approx 0$, and the focal point is moved off to infinity at this level of approximation. Furthermore, we have $\phi(x,y) \approx \phi_0(x,y) + (1/k)\phi_1(x,y) + (1/k^2)\phi_2(x,y) = n_0 x + O(k^{-4})$ for $y \approx 0$, as is easily seen. Thus with E given as in (10.1.234), we obtain

$$u(x,y) = \frac{1}{\sqrt{k}}\, a(x,y) e^{ik\phi(x,y)} \approx \left(\frac{2}{n_1 \alpha^2 k}\right)^{1/2} \exp\left(-\frac{y^2}{\alpha^2}\right) e^{ikn_0 x}, \quad (10.1.236)$$

using the given approximation. We have replaced $a(x,y) \approx E$ by $a(x,y) \approx E\exp(-y^2/\alpha^2)$ in (10.1.236) since the preceding was valid near $y = 0$, and the field, which has the form of a beam, is exponentially small away from $y = 0$.

We have shown that if E is given by (10.1.234), the resulting field is a beam that travels without change of shape to a high level of approximation for large k. Since this is effectively a self-trapped beam for which an (exact) balance between the nonlinear self-focusing effect and the defocusing diffraction effect exists, it is of interest to compare (10.1.236) with the approximate form (10.1.187) of the exact self-trapped solution found earlier. The parameter c that occurs in (10.1.187) may be interpreted as an inverse beam half-width for that solution, since when $cy > 1$, sech(cy) is small. Thus if $c = 1/\alpha$, the amplitude term in (10.1.187) evaluated at the center of the beam is equal to E as defined in (10.1.234). We recall that α is the half-width of the beam (10.1.236). Although the beams (10.1.187) and (10.1.236) have somewhat different phase and amplitude representations, they have a common beam half-width α and have the same amplitude E at the center of the beam. In addition, they are both approximate general plane waves and are self-trapped.

We now present a boundary layer analysis of the field in the region $x = O(k)$, where diffraction effects become important. The phase in that region should equal $n_0 x$ plus lower-order correction terms, and the amplitude should be bounded or decreasing as k increases. From (10.1.202) and (10.1.206) we see that $(1/k)\phi_1(x,y) = O(1)$ and $(1/k)a_1(x,y) = O(k)$ when $x = O(k)$. However, $\phi_0(x,y) = O(k)$ and $a_0(x,y) = O(1)$ when $x = O(k)$. Since we cannot have $(1/k)a_1(x,y)$ be of higher order than $a_0(x,y)$, it must be assumed that $a_0(x,y) = E\exp(-y^2/a^2) = O(k^{-s})$ for some positive s, and this means that $E = O(k^{-s})$. Then $(1/k)\phi_1(x,y) = O(k^{-2s})$ and $(1/k)a_1(x,y) = O(k^{1-3s})$, so that these correction terms are of reduced orders in k.

To proceed we set $\phi(x,y) = n_0 x + (1/k)\tilde{\phi}(x,y)$ in (10.1.189)–(10.1.190), since when $x = O(k)$ we have seen that the correction term to $\phi_0(x,y) = n_0 x$ is one order of k lower than ϕ_0 itself. With $r = \frac{1}{2}$ in (10.1.189)–(10.1.190), we have $\{(2n_0/k)\partial\tilde{\phi}(x,y)/\partial x\, a(x,y) + (1/k^2)(\nabla\tilde{\phi}(x,y))^2 a(x,y) - (n_1/k)a^3(x,y) + \cdots\} - \nabla^2 a(x,y)/k^2 = 0$, and $2n_0\, \partial a(x,y)/\partial x + (2/k)\,\nabla\tilde{\phi}(x,y) \cdot \nabla a(x,y) + (1/k)\, a(x,y)\nabla^2\tilde{\phi}(x,y) = 0$. Then with $x = k\xi$, we set $a(x,y) = (1/k^s)\hat{a}(\xi,y)$,

and $\tilde{\phi}(x,y) = (1/k^{2s-1})\hat{\phi}(\xi,y)$, where $s > 0$ and is to be determined. This implies that the amplitude $a(x,y)$ is $O(k^{-s})$ and the correction term to $\phi_0(x,y) = n_0 x$ is $O(k^{-2s})$, as was required previously. We obtain

$$\frac{1}{k^{1+3s}}\left[2n_0\frac{\partial\hat{\phi}(\xi,y)}{\partial\xi}\hat{a}(\xi,y) + \left(\frac{\partial\hat{\phi}(\xi,y)}{\partial y}\right)^2\hat{a}(\xi,y) - n_1\hat{a}^3(\xi,y)\right] - \frac{1}{k^{2+s}}\frac{\partial^2\hat{a}(\xi,y)}{\partial y^2} + \cdots = 0, \quad (10.1.237)$$

$$\frac{1}{k^{1+s}}\left[2n_0\frac{\partial\hat{a}(\xi,y)}{\partial\xi}\right] + \frac{1}{k^{3s}}\left[2\frac{\partial\hat{\phi}(\xi,y)}{\partial y}\frac{\partial\hat{a}(\xi,y)}{\partial y} + \hat{a}(\xi,y)\frac{\partial^2\hat{\phi}(\xi,y)}{\partial y^2}\right] + \cdots = 0, \quad (10.1.238)$$

where we have neglected lower-order terms. We must choose $s = \frac{1}{2}$ so that the significant terms in (10.1.237)–(10.1.238) include the leading terms in (10.1.212)–(10.1.213) valid in the focal region. However, we do have an additional second derivative term in the amplitude to account for diffraction.

The leading terms in the expansions of $\hat{\phi}(\xi,y)$ and $\hat{a}(\xi,y)$, which we denote by $\hat{\phi}_0(\xi,y)$ and $\hat{a}_0(\xi,y)$, satisfy

$$2n_0\frac{\partial\hat{\phi}_0(\xi,y)}{\partial\xi} + \left(\frac{\partial\hat{\phi}_0(\xi,y)}{\partial y}\right)^2 - n_1\hat{a}_0^2(\xi,y) - \frac{1}{\hat{a}_0(\xi,y)}\frac{\partial^2\hat{a}_0(\xi,y)}{\partial y^2} = 0, \quad (10.1.239)$$

$$2n_0\frac{\partial\hat{a}_0(\xi,y)}{\partial\xi} + 2\frac{\partial\hat{\phi}_0(\xi,y)}{\partial y}\frac{\partial\hat{a}_0(\xi,y)}{\partial y} + \hat{a}_0(\xi,y)\frac{\partial^2\hat{\phi}_0(\xi,y)}{\partial y^2} = 0. \quad (10.1.240)$$

Introducing the function

$$\hat{v}(\xi,y) = \hat{a}_0(\xi,y)\exp[i\hat{\phi}_0(\xi,y)], \quad (10.1.241)$$

we easily find that

$$2in_0\frac{\partial\hat{v}(\xi,y)}{\partial\xi} + \frac{\partial^2\hat{v}(\xi,y)}{\partial y^2} + n_1|\hat{v}(\xi,y)|^2\hat{v}(\xi,y) = 0. \quad (10.1.242)$$

This *nonlinear Schrödinger equation* reduces to the *linear Schrödinger equation* (10.1.103) or (10.1.175) when $n_1 = 0$. Both equations describe the far field in the region $x = O(k)$ where diffraction effects are important. In terms of $u(x,y)$ our result is now given as

$$u(x,y) \approx \frac{1}{k}\hat{v}\left(\frac{1}{k}x, y\right)e^{ikn_0x}, \quad (10.1.243)$$

where $\hat{v}(x/k,y)$ satisfies the equation

$$2ikn_0\frac{\partial\hat{v}(x/k,y)}{\partial x} + \frac{\partial^2\hat{v}(x/k,y)}{\partial y^2} + n_1|\hat{v}(x/k,y)|^2\hat{v}(x/k,y) = 0. \quad (10.1.244)$$

The *nonlinear Schrödinger equation* (10.1.242) is one of a class of nonlinear equations for which an exact solution of the initial value problem can be found. The inverse scattering method and related methods that yield these solutions are too lengthy and complicated to be considered here. A *self-trapped solution* of (10.1.242) may be found by setting $\hat{\phi}_0(\xi, y) = \hat{\lambda}\xi$, $\hat{a}_0(\xi, y) = \hat{a}_0(y)$ in (10.1.239)–(10.1.240). Then (10.1.240) is identically satisfied and (10.1.239) reduces to the ODE $2n_0\hat{\lambda}\hat{a}_0(y) - n_1\hat{a}_0^3(y) - \hat{a}_0''(y) = 0$. This has the form of (10.1.183), and a solution that vanishes as $|y| \to \infty$ is $\hat{a}_0(y) = \left[4n_0\hat{\lambda}/n_1\right]^{1/2} \operatorname{sech}[(2n_0\hat{\lambda})^{1/2}y]$, in view of (10.1.184).

The field $u(x, y)$ is given as

$$u(x, y) \approx \left(\frac{4n_0\hat{\lambda}}{n_1k^2}\right)^{1/2} \operatorname{sech}[(2n_0\hat{\lambda})^{1/2}y] \exp\left[i\left(kn_0x + \frac{\hat{\lambda}x}{k}\right)\right] \tag{10.1.245}$$

on using (10.1.241) and (10.1.243). If we set $\hat{\lambda} = c^2/2n_0$, (10.1.245) is identical to (10.1.187). We are interested in a solution of (10.1.239)–(10.1.240) that matches (10.1.236). Thus, we set $\hat{\phi}_0(\xi, y) = 0$, $\hat{a}_0(\xi, y) = \hat{E}\exp\left(-y^2/\alpha^2\right)$, with $\hat{E}$ to be determined. Inserting this into (10.1.239)–(10.1.240) yields $-(2\hat{E}/\alpha^2)\left(1 - 2y^2/\alpha^2\right)$ 10.1.5 $\exp\left(-y^2/\alpha^2\right) = -n_1\hat{E}^3\exp\left(-3y^2/\alpha^2\right)$, as (10.1.240) is satisfied identically. Near the center of the beam where $y \approx 0$, we may approximate $\exp[-2y^2/\alpha^2]$ by $1 - 2y^2/\alpha^2$. Doing so in the above shows that $\hat{E}$ must be given as $\hat{E} = 2/n_1\alpha^2$. Thus the field $u(x, y)$ near $y = 0$ is

$$u(x, y) \approx \left(\frac{2}{n_1\alpha^2k^2}\right)^{1/2} \exp\left(-\frac{y^2}{\alpha^2}\right) e^{ikn_0x}. \tag{10.1.246}$$

The expression (10.1.246) for $u(x, y)$ agrees with (10.1.236), which was obtained from the geometrical optics solution under the assumption that the self-focusing and diffraction lengths are equal. The field (10.1.246) represents a self-trapped beam since the amplitude is independent of x and decays exponentially in $|y|$. If we define $(2n_0\hat{\lambda})^{-1/2} = \alpha$ to be the beam half-width of (10.1.245), we obtain for (10.1.245)

$$u(x, y) \approx \left(\frac{2}{n_1\alpha^2k^2}\right)^{1/2} \operatorname{sech}\left(\frac{y}{\alpha}\right) \exp\left[i\left(kn_0x + \frac{x}{2n_0\alpha^2k}\right)\right]. \tag{10.1.247}$$

The fact that (10.1.246) and (10.1.247) are both self-trapped beams of half-width α and $|u(x, 0)|^2 = 2/n_1\alpha^2k^2 \equiv E_0^2$, suggests that E_0 plays a critical role in determining whether beams are self-focused, self-trapped, or effectively propagate as in a linear medium. If a beam $u(x, y)$ has $|u(x, 0)| > E_0$, the focusing effect is stronger than the defocusing diffraction effects and self-focusing occurs. If $|u(x, 0)| = E_0$, self-trapping occurs, and if $|u(x, 0)| < E_0$, the nonlinearity contributes small corrections to linear wave propagation. The concept of a *critical intensity* E_0^2 associated with beam propagation has been discussed in the literature, and its properties are somewhat consistent with these results. The validity of this relationship is difficult to verify in general.

Exercises 10.1

10.1.1. Let $u_I(x,y,z) = (1/4\pi r)e^{iknr}$ with $r^2 = (x-\xi)^2 + (y-\eta)^2 + (z-\zeta)^2$ be the free-space Green's function for the reduced wave equation (10.1.1) in (x,y,z)-space (n = constant) with source point in $z > 0$. Use the method of images to construct the Green's function for the half-space $z > 0$ in the form $K(x,y,z;\xi,\eta,\zeta) = u_I(x,y,z) + u_S(x,y,z)$, if $K(x,y,0;\xi,\eta,\zeta) = 0$. Note that both $u_I(x,y,z)$ and $u_S(x,y,z)$ are of the form (10.1.7) and show that the phase and amplitude terms of $u_I(x,y,z)$ and $u_S(x,y,z)$ are equal on $z = 0$. Also show that $u_S(x,y,z)$ satisfies the outgoing condition at $z = 0$.

10.1.2. For the two-dimensional problem (10.1.14)–(10.1.15), determine the Green's function $K(x,y;\xi,\eta)$ if the Dirichlet condition (10.1.21) is replaced by the Neumann condition $\partial K(0,y;\xi,\eta)/\partial x = 0$. *Hint*: Use the method of images. Expand the result asymptotically and show how the leading amplitude term and the phase term of the scattered wave $u_S(x,y)$ can be specified by applying the boundary condition to the asymptotic expansion of the solution.

10.1.3. (a) Apply the method of images to construct the Green's function for (10.1.1) (n = constant) in the quarter-plane $x > 0$, $y > 0$ if the Green's function is required to vanish on the boundary. *Hint*: If (ξ,η) is the source point for the free-space Green's function in the first quadrant, there are image sources at $(-\xi,\eta)$, $(-\xi,-\eta)$, and $(\xi,-\eta)$. (b) Expand the Green's function asymptotically and show how the leading terms may be obtained by expressing the incident and reflected waves in the asymptotic form (10.1.7). *Hint*: Consider that the incident rays may undergo a double reflection from $x = 0$ and $y = 0$.

10.1.4. Solve the problem of Exercise 10.1.3 if the Green's function $K(x,y;\xi,\eta)$ satisfies the boundary conditions: $\partial K(0,y;\xi,\eta)/\partial x = 0$, $K(x,0;\xi,\eta) = 0$.

10.1.5. Consider the equation $u''(x) + k^2n^2(x)u(x) = 0$, which is the one-dimensional version of (10.1.1). Obtain an asymptotic solution of this equation for large k by expanding $u(x)$ as in (10.1.7) and (10.1.11). Determine and solve the equations for the phase and amplitude terms. (This technique in one dimension is known as the *WKB method*, and a large body of results is available for this method in the literature.)

10.1.6. Determine the rays and the phase function $\phi(x,y)$ if $\phi(x,y) = 1$ on the circular cylinder $x^2 + y^2 = a^2$ and the index of refraction $n(x,y) = 1$. Find both sets of rays and phase functions.

10.1.7. Find the rays and the phase function $\phi(x,y)$ in the two-dimensional case if we set $\phi(x,y) = \theta$ (the polar angle in polar coordinates $x = r\cos\theta$, $y = r\sin\theta$) on the unit circle $x^2 + y^2 = 1$. Also solve the leading order transport equation (10.1.54) for $v_0(x,y)$ in this case. [Let $n(x,y) = 1$.]

10.1.8. Show that for a cylindrical wave with phase $\phi = nr = n\sqrt{x^2+y^2}$ and a constant refractive index n, the amplitude term v_0 decays like $1/\sqrt{r}$, by using the relation (10.1.70).

10.1.9. Consider the asymptotic scattering problem for $u(\mathbf{x}) = u_I(\mathbf{x}) + u_S(\mathbf{x})$, where $u_I(\mathbf{x})$ and $u_S(\mathbf{x})$ are given as in (10.1.80)–(10.1.81). Determine the appropriate conditions on the phase term $\phi(\mathbf{x})$ and the amplitude terms $v_j(\mathbf{x})$ if the following boundary conditions are given on the scattering surface: (a) $\partial u(\mathbf{x})/\partial n = 0$; (b) $\partial u(\mathbf{x})/\partial n + hu(\mathbf{x}) = 0$.

10.1.10. Obtain the geometrical optics solution of the problem of the reflection of a cylindrical wave by a parabola considered in the text, if the field $u(x, y)$ satisfies the Neumann condition $\partial u(x, y)/\partial n = 0$ on the parabola.

10.1.11. Assume that the plane wave $u_I(x, y) = e^{ikx}$ is incident on the parabola (10.1.84) from the left and that the total field $u(x, y) = u_I(x, y) + u_S(x, y)$ vanishes on the parabola. Determine the asymptotic expansion of the scattered wave $u_S(x, y)$.

10.1.12. Assume that the plane wave $u_I(x, y) = e^{iky}$ is incident on the parabola (10.1.84) and that the total field $u(x, y) = u_I(x, y) + u_S(x, y)$ vanishes on the parabola. Determine the geometrical optics expansion of the scattered field $u_S(x, y)$. Obtain only the leading amplitude terms for $u_S(x, y)$ and leave it in parametric form. Show that $u_S(x, y)$ is singular at the vertex of the parabola and that the field is discontinuous across the shadow boundary $x = -a/2,\ y > 0$.

10.1.13. Rework the problem of the asymptotic expansion at a caustic under the assumption that the caustic curve is the circle $x^2 + y^2 = a^2$ and that $V_0(\sigma_0) = 1$. Obtain expressions for the incident and outgoing geometrical optics fields and for the field in the boundary layer in terms of variables associated with the circular caustic. Determine the asymptotic value of the field on the caustic $x^2 + y^2 = a^2$.

10.1.14. Consider the problem of the scattering by a half-plane as presented in the text. Introduce a boundary layer at the reflection boundary $y = 0,\ z < 0$ and proceed as in the text to obtain a boundary layer solution that yields a smooth transition across the reflection boundary.

10.1.15. Construct the geometrical optics solution for the problem of the scattering by a half-plane if the boundary condition in (10.1.136) is replaced by the Neumann condition $\partial u(x, y, 0)/\partial z = 0$ for $y < 0$.

10.1.16. In two dimensions, consider the problem of diffraction by a slit. Let the incident plane wave $u_I(x, y) = e^{ikx}$ approach the y-axis, which has an opening or slit for $|y| < 1$. For $x = 0$ and $|y| > 1$, the boundary condition for the total field $u(x, y) = u_I(x, y) + u_S(x, y)$ is given as $u(0, y) = 0,\ |y| > 1$. Using appropriate outgoing conditions, determine the geometrical optics field and show that there are two shadow and reflection boundaries. Show how boundary layer methods can be used to obtain smooth transitions across these shadow and reflection boundaries. Compare these results with those that might be obtained by applying the geometrical theory of diffraction at each edge—that is, at $(0, -1)$ and $(0, 1)$.

10.1.17. Use the parabolic equation $2ikv_x(x, y) + v_{yy}(x, y) = 0,\ x > 0$, with the data $v(0, y) = \begin{cases} 1, & |y| < 1, \\ 0, & |y| > 1, \end{cases}$ to obtain an approximate solution to the problem of Exercise 10.1.16 in the half-plane $x > 0$. Compare the solution obtained by the

parabolic equation method with that obtained in Exercise 10.1.16 using geometrical optics and boundary layer methods.

10.1.18. Obtain the geometrical optics solution of the problem of the scattering and diffraction of a plane wave by a circular aperture. That is, consider the plane $z = 0$ with a circular hole given as $x^2 + y^2 < a^2$ and a plane wave $u_I(x, y, z) = e^{ikz}$ approaching the (x, y)-plane. Let the total field $u(x, y, z) = u_I(x, y, z) + u_S(x, y, z)$ vanish on $z = 0$, $x^2 + y^2 > a^2$. Determine the geometrical optics field $u_S(x, y, z)$ and show that the total field $u(x, y, z)$ is discontinuous across the cylinder $x^2 + y^2 = a^2$.

10.1.19. Apply the parabolic equation method for the problem of Exercise 10.1.18 in the region $x > 0$. Obtain $2ikv_z(x, y, z) + v_{xx}(x, y, z) + v_{yy}(x, y, z) = 0$, $z > 0$, with the boundary condition $v(x, y, 0) = \begin{cases} 1, & x^2 + y^2 < a^2, \\ 0, & x^2 + y^2 > a^2. \end{cases}$ Noting that the fundamental solution for this parabolic equation is $V(x, y, z) = -(1/4\pi z) \exp\left(ikr^2/2z\right)$, where $r^2 = (x - \xi)^2 + (y - \eta)^2$, obtain the solution of the Cauchy problem for the parabolic equation as an integral over $x^2 + y^2 < a^2$. Introduce polar coordinates and show that the double integral can be expressed as a single integral involving the Bessel function J_0.

10.1.20. Solve the problem of scattering by a circular cylinder with the Neumann condition $\partial u(x, y)/\partial n = 0$ on $x^2 + y^2 = a^2$.

10.1.21. Write (10.1.173) in the form $u(x, y) = v(x, y)e^{ikx}$ and show that $v(x, y)$ is a solution of the paraxial wave equation (10.1.174) and satisfies the boundary condition (10.1.169).

10.1.22. If $v(x, z)$ and $v(y, z)$ are solutions of the paraxial wave equations $2ikv_z(x, z) + v_{xx}(x, z) = 0$ and $2ikv_z(y, z) + v_{yy}(y, z) = 0$, respectively, show that $v(x, y, z) = v(x, z)v(y, z)$ is a solution of the paraxial wave equation in $3D$ $2ikv_z(x, y, z) + v_{xx}(x, y, z) + v_{yy}(x, y, z) = 0$.

10.1.23. Use the results of the discussion of the propagation of a Gaussian beam given in the text and Exercise 10.1.22 to obtain the Gaussian beam

$$u(x, y, z) = A\left(\frac{-ika}{z - ika}\right) \exp\left[ikz + \frac{ik(x^2 + y^2)}{2(z - ika)}\right]$$

in three dimensions.

10.1.24. Verify that (10.1.184) is a solution of (10.1.183).

10.1.25. Discuss the propagation of a beam in a nonlinear medium in the three-dimensional case. Introduce cylindrical coordinates (r, θ, z) and assume that the initial conditions at $z = 0$ are $\phi_j(r, \theta, 0) = 0$, $j \geq 0$, $a_0(r, \theta, 0) = E \exp\left(-r^2/\alpha^2\right)$, $a_j(r, \theta, 0) = 0$, $j \geq 1$. Proceed as in the two-dimensional case and assume that the phase and amplitude are independent of θ. Determine an approximate focal distance z_f as in (10.1.209) and a focal length $\hat{z}_f$ as in (10.1.225), based on the solution of the appropriate nonlinear geometrical optics equations obtained by proceeding as in the two-dimensional problem. Note that the equation for f that occurs in three dimensions is easier to solve that its two-dimensional analog.

10.2 THE PROPAGATION OF DISCONTINUITIES AND SINGULARITIES FOR HYPERBOLIC EQUATIONS

Discontinuities and *singularities* in the data for (linear) *hyperbolic equations* are propagated along *characteristics*, as we have demonstrated on a number of occasions. The behavior of solutions in the neighborhood of characteristics across which discontinuities or other rapid variations of the solution occur can be studied separately. Although the results obtained do not, in general, yield a full description of the solution, they do describe it in a region of interest, such as at a wave front. Other approximate methods of the type given in Sections 10.1 and 5.7 may be used to complete the description of the solution in certain other regions of interest.

Solutions with Jump Discontinuities

To begin, we consider the *hyperbolic equation*

$$u_{tt}(\mathbf{x},t) - \nabla \cdot (p(\mathbf{x})\nabla u(\mathbf{x},t)) + 2\lambda(\mathbf{x})u_t(\mathbf{x},t) + q(\mathbf{x})u(\mathbf{x},t) = 0. \qquad (10.2.1)$$

This has the general form of the hyperbolic equations considered in Chapter 4 except that we have put $\rho(\mathbf{x}) = 1$ and $F(\mathbf{x},t) = 0$ in (4.1.10), and added a damping term $2\lambda(\mathbf{x})u_t(\mathbf{x},t)$, where $\lambda(\mathbf{x})$ is a positive function. The coefficients $p(\mathbf{x})$ and $q(\mathbf{x})$ satisfy the conditions given in Section 4.1.

We study the form of the solutions of (10.2.1) near a *wave front surface* $\phi(\mathbf{x},t) = 0$. [In the one-dimensional case we replace $\mathbf{x}$ by x and $\phi(x,t) = 0$ is a curve.] It is assumed that the solution $u(\mathbf{x},t)$ vanishes on one side of the surface and is a nonzero function that satisfies (10.2.1) on the other side of the surface. The solution is represented as

$$u(\mathbf{x},t) = v(\mathbf{x},t)H(\phi(\mathbf{x},t)), \qquad (10.2.2)$$

where $H(\phi)$ is the Heaviside function, with $H(\phi) = 0$ when $\phi < 0$ and $H(\phi) = 1$ when $\phi \geq 0$. The representation (10.2.2) requires that $v(\mathbf{x},t)$ be a solution of (10.2.1) in the region $\phi(\mathbf{x},t) > 0$ and that the value of $v(\mathbf{x},t)$ on $\phi(\mathbf{x},t) = 0$ equals the jump in the solution $u(\mathbf{x},t)$ across the wave front.

To determine the conditions that (10.2.2) places on $\phi(\mathbf{x},t)$ and $v(\mathbf{x},t)$, we insert (10.2.2) into (10.2.1). This gives, on using $H'(\phi) = \delta(\phi)$,

$$\delta'(\phi)\{\phi_t^2 - p(\nabla\phi)^2\}v - \delta(\phi)\{2\phi_t v_t - 2p\nabla\phi\cdot\nabla v + 2\lambda\phi_t v - (\nabla\phi\cdot\nabla p)v$$

$$+ [\phi_{tt} - p\nabla^2\phi]v\} + H(\phi)\{v_{tt} - \nabla\cdot(p\nabla v) + 2\lambda v_t + qv\} = 0, \qquad (10.2.3)$$

with $\nabla = \partial/\partial x$ in the one-dimensional case. Since $v(\mathbf{x},t)$ is a solution of (10.2.1), the coefficient of $H(\phi)$ in (10.2.3) must vanish in the region where $H(\phi) \neq 0$. The delta function $\delta(\phi)$ and its derivative $\delta'(\phi)$ both vanish when $\phi \neq 0$. Since $\delta'(\phi)$ is more singular than $\delta(\phi)$ at $\phi = 0$, we require that their coefficients vanish separately if (10.2.3) is to be satisfied. This implies that $\phi(\mathbf{x},t) = 0$ must be a solution of the

characteristic equation for (10.2.1):

$$\phi_t^2(\mathbf{x},t) - p(\mathbf{x})(\nabla\phi(\mathbf{x},t))^2 = 0. \tag{10.2.4}$$

This result is not unexpected since we have already shown that all discontinuities must occur across characteristics.

The vanishing of the coefficient of $\delta(\phi)$ in (10.2.3) implies that on the characteristic surface $\phi(\mathbf{x},t) = 0$ the function $v(\mathbf{x},t)$ must satisfy

$$2\phi_t(\mathbf{x},t)v_t(\mathbf{x},t) - 2p(\mathbf{x})\nabla\phi(\mathbf{x},t)\cdot\nabla v(\mathbf{x},t) + 2\lambda(\mathbf{x})\phi_t(\mathbf{x},t)v(\mathbf{x},t)$$
$$+ [\phi_{tt}(\mathbf{x},t) - \nabla\cdot(p(\mathbf{x})\nabla\phi(\mathbf{x},t))]v(\mathbf{x},t) = 0. \tag{10.2.5}$$

The appropriate solution $\phi(\mathbf{x},t)$ of (10.2.4) may be determined by the method of characteristics. The relevant equations are given in Exercise 2.4.13. The initial conditions for determining the requisite characteristic function $\phi(\mathbf{x},t)$ must be found from the data for the hyperbolic equation (10.2.1). Then (10.2.5) can be solved as we demonstrate. Again the conditions on $v(\mathbf{x},t)$ in (10.2.5) come from the data for (10.2.1).

Before doing so, we use the same approach on the parabolic equation (4.2.5), where we put $\rho(\mathbf{x}) = 1$. Thus, we insert (10.2.2) into (4.2.5) and obtain

$$-\delta'(\phi)p(\nabla\phi)^2v+\delta(\phi)[\phi_tv-2p\nabla\phi\cdot\nabla v-\nabla\cdot(p\nabla)v]+H(\phi)[v_t-\nabla\cdot(p\nabla v)+qv] = 0. \tag{10.2.6}$$

Arguing as before, we conclude that $(\nabla\phi(\mathbf{x},t))^2 = 0$, so that $\phi = \phi(t)$. Any jumps that occur in the solution must occur across the characteristics t = constant. Since ϕ is independent of $\mathbf{x}$, the coefficient of $\delta(\phi)$ reduces to $\phi'(t)v(\mathbf{x},t)$, and this must vanish on $\phi(t) = 0$. But $\phi(t)$ can be expressed in the form $\phi(t) = t - c$. (It vanishes on the characteristic $t = c$.) Consequently, $\phi'(t) \neq 0$ anywhere. Therefore, we must have $v(\mathbf{x},t) = 0$ on the characteristic, and the solution of the parabolic equation cannot have a jump across the characteristic. This is consistent with our results in Chapter 6.

We have again determined that jumps in the solutions of homogeneous hyperbolic and parabolic equations with smooth coefficients can occur only across characteristics. Since elliptic equations do not have real characteristics, the foregoing analysis applied to the elliptic equation (4.1.9) with $F(\mathbf{x}) = 0$ shows that its solutions cannot have any jumps. In the parabolic case, even though there are real characteristics, no jumps can occur across them. Only in the hyperbolic case can there be nonzero jumps.

Bicharacteristics and the Propagation of Jump Discontinuities

We apply the method of characteristics to (10.2.4). Let $\alpha = \phi_t$, $\boldsymbol{\beta} = \nabla\phi$. Then

$$F(\mathbf{x},\alpha,\boldsymbol{\beta}) = \alpha^2 - p(\mathbf{x})|\boldsymbol{\beta}|^2 = 0. \tag{10.2.7}$$

The characteristic curves for (10.2.4) are called *bicharacteristics* to distinguish them from the characteristics of the hyperbolic equation (10.2.1). Let the bicharacteristic

curves be represented as $(\mathbf{x}(s),\ t(s),\ \phi(s),\ \alpha(s),\ \boldsymbol{\beta}(s))$. Then we have the equations (see Exercise 2.4.13)

$$\frac{d\mathbf{x}}{ds} = -2p\boldsymbol{\beta},\ \frac{dt}{ds} = 2\alpha,\ \frac{d\phi}{ds} = -2[\alpha^2 - p|\boldsymbol{\beta}|^2] = 0,\ \frac{d\alpha}{ds} = 0,\ \frac{d\boldsymbol{\beta}}{ds} = |\boldsymbol{\beta}|^2\nabla p, \tag{10.2.8}$$

where $d\phi/ds = 0$ since $F = 0$ along the bicharacteristics.

It follows from (10.2.8) that ϕ = constant on the bicharacteristics. Thus if $\phi = 0$ initially, $\phi(\mathbf{x}, t) = 0$ is the resulting characteristic surface. Also, (10.2.8) shows that α = constant along the bicharacteristics. Then (10.2.8) indicates that the parameter s can be replaced by the (time) parameter t along the bicharacteristics, with $dt = 2\alpha\, ds$. This yields $d\mathbf{x}/dt = -(p/\alpha)\boldsymbol{\beta}$, and a similar transformation can be carried out for the other equations in (10.2.8).

The bicharacteristics can be used to reduce (10.2.5) to an ODE. Given the (bicharacteristic) curve $\mathbf{x} = \mathbf{x}(t),\ t = t$ on the characteristic $\phi(\mathbf{x}, t) = 0$, we have

$$2\phi_t v_t - 2p\nabla\phi\cdot\nabla v + 2\lambda\phi_t v = 2\alpha v_t - 2p\boldsymbol{\beta}\cdot\nabla v + 2\lambda\alpha v = 2\alpha\left(v_t - \frac{p}{\alpha}\,\boldsymbol{\beta}\cdot\nabla v + \lambda v\right)$$

$$= 2\alpha[v_t + \mathbf{x}'(t)\cdot\nabla v + \lambda v] = 2\alpha\left(\frac{dv}{dt} + \lambda v\right), \tag{10.2.9}$$

in view of the above, with $v = v(\mathbf{x}(t), t)$. Thus (10.2.5) reduces to

$$\frac{dv}{dt} + \lambda v + \frac{1}{2\alpha}\,[\phi_{tt} - \nabla\cdot(p\nabla\phi)]v = 0. \tag{10.2.10}$$

on the bicharacteristics, along which $\alpha = \phi_t$ = constant.

If $p(\mathbf{x}) = p$ is a constant and $\phi(\mathbf{x}, t) = 0$ is the plane wave front $\phi(\mathbf{x}, t) = \sqrt{p}\,t - \mathbf{x}\cdot\boldsymbol{\gamma} - c = 0$, where $\boldsymbol{\gamma}$ and c are constants, with $|\boldsymbol{\gamma}| = 1$, (10.2.10) becomes $dv/dt + \lambda v = 0$. Since $\lambda > 0$, this implies that $v(\mathbf{x}, t)$—when evaluated on the characteristic— decays (exponentially) with increasing time. As a result, the presence of the damping term $2\lambda(\mathbf{x})u_t(\mathbf{x}, t)$ in (10.2.1) has the effect of smoothing out the discontinuity (10.2.2).

In the two-dimensional case, if $p(\mathbf{x}) = p$ is a constant and $\phi = \sqrt{p}\,t - r = 0$ is a cylindrical wave front, (10.2.10) takes the form $dv/dt + \lambda v + (p/2\alpha r)v = 0$. Since we can replace r by $\sqrt{p}\,t$ on the wave front and $\alpha = \phi_t = \sqrt{p}$ on the bicharacteristics, we obtain the solution [assuming that $\lambda(\mathbf{x}) = \lambda$ is a constant] $v = (c/\sqrt{t})e^{-\lambda t}$, where c is a constant that can vary from one bicharacteristic to the next. v decays exponentially because of the damping effect, and it decays algebraically because of the spreading of the bicharacteristics. Also, v is singular at $t = 0$ because all the bicharacteristics intersect there.

In three dimensions, if $p(\mathbf{x}) = p$ is a constant and we consider the spherical wave front $\phi = \sqrt{p}\,t - r = 0$, we find that if $\lambda(\mathbf{x}) = \lambda$ is a constant, $v = (c/t)e^{-\lambda t}$, where the constant c may have different values on different bicharacteristics. We again observe the combined exponential and algebraic decay of v and its singularity at $t = 0$.

Although our result only describes the variation of the solution within the discontinuity region, it is possible to extend it to obtain an expression valid in the neighborhood of the discontinuity region as well. To do so we consider the characteristic surfaces $\phi(\mathbf{x}, t) = \text{constant}$, which are parallel to a prescribed characteristic surface $\phi(\mathbf{x}, t) = 0$. For small values of the constant, it is possible to expand $v(\mathbf{x}, t)$ in a Taylor series of the form

$$v(\mathbf{x}, t) = \sum_{j=0}^{\infty} v_j(\mathbf{x}, t) \frac{\phi^j(\mathbf{x}, t)}{j!}, \tag{10.2.11}$$

where the $v_j(\mathbf{x}, t)$ are evaluated on the characteristic $\phi(\mathbf{x}, t) = 0$. [Effectively, we are introducing ϕ as a new coordinate in addition to a set of coordinates on the characteristic surface $\phi(\mathbf{x}, t) = 0$.] We insert (10.2.11) into (10.2.2) and obtain

$$u(\mathbf{x}, t) = \left(\sum_{j=0}^{\infty} v_j(\mathbf{x}, t) \frac{\phi^j(\mathbf{x}, t)}{j!} \right) H(\phi(\mathbf{x}, t)). \tag{10.2.12}$$

The generalized functions

$$H_j(x) = \left[\frac{x^j}{j!} \right] H(x), \qquad j = 0, 1, 2, \ldots, \tag{10.2.13}$$

where $H(x)$ is the Heaviside function, have the property that $H_j'(x) = H_{j-1}(x), j \geq 1$ [see (7.2.41) and Example 7.1]. Thus we write (10.2.12) as

$$u(\mathbf{x}, t) = \sum_{j=0}^{\infty} v_j(\mathbf{x}, t)\, H_j(\phi(\mathbf{x}, t)). \tag{10.2.14}$$

Inserting (10.2.14) into (10.2.1) and using the above, we again find that $\delta'(\phi)$ is the most singular term and this implies that $\phi(\mathbf{x}, t)$ must satisfy the characteristic equation (10.2.4). Equating the coefficients of $\delta(\phi)$ and $H_j(\phi)$ to zero yields the system of *transport equations*

$$2 \frac{\partial \phi}{\partial t} \frac{\partial v_0}{\partial t} - 2p \nabla \phi \cdot \nabla v_0 + 2\lambda \frac{\partial \phi}{\partial t} v_0 + \left[\frac{\partial^2 \phi}{\partial t^2} - \nabla \cdot (p \nabla \phi) \right] v_0 = 0, \tag{10.2.15}$$

$$2 \frac{\partial \phi}{\partial t} \frac{\partial v_j}{\partial t} - 2p \nabla \phi \cdot \nabla v_j + 2\lambda \frac{\partial \phi}{\partial t} v_j + \left[\frac{\partial^2 \phi}{\partial t^2} - \nabla \cdot (p \nabla \phi) \right] v_j$$
$$= -\frac{\partial^2 v_{j-1}}{\partial t^2} + \nabla \cdot (p \nabla v_{j-1}) - 2\lambda \frac{\partial v_{j-1}}{\partial t} - q v_{j-1}, \qquad j \geq 1. \tag{10.2.16}$$

In view of (10.2.9), each of the transport equations can be expressed as an ODE along the bicharacteristics.

In solving these equations in terms of the discontinuous data given for the problem, we do not merely determine the $v_j(\mathbf{x}, t)$ along the bicharacteristics that issue from the discontinuity region. Instead, the data are represented in the form (10.2.14) at $t = 0$

or on the boundary, and the $v_j(\mathbf{x}, t)$ are obtained in terms of the coefficients in these series. Thereby the $v_j(\mathbf{x}, t)$ are specified over the entire region where the solution is valid. This more general result yields the jumps in $u(\mathbf{x}, t)$ and its normal derivatives at $\phi(\mathbf{x}, t) = 0$.

This problem can be simplified if the characteristic surface $\phi(\mathbf{x}, t) = 0$ is expressed as $\phi(\mathbf{x}, t) = t - \psi(\mathbf{x}) = 0$. Then $\psi(\mathbf{x})$ = constant represents the position of the wave fronts at different times t. [Often, it is convenient to write $\phi(\mathbf{x}, t)$ as $\phi = \psi(\mathbf{x}) - t$.] The functions in (10.2.11), evaluated on the characteristic, may now be taken to be functions of $\mathbf{x}$, and (10.2.14) becomes

$$u(\mathbf{x}, t) = \sum_{j=0}^{\infty} v_j(\mathbf{x}) H_j(t - \psi(\mathbf{x})). \tag{10.2.17}$$

The characteristic equation (10.2.4) becomes $(\nabla\psi(\mathbf{x}))^2 = 1/p(\mathbf{x})$, which has the form of the *eiconal equation*. The transport equations (10.2.15)–(10.2.16) take the form, with $v_{-1} = 0$,

$$2p\nabla\psi\cdot\nabla v_j + 2\lambda v_j + \nabla\cdot(p\nabla\psi)v_j = \nabla\cdot(p\nabla v_{j-1}) - qv_{j-1}, \qquad j \geq 0. \tag{10.2.18}$$

The expressions (10.2.14) or (10.2.17) are to be interpreted as *weak solutions* of (10.2.1) since they involve generalized functions. To see how initial conditions are to be assigned for the bicharacteristic or ray equations and for the transport equations for the $v_j(\mathbf{x})$, in our discussion below we consider initial and boundary value and Cauchy problems for the one-dimensional telegrapher's equation. The results obtained below for various problems involving that equation suggest how the general problem ought to be treated.

Functions with Jump Discontinuities and Heaviside Functions

The use of the expansion (10.2.14) requires that the data for the problems be expressed directly in terms of the Heaviside function $H(x)$ or the Dirac delta function $\delta(x)$. Suppose that the data are given in terms of a discontinuous function $F(x) = \begin{cases} f(x), & x < 0, \\ g(x), & x > 0. \end{cases}$ We assume that $f(x)$ and $g(x)$ are smooth functions in their intervals of definition and have finite one-sided limits at $x = 0$. If $f(x)$ can be defined or extended as a smooth function for all x, we can write $F(x) = f(x) + H(x)[g(x) - f(x)]$. Furthermore, if $g(x) - f(x)$ can be expanded in a Taylor series around $x = 0$, we can express the result as

$$F(x) = f(x) + \sum_{j=0}^{\infty} [F^{(j)}(x)]_{x=0} H_j(x), \tag{10.2.19}$$

where $[F^{(j)}(x)]_{x=0} = g^{(j)}(0) - f^{(j)}(0)$ represents the jump in the jth derivative of $F(x)$ at $x = 0$. The $H_j(x)$ are defined as in (10.2.13).

More generally, if $F(x)$ is a smooth function for $x \neq \alpha$ with at most jumps in its derivatives at $x = \alpha$, we can represent $F(x)$ as in (7.2.42) and obtain

$$F(x) = \tilde{F}(x) + \sum_{j=0}^{\infty} [F^{(j)}(x)]_{x=\alpha} H_j(x - \alpha), \tag{10.2.20}$$

where $\tilde{F}(x)$ is the difference between $F(x)$ and the sum on the right side of (10.2.20). An example is given in Example 7.2 [where we replace $f(x)$ by $F(x)$] and the representation (10.2.20) is obtained. If $F(x)$ has jump discontinuities at a finite number of points α_j, $j = 1, 2, \ldots, n$, a representation in terms of Heaviside functions can again be given, as will be seen below.

If $F(x)$ is given as in (10.2.19) or (10.2.20), its generalized derivatives can be obtained by differentiating the series term by term. Apart from the addition of a finite number of terms that involve the delta function and its derivatives, the derivatives of $F(x)$ are still represented in terms of a series of the functions $H_j(x)$ or $H_j(x - \alpha)$. We make use of the representations (10.2.19) and (10.2.20) in our discussion below.

Initial Value Problem for the Telegrapher's Equation: Jump Discontinuities

We consider the initial value problem for the *telegrapher's equation*

$$u_{tt}(x,t) - \gamma^2 u_{xx}(x,t) + 2\lambda u_t(x,t) = 0, \qquad -\infty < x > \infty, \; t > 0, \tag{10.2.21}$$

with the initial data

$$u(x,0) = \begin{cases} -1, & x < -1, \\ \sin(x), & -1 < x < 1, \\ 1, & x > 1, \end{cases} \qquad u_t(x,0) = 0. \tag{10.2.22}$$

The function $u(x,0)$ can be written in terms of the Heaviside function as

$$u(x,0) = -1 + (1 + \sin(x))H(x+1) + (1 - \sin(x))H(x-1). \tag{10.2.23}$$

To account for the jumps in the initial data at $x = 1$ and $x = -1$, we introduce the expansions

$$U(x,t) = \sum_{j=0}^{\infty} v_j(x,t) H_j[\phi(x,t)] + \sum_{j=0}^{\infty} w_j(x,t) H_j[\psi(x,t)]. \tag{10.2.24}$$

Inserting (10.2.24) into (10.2.21), we see that $\phi(x,t)$ and $\psi(x,t)$ must satisfy the characteristic equation $\phi_t^2 - \gamma^2 \phi_x^2 = 0$ and that the $v_j(x,t)$ and $w_j(x,t)$ must satisfy the (appropriate) transport equations (10.2.15)–(10.2.16) specialized to (10.2.21). At $t = 0$, $U(x,0) = \sum_{j=0}^{\infty} v_j(x,0) H_j[\phi(x,0)] + \sum_{j=0}^{\infty} w_j(x,0) H_j[\psi(x,0)]$, and this can account only for the singular terms in the data (10.2.23). This implies that

we should set $\phi(x,0) = x+1$, $\psi(x,0) = x-1$. We must use both solutions of the characteristic equations for ϕ and ψ that satisfy the conditions above and they are $\tilde{\phi}(x,t) = x+1-\gamma t$, $\hat{\phi}(x,t) = x+1+\gamma t$, $\tilde{\psi}(x,t) = x-1-\gamma t$, $\hat{\psi}(x,t) = x-1+\gamma t$.

For each characteristic line $\tilde{\phi}(x,t) = \hat{\phi}(x,t) = \tilde{\psi}(x,t) = \hat{\psi}(x,t) = 0$, we introduce a series of the form (10.2.14), and thus we replace (10.2.24) by

$$U(x,t) = \sum_{j=0}^{\infty} \tilde{v}_j(x,t)H_j[x+1-\gamma t] + \sum_{j=0}^{\infty} \hat{v}_j(x,t)H_j[x+1+\gamma t]$$

$$+ \sum_{j=0}^{\infty} \tilde{w}_j(x,t)H_j[x-1-\gamma t] + \sum_{j=0}^{\infty} \hat{w}_j(x,t)H_j[x-1+\gamma t]. \quad (10.2.25)$$

Considering only the leading terms we have $U(x,0) = [\tilde{v}_0(x,0) + \hat{v}_0(x,0)]H_0[x+1]+[\tilde{w}_0(x,0)+\hat{w}_0(x,0)]H_0[x-1]+\cdots = (1+\sin(x))H[x+1]+(1-\sin(x))H[x-1]$ and $U_t(x,0) = [-\gamma\tilde{v}_0(x,0) + \gamma\hat{v}_0(x,0)]\delta[x+1] + [-\gamma\tilde{w}_0(x,0) + \gamma\hat{w}_0(x,0)]\delta[x-1]+\cdots = 0$, since we are only concerned with the singular part of the data. Equating coefficients of like terms in the equations above and solving the resulting equations yields $\tilde{v}_0(x,0) = \hat{v}_0(x,0) = (1/2)(1+\sin(x))$, $\tilde{w}_0(x,0) = \hat{w}_0(x,0) = (1/2)(1-\sin(x))$. The transport equation for the $\tilde{v}_0(x,t)$, $\hat{v}_0(x,t)$, $\tilde{w}_0(x,t)$, and $\hat{w}_0(x,t)$ are easily solved, and we obtain

$$U(x,t) = \frac{1}{2}[1+\sin(x-\gamma t)]e^{-\lambda t}H(x+1-\gamma t)+\frac{1}{2}[1+\sin(x+\gamma t)]e^{-\lambda t}H(x+1+\gamma t)$$

$$+\frac{1}{2}[1-\sin(x-\gamma t)]e^{-\lambda t}H(x-1-\gamma t)+\frac{1}{2}[1-\sin(x+\gamma t)]e^{-\lambda t}H(x-1+\gamma t)+\cdots. \quad (10.2.26)$$

The coefficients of the Heaviside functions in (10.2.26) are not evaluated on the relevant characteristics $x+1\pm\gamma t = 0$ or $x-1\pm\gamma t = 0$. Although our original expansion forms required the transport coefficients to be evaluated on the characteristics, we showed in the preceding that these expansions are formal solutions even if the coefficients are not evaluated on the characteristics. To obtain a representation where each coefficient in the expansion is evaluated on the relevant characteristic, we expand each of the coefficients in (10.2.26) in powers of the relevant characteristic variables. If this is done, we obtain to leading terms

$$U(x,t) = \frac{1}{2}[1-\sin(1)]e^{-\lambda t}[H(x+1-\gamma t) + H(x+1+\gamma t)$$

$$+ H(x-1-\gamma t) + H(x-1+\gamma t)] + \cdots \quad (10.2.27)$$

and the jump across each characteristic is $\frac{1}{2}[1-\sin(1)]e^{-\lambda t}$.

The solution $u(x,t)$ of (10.2.21)–(10.2.22) may be represented as

$$u(x,t) = v(x,t) + U(x,t), \quad (10.2.28)$$

with $U(x,t)$ given as above. In view of (10.2.22)–(10.2.23), the initial data for $v(x,t)$ are $v(x,0) = -1+[\sin(x)+\sin(1)]H(x+1)-[\sin(x)-\sin(1)]H(x-1)+\cdots$, and $v_t(x,0) = \lambda[\sin(1)-1][H(x+1)+H(x-1)]+\cdots$. In addition, $U(x,t)$ gives rise to an inhomogeneous term that occurs in the equation for $v(x,t)$. Although $u(x,0)$ is discontinuous at $x=-1$ and $x=1$, $v(x,0)$ is continuous for all x. Consequently, $v(x,t)$ is a smoother function than $u(x,t)$. We do not solve for $v(x,t)$.

Initial Value Problem for the Telegrapher's Equation: Singular Solutions

Next we show how to treat problems in which the initial data for the telegrapher's equation do not lead to jump discontinuities but result in stronger or, possibly, weaker singularities in the solution. For instance, the initial value problem for the telegrapher's equation in Section 1.2 has $u(x,0) = \delta(x)$, so that the solution is expected to have a delta function singularity. Also, the problem considered for $v(x,t)$ in (10.2.28) leads to a solution that is continuous everywhere but has discontinuous derivatives.

These problems can be dealt with by observing that (10.2.14) can be replaced by the expansion

$$u(x,t) = \sum_{j=0}^{\infty} v_j(x,t) S_j[\phi(x,t)], \tag{10.2.29}$$

where the (generalized) functions $S_j[x]$ are related to each other by

$$S_j'[x] = S_{j-1}[x] \tag{10.2.30}$$

but are otherwise arbitrary. On inserting (10.2.29) into the telegrapher's equation, using (10.2.30) and equating the coefficients of the $S_j[\phi]$ to zero, we conclude that $\phi(x,t)$ must satisfy the characteristic equation and the $v_j(x,t)$ must satisfy the transport equations for the telegrapher's equation, as found above. For example, if we put $S_0[x] = H[x]$, the Heaviside function, (10.2.30) yields $S_j[x] = H_j[x]$ and (10.2.29) reduces to (10.2.14).

For a given problem we may use the general series (10.2.29) and then decide on the basis of the data what form S_0 and the S_j should take. As an example, we consider the problem of Section 1.2 in which $u(x,t)$ satisfies

$$u_{tt}(x,t) - \gamma^2 u_{xx}(x,t) + 2\lambda u_t(x,t) = 0, \qquad -\infty < x > \infty,\ t > 0, \tag{10.2.31}$$

with the initial data

$$u(x,0) = \delta(x), \qquad u_t(x,0) = 0. \tag{10.2.32}$$

Expanding $u(x,t)$ as in (10.2.29), we obtain $u(x,0) = \sum_{j=0}^{\infty} v_j(x,0) S_j[\phi(x,0)] = \delta(x)$. Since the S_j are increasingly less singular for increasing j, we conclude that $S_0[x] = \delta(x)$ and $\phi(x,0) = x$, so that $S_1[x] = H[x]$, $S_j[x] = H_{j-1}[x]$, $j \geq 1$, with the H_j defined as in (10.2.13). Again we must consider the two characteristics

$\tilde{\phi}(x,t) = x - \gamma t = 0$, $\hat{\phi}(x,t) = x + \gamma t = 0$. It is easily shown on using the transport equations for $\tilde{v}_0(x,t)$ and $\hat{v}_0(x,t)$ that

$$u(x,t) = \frac{1}{2}e^{-\lambda t}\delta[x + \gamma t] + \frac{1}{2}e^{-\lambda t}\delta[x - \gamma t] + \cdots, \tag{10.2.33}$$

to leading terms. This confirms the result obtained in Section 1.2.

As another example we consider the initial value problem that yields the causal fundamental solution for the telegrapher's equation, as follows from Chapter 7. The function $u(x,t)$ satisfies (10.2.31) and the data are

$$u(x,0) = 0, \qquad u_t(x,0) = \delta(x). \tag{10.2.34}$$

Expanding $u(x,t)$ as in (10.2.29), we have $u(x,0) = \sum_{j=0}^{\infty} v_j(x,0)S_j[\phi(x,0)] = 0$, and $u_t(x,0) = v_0(x,0)\partial\phi(x,0)/\partial t\, S_{-1}[\phi(x,0)] + \sum_{j=0}^{\infty}\{v_{j+1}(x,0)\partial\phi(x,0)/\partial t + \partial v_j(x,0)/\partial t\}S_j[\phi(x,0)] = \delta(x)$.

Since $S_{-1}[x] = S_0'[x]$ is the most singular term in the series, we conclude that $S_0'[x] = \delta(x)$, $\phi(x,0) = x$, so that $S_0[x] = H[x]$. Consequently, the series (10.2.29) for this problem is identical with (10.2.14) since $S_j[x] = H_j[x]$.

Adapting the result (6.7.51), it is easy to show that the causal fundamental solution of the telegrapher's equation is

$$u(x,t) = \frac{1}{2\gamma}e^{-\lambda t}I_0\left(\frac{\lambda}{\gamma}\sqrt{\gamma^2t^2 - x^2}\right)H(\gamma t - x)H(\gamma t + x), \tag{10.2.35}$$

where $I_0[z]$ is the modified Bessel function of zero order. On using the Taylor series for $I_0[z]$ and considering (10.2.35) near each characteristic $x - \gamma t = 0$ and $x + \gamma t = 0$ for $t > 0$, it can be shown that the expansion based on the use of (10.2.29) agrees completely with the expression obtained from the exact solution (10.2.35). To accomplish this, however, the terms in the expansion must be evaluated on the appropriate characteristics $x - \gamma t = 0$ and $x + \gamma t = 0$. This is considered in the exercises.

General Singularity Expansions

The foregoing shows that singularity expansions of the *hyperbolic equation* (10.2.1) need not be restricted to the form (10.2.14). They can be given as

$$u(\mathbf{x},t) = \sum_{j=0}^{\infty} v_j(\mathbf{x},t)S_j[\phi(\mathbf{x},t)], \tag{10.2.36}$$

where the $S_j[z]$ satisfy the relations $S_j'[z] = S_{j-1}[z]$. This represents a formal solution of (10.2.1) if $\phi(\mathbf{x},t)$ satisfies the *characteristic equation*

$$\phi_t^2(\mathbf{x},t) - p(\mathbf{x})(\nabla\phi(\mathbf{x},t))^2 = 0 \tag{10.2.37}$$

and the $v_j(\mathbf{x}, t)$ satisfy the transport equations

$$2\frac{\partial\phi}{\partial t}\frac{\partial v_j}{\partial t} - 2p\nabla\phi\cdot\nabla v_j + 2\lambda\frac{\partial\phi}{\partial t}v_j + \left[\frac{\partial^2\phi}{\partial t^2} - \nabla\cdot(p\nabla\phi)\right]v_j$$

$$= -\frac{\partial^2 v_{j-1}}{\partial t^2} + \nabla\cdot(p\nabla v_{j-1}) - 2\lambda\frac{\partial v_{j-1}}{\partial t} - qv_{j-1}, \quad j \geq 0, \quad (10.2.38)$$

where it is assumed that $v_{-1}(\mathbf{x}, t) = 0$.

Additionally, the expansion (10.2.36) no longer need be restricted to describing the *propagation of singularities*. For example, if we set

$$S_0[\phi] = e^{i\omega\phi}, \qquad S_j[\phi] = \frac{1}{(i\omega)^j}e^{i\omega\phi}, \quad (10.2.39)$$

where ω is a large (real) parameter, (10.2.36) may be regarded as an asymptotic solution of (10.2.1), as would arise if the problem contains rapidly oscillating data. For instance, if initially we have

$$u(\mathbf{x}, 0) = f(\mathbf{x})e^{i\omega g(\mathbf{x})}, \qquad u_t(\mathbf{x}, 0) = 0, \quad (10.2.40)$$

where ω is large and $g(\mathbf{x})$ is a real-valued function, we look for a solution of (10.2.1) in the form (10.2.36) with the S_j given by (10.2.39). From the initial data (10.2.40) we conclude that $\phi(\mathbf{x}, 0) = g(\mathbf{x})$. There are two solutions $\tilde{\phi}(\mathbf{x}, t)$ and $\hat{\phi}(\mathbf{x}, t)$ of (10.2.37) that satisfy this initial condition. For each solution we construct a series (10.2.36) with coefficients $\tilde{v}_j(\mathbf{x}, t)$ and $\hat{v}_j(\mathbf{x}, t)$, respectively. Inserting these series into the data (10.2.40) determines initial conditions for $\tilde{v}_j(\mathbf{x}, t)$ and $\hat{v}_j(\mathbf{x}, t)$, each of which satisfies the transport equations (10.2.38).

With $\lambda = q = 0$ and $p = \gamma^2$ = constant in (10.2.1), the choice $\phi(\mathbf{x}, t) = \psi(\mathbf{x}) - t$, $v_j(\mathbf{x}, t) = v_j(\mathbf{x})$, with the S_j defined as in (10.2.39), yields the *geometrical optics expansion* for the *reduced wave equation*. In fact, the series

$$v(\mathbf{x})e^{i\omega\psi(\mathbf{x})} = \sum_{j=0}^{\infty} v_j(\mathbf{x})\frac{1}{(i\omega)^j}e^{i\omega\psi(\mathbf{x})} \quad (10.2.41)$$

is a solution of the reduced wave equation as follows on setting $u(\mathbf{x}, t) = e^{-i\omega t}v(\mathbf{x})$ $e^{i\omega\psi(\mathbf{x})}$. With $\omega = k$, it reduces to the form considered in Section 10.1. In view of the fact that solutions of (10.2.1) undergo rapid variations across (characteristic) singularity regions, it is not surprising that solutions of the type (10.2.36) and (10.2.39) with rapidly oscillating terms are similar in form to solutions that describe the propagation of singularities. In this section, however, we are only interested in discussing the propagation of singularities for solutions of (10.2.1) and do not pursue these matters any further.

Initial Value Problem for the Two-Dimensional Wave Equation: Jump Discontinuities

The preceding examples have dealt with problems in one dimension. Next we consider an initial value problem for the wave equation in two dimensions. Let $u(x, y, t)$ satisfy

the *wave equation*

$$u_{tt}(x,y,t) - \gamma^2(u_{xx}(x,y,t) + u_{yy}(x,y,t)) = 0, \qquad -\infty < x,\ y > \infty,\ t > 0, \tag{10.2.42}$$

where γ is a constant. The initial data for $u(x,y,t)$ have a jump discontinuity on the circle $x^2 + y^2 = a^2$ and are given as

$$u(x,y,0) = \alpha + (\beta - \alpha)H[\rho - a], \qquad u_t(x,y,0) = 0. \tag{10.2.43}$$

Here, α and β are constants, $\rho^2 = x^2 + y^2$, and $H[x]$ is the Heaviside function. We expand $u(x,y,t)$ as

$$u(x,y,t) = \alpha + \sum_{j=0}^{\infty} v_j(x,y,t)S_j[\phi(x,y,t)]. \tag{10.2.44}$$

At $t = 0$ we have $u(x,y,0) = \alpha + \sum_{j=0}^{\infty} v_j(x,y,0)S_j[\phi(x,y,0)] = \alpha + (\beta - \alpha)H[\rho - a]$. This suggests that we set $S_0[\phi] = H[\phi]$, $S_j[\phi] = H_j[\phi]$, with the H_j defined as in (10.2.13) and $\phi(x,y,0) = \rho - a$. Since $\phi(x,y,t)$ must satisfy the characteristic equation $\phi_t^2(x,y,t) - \gamma^2(\phi_x^2(x,y,t) + \phi_y^2(x,y,t)) = 0$, we obtain two possible solutions $\tilde{\phi}(x,y,t) = \rho - a - \gamma t$, $\hat{\phi}(x,y,t) = \rho - a + \gamma t$. To account for $\tilde{\phi}(x,y,t)$ and $\hat{\phi}(x,y,t)$ we replace (10.2.44) by

$$u(x,y,t) = \alpha + \sum_{j=0}^{\infty} \tilde{v}_j(x,y,t)H_j[\tilde{\phi}(x,y,t)] + \sum_{j=0}^{\infty} \hat{v}_j(x,y,t)H_j[\hat{\phi}(x,y,t)]. \tag{10.2.45}$$

Inserting (10.2.45) into the initial conditions (10.2.43), we easily conclude that $\tilde{v}_0(x,y,0) = \hat{v}_0(x,y,0) = (1/2)(\beta - \alpha)$. The transport equations for $\tilde{v}_0(x,y,t)$ and $\hat{v}_0(x,y,t)$ are $\partial\tilde{v}_0(x,y,t)/\partial t + \gamma\,\partial\tilde{v}_0(x,y,t)/\partial x + (\gamma/2\rho)\tilde{v}_0(x,y,t) = 0$, and $\partial\hat{v}_0(x,y,t)/\partial t - \gamma\,\partial\hat{v}_0(x,y,t)/\partial x - (\gamma/2\rho)\hat{v}_0(x,y,t) = 0$, with the solutions

$$\tilde{v}_0(x,y,t) = \frac{1}{2}(\beta - \alpha)\left(\frac{\rho - \gamma t}{\rho}\right)^{1/2}, \qquad \hat{v}_0(x,y,t) = \frac{1}{2}(\beta - \alpha)\left(\frac{\rho + \gamma t}{\rho}\right)^{1/2}. \tag{10.2.46}$$

These expressions may be expanded in powers of $\tilde{\phi}$ and $\hat{\phi}$. Doing so yields as the leading terms in the expansion of $u(x,y,t)$,

$$u = \alpha + \frac{\beta - \alpha}{2}\left(\frac{a}{a + \gamma t}\right)^{1/2} H[\rho - a - \gamma t] + \frac{\beta - \alpha}{2}\left(\frac{a}{a - \gamma t}\right)^{1/2} H[\rho - a + \gamma t] + \cdots. \tag{10.2.47}$$

The coefficients of the Heaviside functions in (10.2.47) could have been determined by expressing the transport equations as ODEs along the bicharacteristics of $\tilde{\phi}(x,y,t) = 0$ and $\hat{\phi}(x,y,t) = 0$, as was done above.

The characteristic surface $\tilde{\phi} = \rho - a - \gamma t = 0$ represents *cylindrical wave fronts* that travel outward from the circle $\rho = a$ with speed $d\rho/dt = \gamma$. The jump in u across the wave fronts decays like $t^{-1/2}$ for increasing t. The characteristic surface $\hat{\phi} = \rho - a + \gamma t = 0$ yields *cylindrical wave fronts* that travel inward from $\rho = a$ with speed γ. At the time $t = a/\gamma$ the wave fronts collapse on the point $\rho = 0$, and (10.2.47) is singular there since the coefficient of $H[\rho - a + \gamma t]$ blows up. The singularity results because the origin is a *focal point* for the bicharacteristics. We do not consider how (10.2.47) should be modified to deal with this singularity.

Modified Singularity Expansions: Fundamental Solution of the Klein-Gordon Equation

The usefulness of the expansion (10.2.36) in solving problems for the hyperbolic equation (10.2.1) depends on the possibility of representing the data for (10.2.1) in the form (10.2.36) and the ability to determine the coefficients in the expansion of the solution in terms of those in the data. It may happen, however, that the bicharacteristics of $\phi(\mathbf{x}, t) = 0$ intersect in the initial or boundary region. This occurs, for example, if $\phi(\mathbf{x}, t) = 0$ is a characteristic cone or conoid with vertex at $t = 0$ in the two- or three-dimensional case. Then, the leading coefficient $v_0(\mathbf{x}, t)$, at least, in (10.2.36) must be singular where the bicharacteristics intersect, and it is not possible to determine its value directly in terms of the data.

As an example, we consider the *causal fundamental solution* for the *Klein-Gordon equation* in three dimensions for $u = u(x, y, z, t)$,

$$u_{tt} - \gamma^2[u_{xx} + u_{yy} + u_{zz}] + c^2 u = 0, \tag{10.2.48}$$

with $-\infty < x,\ y,\ z < \infty,\ t > 0$ and constant coefficients. As shown in Chapter 7, it is a solution of (10.2.48) with the initial data

$$u(x, y, z, 0) = 0, \qquad u_t(x, y, z, 0) = \delta(x)\delta(y)\delta(z). \tag{10.2.49}$$

It is given as (see Exercise 7.4.9)

$$u(x, y, z, t) = \frac{1}{4\pi\gamma r}\,\delta[\gamma t - r] - \frac{cJ_1\left(\frac{c}{\gamma}\sqrt{\gamma^2 t^2 - r^2}\right)}{4\pi\gamma^2\sqrt{\gamma^2 t^2 - r^2}}\,H[\gamma t - r] \tag{10.2.50}$$

for $t > 0$, where $r^2 = x^2 + y^2 + z^2$ and J_1 is the Bessel function of order 1.

Making use of the Taylor series for J_1 yields

$$u = \frac{1}{4\pi\gamma r}\,\delta[\gamma t - r] - \frac{c}{4\pi\gamma^2}\sum_{j=0}^{\infty}\frac{(-1)^j(c/2\gamma)^{2j+1}}{j!(j+1)!}(\gamma t + r)^j(\gamma t - r)^j H[\gamma t - r]. \tag{10.2.51}$$

Since $[\gamma t + r]^j = [2r + \gamma t - r]^j$ and can be expanded in powers of $\gamma t - r$, we see that (10.2.51) has the form $u(x, y, z, t) = \sum_{j=0}^{\infty} v_j(x, y, z, t)S_j[\gamma t - r]$, where

$S_0[\phi] = \delta[\phi]$, $S_j[\phi] = H_{j-1}[\phi]$, $j \geq 1$, with $H_j[\phi]$ defined as in (10.2.13). The transport coefficients v_j are evaluated on the characteristic surface $\phi = \gamma t - r = 0$. The leading terms are $v_0 = 1/4\pi\gamma r$, $v_1 = -c^2/8\pi\gamma^3$.

If we attempt to reproduce these results by using the direct expansion method and set $u(x, y, z, t) = \sum_{j=0}^{\infty} v_j(x, y, z, t) S_j[\phi(x, y, z, t)]$, with the S_j given as above, we find that it is not possible to determine the v_j and the S_j directly in terms of the initial data, because of the singularity that results from the fact that the bicharacteristics of the forward characteristic cone for the problem all intersect at $r = 0$. Therefore, we now present an alternative method for constructing causal fundamental solutions of hyperbolic equations in which the singularity expansion is modified.

We represent the *causal fundamental solution* (10.2.50) of the Klein-Gordon equation in terms of the square of the *hyperbolic distance* σ defined as $\sigma = t^2 - r^2/\gamma^2$ (see Example 6.14). It has the form, for $t > 0$,

$$u(x, y, z, t) = \frac{1}{2\pi\gamma^3}\,\delta[\sigma] - \frac{c}{4\pi\gamma^3}\,\frac{J_1[c\sqrt{\sigma}]}{\sqrt{\sigma}}\,H[\sigma], \tag{10.2.52}$$

as follows on using the properties of the delta function given in Section 7.2.

On expanding J_1 in a Taylor series, as in (10.2.51), we find that $u(x, y, z, t)$ takes the form

$$u(x, y, z, t) = \frac{1}{2\pi\gamma^3}\,\delta[\sigma] + \frac{1}{2\pi\gamma^3}\sum_{j=1}^{\infty}\frac{(-1)^j c^{2j}}{2^{2j} j!}\,H_{j-1}[\sigma]. \tag{10.2.53}$$

Then, if we set $S_0(\sigma) = \delta(\sigma)$ and $S_j(\sigma) = H_{j-1}(\sigma)$, $j \geq 1$, we have $S_j'(\sigma) = S_{j-1}(\sigma)$ and $u(x, y, z, t)$ can be written as

$$u(x, y, z, t) = \sum_{j=0}^{\infty} v_j S_j[\sigma], \qquad v_j = \frac{1}{2\pi\gamma^3}\,\frac{(-1)^j c^{2j}}{2^{2j} j!}. \tag{10.2.54}$$

While this series has the form of (10.2.36) and the S_j satisfy $S_j' = S_{j-1}$, it differs from (10.2.36) in that the surfaces σ = constant are not characteristics of the Klein-Gordon equation unless the constant equals zero. In fact, $\sigma_t^2 - \gamma^2(\nabla\sigma)^2 = 4\sigma$, and only $\sigma = 0$, which is the characteristic cone, satisfies the characteristic equation for the Klein-Gordon equation. We further observe that with $S_{-1}(\sigma) = \delta'(\sigma)$ and $S_{-2}(\sigma) = \delta''(\sigma)$ we have

$$\sigma S_j(\sigma) = j S_{j+1}(\sigma), \qquad j = -2, -1, 0, \ldots. \tag{10.2.55}$$

The coefficients v_j in the representation (10.2.54) were determined from the exact fundamental solution. We now show how to determine the coefficients in (10.2.54) directly from the Klein-Gordon equation and the initial conditions. To do so, we insert (10.2.54) into (10.2.48) and make use of (10.2.55) to obtain

$$\sum_{j=0}^{\infty}\left[4\left(t\frac{\partial v_j}{\partial t} + \mathbf{r}\cdot\nabla v_j + j v_j\right) + \left(\frac{\partial^2 v_{j-1}}{\partial t^2} - \gamma^2\nabla^2 v_{j-1} + c^2 v_{j-1}\right)\right] S_{j-1}(\sigma) = 0 \tag{10.2.56}$$

after some elementary calculations, with $v_{-1} = 0$. On equating the coefficients of the S_j to zero, the expression in the brackets in (10.2.56) yields the transport equations for the v_j.

As we only need to integrate these equations over the forward characteristic cone $t = r/\gamma$, and $\mathbf{r} \cdot \nabla = r(\partial/\partial r)$ on the cone, we have $t\,\partial/\partial t + r\,\partial/\partial r = r\,d/dr$ on $t = r/\gamma$. As a result, the left side of each equation for the v_j has the form $4r\,dv_j/dr + 4jv_j$. Since $v_{-1} = 0$, the ODE satisfied by v_0 on the characteristic cone is $r\,dv_0/dr = 0$. We require that each v_j be bounded at the vertex of the cone where $r = 0$, and this implies that v_0 must be a constant. Then the equations for the remaining v_j become $4r\,dv_j/dr + 4jv_j = -c^2 v_{j-1}$. The boundedness requirement implies that all the v_j are constants, and by induction we conclude that

$$v_j = \frac{(-1)^j c^{2j}}{2^{2j} j!}\, v_0, \qquad j = 1, 2, \ldots. \tag{10.2.57}$$

With $v_0 = 1/2\pi\gamma^3$, the result agrees with (10.2.54).

While we have specified the value of v_0 by using the exact result (10.2.54), the determination of v_0 may be based on comparing our solution with the causal fundamental solution of the wave equation near the singular point at the vertex of the characteristic cone, where the most singular part of both fundamental solutions must agree. Additionally, we have prescribed the form of the S_j in the expansion (10.2.54) that was inserted into the Klein-Gordon equation. However, this was not necessary. As seen in the following, the form of the S_j can also be determined from the conditions given for the problem.

Modified Singularity Expansions: Fundamental Solutions of Hyperbolic Equations

We consider the hyperbolic equation for $u(\mathbf{x}, t)$,

$$L(u(\mathbf{x}, t)) = u_{tt}(\mathbf{x}, t) - \nabla \cdot (p(|\mathbf{x}|)\nabla u(\mathbf{x}, t)) + 2\lambda u_t(\mathbf{x}, t) + q(\mathbf{x}, t)u(\mathbf{x}, t) = 0, \tag{10.2.58}$$

in two or three dimensions, where $|\mathbf{x}| = r$ is the distance from the origin and λ is a constant. With the initial data $u(\mathbf{x}, 0) = 0$ and $u_t(\mathbf{x}, 0) = \delta(\mathbf{x})$ [the delta function $\delta(\mathbf{x})$ is singular at $r = 0$], the solution of (10.2.58) for $t > 0$ over the entire space R_2 or R_3 yields the causal fundamental solution for (10.2.58). We construct an expansion of the fundamental solution in terms of the square of the hyperbolic distance associated with (10.2.58) and determine the leading term in the two- and three-dimensional cases. This method can also be used for the one-dimensional problem. In that case, however, the solution can be determined completely as an expansion in terms of the relevant characteristics, as we have shown.

The hyperbolic distance element s for (10.2.58) is given as $ds^2 = dt^2 - (1/p(r))\,dr^2$, so that σ, the square of the hyperbolic distance measured from $r = 0$ and $t = 0$, is

given as $\sigma = t^2 - \left(\int_0^r 1/\sqrt{p(s)}\, ds\right)^2$. We easily verify that $\sigma(\mathbf{x}, t)$ satisfies

$$\sigma_t^2 - p(r)(\nabla\sigma)^2 = 4\sigma. \tag{10.2.59}$$

We express the fundamental solution of (10.2.58) as $u = u(r, t)$ and expand it in the series

$$u(r,t) = \sum_{j=0}^{\infty} v_j(r,t) S_j(\sigma), \qquad S_j'(\sigma) = S_{j-1}(\sigma), \tag{10.2.60}$$

with the $v_j(r, t)$ and the S_j to be determined. On inserting (10.2.60) into (10.2.58), we obtain

$$\sum_{j=0}^{\infty} \{[\sigma_t^2 - p(r)(\nabla\sigma)^2] v_j(r,t) S_{j-2}(\sigma) + [2\sigma_t\, \partial v_j(r,t)/\partial t - 2p(r)\nabla\sigma\cdot\nabla v_j(r,t)$$

$$+ 2\lambda\sigma_t v_j(r,t) + (\sigma_{tt} - \nabla\cdot(p(r)\nabla\sigma)) v_j(r,t)] S_{j-1}(\sigma) + L[v_j(r,t)] S_j(\sigma)\} = 0. \tag{10.2.61}$$

Now $\sigma = 0$ represents a characteristic conoid for (10.2.58) with vertex at $(r, t) = (0, 0)$. We determine the form of the S_j in (10.2.60) by requiring that the solutions of the (transport) equations for the $v_j(r, t)$ be nonsingular at the vertex of the characteristic conoid.

We begin with the *two-dimensional case*. Only the forward characteristic conoid $t = \int_0^r ds/\sqrt{p(s)}$ concerns us and we have as the differential operator for $v_0(r, t)$,

$$4\sqrt{p(r)} \int_0^r \frac{ds}{\sqrt{p(s)}} \left\{ \frac{dv_0}{dr} + \frac{\lambda v_0}{\sqrt{p(r)}} + \frac{v_0}{2r} + \frac{p'(r) v_0}{4p(r)} + \frac{v_0}{\sqrt{p(r)} \int_0^r ds/\sqrt{p(s)}} \right\}, \tag{10.2.62}$$

with differentiation taken over the forward characteristic conoid.

On collecting the coefficients of $v_0(r, t)$ in the braces in (10.2.62) we determine that they behave like $3/2r + O(1)$ near $r = 0$. Consequently, if we equate (10.2.62) to zero to obtain a transport equation for $v_0(r, t)$, the solution is singular at $r = 0$. However, in view of (10.2.59), the leading term in the expansion (10.2.61) is given as $4\sigma S_{-2}(\sigma) v_0(r, t)$. Thus if we equate this term to $-6S_{-1}(\sigma) v_0(r, t)$, we find that

$$4\sigma S_{-2}(\sigma) + 6S_{-1}(\sigma) = 4\left[\sigma S_0''(\sigma) + \frac{3}{2} S_0'(\sigma)\right] = 0, \tag{10.2.63}$$

and on equating the resulting coefficient of $S_{-1}(\sigma)$ in (10.2.61) to zero, we now have

$$\frac{dv_0(r,t)}{dr} + \left[\frac{\lambda}{\sqrt{p(r)}} + \frac{1}{2r} + \frac{p'(r)}{4p(r)} - \left(2\sqrt{p(r)} \int_0^r \frac{ds}{\sqrt{p(s)}} \right)^{-1} \right] v_0(r,t) = 0, \tag{10.2.64}$$

and the singularity in the coefficient of $v_0(r, t)$ is eliminated.

Noting the properties of the function $f_a(x)$ defined in Example 7.1, we determine from (10.2.63) that $S_0(\sigma) = (1/\sqrt{\sigma})H(\sigma)$. The solution of (10.2.64) is

$$v_0(r,t) = c_0 \left[\frac{1}{r\sqrt{p(r)}} \int_0^r \frac{ds}{\sqrt{p(s)}}\right]^{1/2} e^{-\lambda t} \quad \text{on} \quad t = \int_0^r \frac{ds}{\sqrt{p(s)}}, \tag{10.2.65}$$

where c_0 is an arbitrary constant. With $p(r)$ = constant and $q(\mathbf{x}) = \lambda = 0$, (10.2.60) must reduce to the fundamental solution of the two-dimensional wave equation. On comparing our result with (6.7.38), we find that $c_0 = 1/2\pi\sqrt{p(0)}$. Then, the causal fundamental solution is given approximately as

$$u(r,t) \approx \frac{1}{2\pi}[p(0)p(r)]^{-1/4} \left[\frac{1}{r}\int_0^r \frac{ds}{\sqrt{p(0)p(s)}}\right]^{1/2} \frac{e^{-\lambda t} H\left[t - \int_0^r ds/\sqrt{p(s)}\right]}{\sqrt{t^2 - \left[\int_0^r ds/\sqrt{p(s)}\right]^2}}. \tag{10.2.66}$$

We have replaced the argument σ of the Heaviside function by $t - \int_0^r ds/\sqrt{p(s)}$, which is valid for $t > 0$. We do not obtain any further terms in the expansion (10.2.60).

The only difference that occurs in the case of *three dimensions* (for our purposes) is that in the Laplacian operator in spherical coordinates the term $\partial/\partial r$ is multiplied by 2 rather than 1 as is the case for cylindrical coordinates. This has the effect that we must set

$$4\sigma S_{-2}(\sigma) + 8S_{-1}(\sigma) = 4\left[\sigma S_0''(\sigma) + 2S_0'(\sigma)\right] = 0, \tag{10.2.67}$$

in order to eliminate the singularity in the equation for $v_0(r,t)$ at the vertex of the characteristic conoid. The solution of this equation for $S_0(\sigma)$ is $S_0(\sigma) = \delta(\sigma)$. On solving for $v_0(r,t)$ and choosing the arbitrary constant by comparing the result with (7.4.30), we obtain

$$u(r,t) \approx \frac{1}{4\pi r}\,[p(0)p(r)]^{-1/4}\, e^{-\lambda t}\, \delta\left[\sqrt{p(0)}\,t - \int_0^r \sqrt{\frac{p(0)}{p(s)}}\,ds\right], \tag{10.2.68}$$

which is valid for $t > 0$. Again we do not obtain further terms in the expansion.

We have shown how the behavior of solutions of hyperbolic equations near singularity regions can be determined without having to find the full solution. Further examples are given in the exercises.

Exercises 10.2

10.2.1. Use expansions of the form (10.2.14) to account for the singularity in the data for the following problem: $u_{tt}(x,t) - \gamma^2 u_{xx}(x,t) + c^2 u(x,t) = 0$, $-\infty < x < \infty$, $t > 0$, $u(x,0) = H(x)$, $u_t(x,0) = \gamma\delta(x)$, $-\infty < x < \infty$.

10.2.2. Solve the problem of Exercise 10.2.1 if the equation is replaced by $u_{tt}(x,t) - \gamma^2 u_{xx}(x,t) - c^2 u(x,t) = 0$, $-\infty < x < \infty$, $t > 0$, and the data remain the same.

10.2.3. Expand the solution (10.2.35) near the characteristic $x - \gamma t = 0$ and show that the first two terms agree with those obtained by applying the expansion method directly.

10.2.4. Consider the solution $u(x,t)$ of the problem (10.2.31) and (10.2.34), as given in (10.2.35). Show that $u_t(x,t) + 2\lambda u(x,t)$ is the solution of (10.2.31)–(10.2.32). Determine the solution and verify the result (10.2.33).

10.2.5. Use an expansion of the form (10.2.36) to discuss the approximate solution of the initial value problem $u_{tt}(x,t) - \gamma^2 u_{xx}(x,t) + 2\lambda u_t(x,t) = 0$, $-\infty < x < \infty$, $t > 0$, $u(x,0) = \begin{cases} 0, & x < 0, \\ \sqrt{x}, & x > 0, \end{cases}$ $u_t(x,0) = 0$. *Hint*: Use the generalized functions $f_a(x)$ of Example 7.1.

10.2.6. Solve the wave equation $u_{tt}(x,y,z) = \nabla^2 u(x,y,z)$ in three dimensions with the Cauchy data $u(x,y,z,0) = 0$, $u_t(x,y,z,0) = \begin{cases} \alpha, & r < a, \\ \beta, & r > a, \end{cases}$ if $r^2 = x^2 + y^2 + z^2$ and α, β, and a are constants. Use an expansion of the form (10.2.36).

10.2.7. Replace c by ic in the fundamental solution (10.2.52) of the Klein- Gordon equation to obtain the solution of the following problem and discuss the solution and its expansion in terms of the hyperbolic distance, in the manner in which the Klein-Gordon equation was analyzed in the text. $u_{tt}(x,y,z) = \gamma^2 \nabla^2 u(x,y,z) - c^2 u(x,y,z,t) = 0$, $-\infty < x,y,z < \infty$, $t > 0$, $u(x,y,z,0) = 0$, $u_t(x,y,z,0) = \delta(x)\delta(y)\delta(z)$.

10.2.8. Let $p(\mathbf{x}) = p(r) = \gamma^2(1 + a^2 r^2)$ in (10.2.58). Determine the explicit forms of $u(r,t)$ as given in (10.2.66) and (10.2.68) in the two- and three-dimensional cases, respectively.

10.2.9. Given the hyperbolic system $\mathbf{u}_t(x,t) + A\mathbf{u}_x(x,t) + B\mathbf{u} = \mathbf{0}$, consider an expansion $\mathbf{u}(x,t) = \sum_{j=0}^{\infty} \mathbf{v}_j(x,t) S_j[\phi(x,t)]$, where the S_j satisfy $S_j' = S_{j-1}$. Insert the expansion for $\mathbf{u}(x,t)$ into the equation for $\mathbf{u}(x,t)$ and equate coefficients of like terms S_j. Show that $\phi(x,t)$ must be a characteristic of the hyperbolic equation and that the $\mathbf{v}_j(x,t)$ are given in terms of eigenvectors of the matrix from which the characteristics are determined.

10.2.10. Express the telegrapher's equation (10.2.21) as a system of equations by setting $u_t(x,t) = v(x,t)$ and $u_x(x,t) = w(x,t)$. Use the initial data (10.2.34) to obtain $\mathbf{u}(x,0) = [v(x,0), w(x,0)] = [\delta(x), 0] = \delta(x)\mathbf{i}$. Solve the Cauchy problem for this system with the preceding data using the method of Exercise 10.2.9. Compare the results with those obtained by solving the telegrapher's equation directly by the expansion method given in the text.

10.3 ASYMPTOTIC SIMPLIFICATION OF EQUATIONS

In the first two sections of this chapter and in Chapter 9, we discussed techniques for obtaining approximate solutions of initial and boundary value problems for PDEs. In this section we consider methods for replacing the equations themselves by simpler, more easily solvable equations. Although the use of perturbation and asymptotic methods also gives rise to the consideration of simpler equations, the approach we use differs from that employed in the aforementioned methods. We show how to construct the simplified equations but do not, in general, indicate how they may be used to solve specific initial and boundary value problems. As shown, general information about the behavior of the solutions of the given full equations can be inferred from their simplified forms.

Asymptotic Simplification of the Dissipative Wave Equation

To fix ideas we consider the *dissipative wave equation* with constant α and c and $c > 0$,

$$u_{tt}(x,t) - c^2 u_{xx}(x,t) + u_t(x,t) - \alpha u_x(x,t) = 0. \tag{10.3.1}$$

Assuming that the point (x_0, t_0) is a discontinuity or singular point for the data for (10.3.1), we study its effect on the solutions of (10.3.1) by setting $x - x_0 = \epsilon\sigma$ and $t - t_0 = \epsilon\tau$. This yields

$$u_{\tau\tau}(\sigma,\tau) - c^2 u_{\sigma\sigma}(\sigma,\tau) + \epsilon[u_\tau(\sigma,\tau) - \alpha u_\sigma(\sigma,\tau)] = 0. \tag{10.3.2}$$

Although we have, in effect, introduced a boundary layer stretching to obtain (10.3.2), we do not proceed as in boundary layer theory by expanding $u(\sigma,\tau)$ in powers of ϵ. Instead, we introduce an approximate factorization of (10.3.2). If $\epsilon = 0$, then (10.3.2) can be factored exactly as was done in (2.1.1), but if $\epsilon \neq 0$ we cannot factor (10.3.2) in a simple fashion. Yet since ϵ is small, it can be factored approximately. With $\partial_\sigma = \partial/\partial\sigma$, $\partial_\tau = \partial/\partial\tau$ in (10.3.2), we have

$$[\partial_\tau^2 - c^2\partial_\sigma^2 + \epsilon\partial_\tau - \epsilon\alpha\partial_\sigma]u = [\partial_\tau - c\partial_\sigma + \epsilon a + O(\epsilon^2)][\partial_\tau + c\partial_\sigma + \epsilon b + O(\epsilon^2)]u = 0, \tag{10.3.3}$$

where a and b are to be specified. (We may assume that a and b are constants.) Multiplying out the two first order operators in (10.3.3) and comparing the results with the second order operator in (10.3.3), we find that $a = (c - \alpha)/2c$, $b = (c + \alpha)/2c$.

Since the order of the operators in (10.3.3) can be interchanged, we see that to order ϵ^2, (10.3.2) is equivalent to the two first order equations

$$\begin{cases} u_\tau(\sigma,\tau) - cu_\sigma(\sigma,\tau) + \epsilon((c-\alpha)/2c)u(\sigma,\tau) = O(\epsilon^2), \\ u_\tau(\sigma,\tau) + cu_\sigma(\sigma,\tau) + \epsilon((c+\alpha)/2c)u(\sigma,\tau) = O(\epsilon^2). \end{cases} \tag{10.3.4}$$

Neglecting the $O(\epsilon^2)$ terms in (10.3.4), we obtain the solutions

$$u(\sigma,\tau) = \begin{cases} f(\sigma + c\tau)\exp(-\epsilon\tau(c-\alpha)/2c), \\ g(\sigma - c\tau)\exp(-\epsilon\tau(c+\alpha)/2c), \end{cases} \tag{10.3.5}$$

where f and g are arbitrary functions. If $|\alpha| < c$, both solutions decay exponentially as $t - t_0 = \epsilon\tau \to \infty$. If $|\alpha| > c$, one solution grows exponentially as $t - t_0 = \epsilon\tau \to \infty$. It is easy to show that the initial value problem for equation (10.3.1) is unstable if $|\alpha| > c$. Since discontinuities in the solutions are carried along the characteristics $\sigma \pm c\tau$ = constant, we see that (10.3.5) is useful in describing the propagation of discontinuities in the solutions of (10.3.1). In fact, the neglected $O(\epsilon^2)$ terms in (10.3.4) are antiderivative or integral operators that have the effect of smoothing out singularities in the solutions. The expressions for $u(\sigma, \tau)$ in (10.3.5) have the general form of the solutions given in the preceding section.

To study the solutions of (10.3.1) at large times and in the far field, we set $x = \sigma/\epsilon$ and $t = \tau/\epsilon$ and obtain

$$\epsilon[u_{\tau\tau}(\sigma, \tau) - c^2 u_{\sigma\sigma}(\sigma, \tau)] + u_\tau(\sigma, \tau) - \alpha u_\sigma(\sigma, \tau) = 0. \tag{10.3.6}$$

It follows from (10.3.6) that $u_{\tau\tau} = \alpha u_{\sigma\tau} - \epsilon[u_{\tau\tau\tau} - c^2 u_{\sigma\sigma\tau}] = \partial(\alpha u_\tau)/\partial\sigma + O(\epsilon) = \alpha^2 u_{\sigma\sigma} + O(\epsilon)$, so that (10.3.6) can be replaced by

$$u_\tau(\sigma, \tau) - \alpha u_\sigma(\sigma, \tau) + \epsilon(\alpha^2 - c^2) u_{\sigma\sigma}(\sigma, \tau) = O(\epsilon^2). \tag{10.3.7}$$

If $|a| < c$, so that the initial value problem for (10.3.1) is stable, we see that on neglecting the $O(\epsilon^2)$ terms in (10.3.7), we have a *diffusion equation*,. In the original variables this equation has the form

$$u_t(x, t) - \alpha u_x(x, t) + \epsilon(\alpha^2 - c^2) u_{xx}(x, t) = 0, \tag{10.3.8}$$

and this shows that for large x and t the dissipative wave equation yields a *wave motion* along the line $x + \alpha t = 0$ modified by a diffusion effect represented by the second derivative term. That is, if $\epsilon = 0$ in (10.3.8), we have undirectional wave motion along $x + \alpha t = 0$. The $O(\epsilon^2)$ terms neglected in (10.3.7) are higher derivatives of $u(\sigma, \tau)$. Since the diffusion effect smooths out the solutions, these terms can be disregarded if they are multiplied by higher powers of ϵ. Our conclusions are consistent with those obtained in Example 5.15 and the discussion in Section 1.2.

We have determined that there are two types of wave motion associated with (10.3.1). In (10.3.4) we obtained approximate equations corresponding to waves moving with speed c, and in (10.3.7)–(10.3.8) we obtained an equation characterizing wave motion with the speed $|\alpha|$. The *principal part* of (10.3.1)—that is, $u_{tt} - c^2 u_{xx}$—yields traveling waves moving to the right or left with speed c. The *reduced part* of (10.3.1)—that is, $u_t - \alpha u_x$—gives rise to a traveling wave moving with speed $|\alpha|$. Wave motions with both speeds play a role in the solution of (10.3.1) in different regions, as we have shown. In this section we show how to construct simplified approximating equations that characterize basic aspects of the given full equations or systems of equations, in a more general setting. Since our method is most easily presented in terms of matrix theory, we restrict our discussion to systems of equations. Single higher-order PDEs can always be expressed in system form.

With $v = u_x$ and $w = u_t$, the dissipative equation (10.3.1) can be written as a first order system,

$$\mathbf{u}_t(x, t) + A\mathbf{u}_x(x, t) + B\mathbf{u}(x, t) = \mathbf{0}, \tag{10.3.9}$$

where

$$\mathbf{u}(x,t) = \begin{bmatrix} v(x,t) \\ w(x,t) \end{bmatrix}, \quad A = \begin{bmatrix} 0 & -1 \\ -c^2 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 \\ -\alpha & 1 \end{bmatrix}. \tag{10.3.10}$$

Then (10.3.2) is equivalent to

$$\mathbf{u}_\tau(\sigma,\tau) + A\mathbf{u}_\sigma(\sigma,\tau) + \epsilon B\mathbf{u}(\sigma,\tau) = \mathbf{0}, \tag{10.3.11}$$

and (10.3.6) corresponds to

$$\epsilon[\mathbf{u}_\tau(\sigma,\tau) + A\mathbf{u}_\sigma(\sigma,\tau)] + B\mathbf{u}(\sigma,\tau) = \mathbf{0}. \tag{10.3.12}$$

We introduce asymptotic decompositions of (10.3.11) and (10.3.12) that correspond to the reduced equations obtained on setting $\epsilon = 0$ in each of these systems of equations.

Eigenspaces and Projection Matrices

To carry out these decompositions for (10.3.11) and (10.3.12) and general systems, we require the following result from matrix theory. Let M be an $n \times n$ matrix and $C(\lambda) = |M - \lambda I|$ its *characteristic polynomial.* We factor $C(\lambda)$ as $C(\lambda) = C_1(\lambda)C_2(\lambda)$, where the factors $C_1(\lambda)$ and $C_2(\lambda)$ contain no common eigenvalues as roots. Introducing the partial fraction decomposition $1/C(\lambda) = c_1(\lambda)/C_1(\lambda) + c_2(\lambda)/C_2(\lambda)$, we have $1 = c_1(\lambda)C_2(\lambda) + c_2(\lambda)C_1(\lambda)$. Then the matrices

$$P_1 = c_1(M)C_2(M), \qquad P_2 = c_2(M)C_1(M) \tag{10.3.13}$$

are *projection matrices* that project vectors into the *eigenspaces* spanned by the eigenvectors corresponding to the factors $C_1(\lambda)$ and $C_2(\lambda)$, respectively. They have the following additional properties $P_1 + P_2 = I$, $P_1P_2 = P_2P_1 = 0$, $P_1M = MP_1$, $P_2M = MP_2$, $P_1^2 = P_1$, $P_2^2 = P_2$. We assume in our discussion that the $n \times n$ matrix M has n linearly independent eigenvectors.

As an example, we construct projection operators for the matrices A and B defined in (10.3.10). The characteristic polynomial for A is $C(\lambda) = |A - \lambda I| = \lambda^2 - c^2 = (\lambda + c)(\lambda - c)$. With $C_1(\lambda) = \lambda + c$ and $C_2(\lambda) = \lambda - c$, we have

$$\frac{1}{\lambda^2 - c^2} = \frac{-1/2c}{\lambda + c} + \frac{1/2c}{\lambda - c} = \frac{(-1/2c)(\lambda - c) + (1/2c)(\lambda + c)}{\lambda^2 - c^2}. \tag{10.3.14}$$

Thus $c_1(\lambda) = -c_2(\lambda) = 1/2c$, $c_1(\lambda)C_2(\lambda) = (-1/2c)(\lambda - c)$, and $c_2(\lambda)C_1(\lambda) = (1/2c)(\lambda + c)$. Finally, the projection operators are given as

$$P_1 = -\frac{1}{2c}[A - cI] = \frac{1}{2c}\begin{bmatrix} c & 1 \\ c^2 & c \end{bmatrix}, \; P_2 = \frac{1}{2c}[A + cI] = \frac{1}{2c}\begin{bmatrix} c & -1 \\ -c^2 & c \end{bmatrix}. \tag{10.3.15}$$

The matrix A has two eigenvalues $\lambda_1 = -c$ and $\lambda_2 = c$. The corresponding eigenvectors are

$$\mathbf{r}_1 = \begin{bmatrix} 1 \\ c \end{bmatrix}, \qquad A\mathbf{r}_1 = -c\mathbf{r}_1, \qquad \mathbf{r}_2 = \begin{bmatrix} 1 \\ -c \end{bmatrix}, \qquad A\mathbf{r}_2 = c\mathbf{r}_2, \quad (10.3.16)$$

and they are linearly independent. We have $P_1\mathbf{r}_1 = \mathbf{r}_1$, $P_1\mathbf{r}_2 = \mathbf{0}$, $P_2\mathbf{r}_1 = \mathbf{0}$, and $P_2\mathbf{r}_2 = \mathbf{r}_2$. Since any two-component vector can be expressed as a linear combination of $\mathbf{r}_1$ and $\mathbf{r}_2$, we see that P_1 and P_2 project vectors into the required eigenspaces (these are one-dimensional in this case).

For the matrix B in (10.3.10), we have $C(\lambda) = |B - \lambda I| = \lambda(\lambda - 1)$. With $C_1(\lambda) = \lambda$ and $C_2(\lambda) = \lambda - 1$, we have

$$\frac{1}{\lambda(\lambda-1)} = \frac{-1}{\lambda} + \frac{1}{\lambda-1} = \frac{-(\lambda-1)+\lambda}{\lambda(\lambda-1)}, \qquad (10.3.17)$$

so that

$$P_1 = -[B - I] = \begin{bmatrix} 1 & 0 \\ \alpha & 0 \end{bmatrix}, \qquad P_2 = B = \begin{bmatrix} 0 & 0 \\ -\alpha & 1 \end{bmatrix}. \qquad (10.3.18)$$

The eigenvectors $\mathbf{r}_1$ and $\mathbf{r}_2$, which correspond to the eigenvalues $\lambda_1 = 0$ and $\lambda_2 = 1$, respectively, of B are

$$\mathbf{r}_1 = \begin{bmatrix} 1 \\ \alpha \end{bmatrix}, \qquad B\mathbf{r}_1 = \mathbf{0}, \qquad \mathbf{r}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad B\mathbf{r}_2 = \mathbf{r}_2. \qquad (10.3.19)$$

Again, $\mathbf{r}_1$ and $\mathbf{r}_2$ are linearly independent, and P_1 and P_2 have all the properties required of projection operators.

Asymptotic Simplification of the System Form of the Dissipative Wave Equation

We now use these projection operators to construct asymptotic decompositions of (10.3.11) and (10.3.12). Beginning with (10.3.11), we set

$$\mathbf{u}(\sigma,\tau) = V(\sigma,\tau)\mathbf{r}_1 + \epsilon W(\sigma,\tau)\mathbf{r}_2, \qquad (10.3.20)$$

where $V(\sigma,\tau)$ and $W(\sigma,\tau)$ are scalar functions (to be determined) that may depend on ϵ, and $\mathbf{r}_1$ and $\mathbf{r}_2$ are the eigenvectors of A given in (10.3.16). Inserting (10.3.20) into (10.3.11) gives

$$(V_\tau - cV_\sigma)\mathbf{r}_1 + \epsilon\left\{(W_\tau + cW_\sigma)\mathbf{r}_2 + V\begin{bmatrix} 0 \\ c-\alpha \end{bmatrix}\right\} + \epsilon^2\left\{-W\begin{bmatrix} 0 \\ c+\alpha \end{bmatrix}\right\} = \mathbf{0}. \qquad (10.3.21)$$

Applying the projection operator P_1 to (10.3.21) (i.e., multiplying across on the left by P_1) gives

$$\left\{V_\tau - cV_\sigma + \epsilon\frac{c-\alpha}{2c}V - \epsilon^2\frac{c+\alpha}{2c}W\right\}\mathbf{r}_1 = \mathbf{0}, \qquad (10.3.22)$$

since $P_1\mathbf{r}_2 = \mathbf{0}$. Applying P_2 to (10.3.21) yields

$$\epsilon\left\{W_\tau + cW_\sigma - \frac{c-\alpha}{2c}V\right\}\mathbf{r}_2 + \epsilon^2\frac{c+\alpha}{2c}W\mathbf{r}_2 = \mathbf{0}. \tag{10.3.23}$$

Retaining terms up to order ϵ in (10.3.22) as was done in our discussion of the scalar example, we find that

$$V_\tau(\sigma,\tau) - cV_\sigma(\sigma,\tau) + \epsilon\frac{c-\alpha}{2c}V(\sigma,\tau) = 0, \tag{10.3.24}$$

and this agrees with (10.3.4). The leading term of (10.3.23) shows that $W(\sigma,\tau)$ can be expressed in terms of $V(\sigma,\tau)$ apart from an arbitrary function of $\sigma - c\tau$. If this arbitrary function is disregarded, we can replace $W(\sigma,\tau)$ in (10.3.22) by a function of $V(\sigma,\tau)$ given in integral form. Higher approximations may be obtained by expanding $W(\sigma,\tau)$ in a power series in ϵ and solving (10.3.23) for the coefficients as functions of $V(\sigma,\tau)$, but we do not pursue this matter here. We do note that if $\mathbf{r}_1$ and $\mathbf{r}_2$ are interchanged in (10.3.20), we obtain the second form of the equations for $u(\sigma,\tau)$ given in (10.3.4) as the equation satisfied by $V(\sigma,\tau)$.

Next we consider the system (10.3.12) and again represent $u(\sigma,\tau)$ as in (10.3.20), with $\mathbf{r}_1$ and $\mathbf{r}_2$ as defined in (10.3.19). Inserting (10.3.20) into (10.3.12) yields

$$\epsilon\left\{V_\tau\mathbf{r}_1 + V_\sigma\begin{bmatrix}-\alpha\\-c^2\end{bmatrix} + W\mathbf{r}_2\right\} + \epsilon^2\left\{W_\tau\mathbf{r}_2 + W_\sigma\begin{bmatrix}-1\\0\end{bmatrix}\right\} = \mathbf{0}, \tag{10.3.25}$$

since $B\mathbf{r}_1 = \mathbf{0}$. Applying the projection operator P_1, as given in (10.3.18), to (10.3.25) gives

$$\{V_\tau - \alpha V_\sigma - \epsilon W_\sigma\}\,\mathbf{r}_1 = \mathbf{0}, \tag{10.3.26}$$

as is easily seen. Applying $P_2 = B$ to (10.3.25) results in

$$\{(\alpha^2 - c^2)V_\sigma + W\}\,\mathbf{r}_2 + \epsilon\{W_\tau + \alpha W_\sigma\}\,\mathbf{r}_2 = \mathbf{0}. \tag{10.3.27}$$

We solve (10.3.27) by expanding $W(\sigma,\tau)$ as

$$W(\sigma,\tau) = \sum_{j=0}^{\infty} W_j(\sigma,\tau)\epsilon^j. \tag{10.3.28}$$

Then $W_0(\sigma,\tau)$ has the form $W_0(\sigma,\tau) = (c^2 - \alpha^2)V_\sigma$. We must also insert (10.3.28) into (10.3.26), and on retaining only terms up to order ϵ, we have

$$V_\tau(\sigma,\tau) - \alpha V_\sigma(\sigma,\tau) - \epsilon(c^2 - \alpha^2)V_{\sigma\sigma}(\sigma,\tau) = 0. \tag{10.3.29}$$

This result has the form of (10.3.7). Clearly, the $W_j(\sigma,\tau)$ in (10.3.28) are all expressed in terms of increasingly higher-order derivatives of $V(\sigma,\tau)$, as is seen in (10.3.27). Thus, in neglecting $O(\epsilon^2)$ terms to obtain (10.3.29), we are disregarding presumably

smoother higher derivative terms in $V(\sigma, \tau)$, in agreement with the discussion given for the scalar version of this example.

For (10.3.12), we do not consider the effect of interchanging $\mathbf{r}_1$ and $\mathbf{r}_2$ in the assumed solution form (10.3.20) because on setting $\epsilon = 0$ in (10.3.12), we obtain $B\mathbf{u}(\sigma, \tau) = \mathbf{0}$, and this has the solution $\mathbf{u}(\sigma, \tau) = V(\sigma, \tau)\mathbf{r}_1$, where $V(\sigma, \tau)$ is arbitrary. It is this expression that was perturbed around in (10.3.20). Similarly, with $\epsilon = 0$ in (10.3.11), we have $u_\tau(\sigma, \tau) + Au_\sigma(\sigma, \tau) = \mathbf{0}$, and this has the solutions $\mathbf{u}(\sigma, \tau) = \hat{V}(\sigma + c\tau)\mathbf{r}_1$ and $\mathbf{u}(\sigma, \tau) = \hat{V}(\sigma - c\tau)\mathbf{r}_2$ with arbitrary $\hat{V}$ and $\hat{V}$. As a result, we considered perturbations corresponding to each of the eigenvectors $\mathbf{r}_1$ and $\mathbf{r}_2$ in our discussion of (10.3.11).

The system obtained on setting $\epsilon = 0$ in (10.3.11) or (10.3.12) is known as the *reduced system*. We constructed asymptotic simplifications of the full systems corresponding to solutions for the reduced system. The solutions of the reduced system are related to the properties of a relevant coefficient matrix. Based on a decomposition of that matrix according to its eigenvectors (even though this was not done explicitly), a decomposition of the full system was carried out. We now show how this technique may be applied to a general class of systems of equations.

Asymptotic Simplification of Systems of Equations

We consider systems of equations of the form

$$\mathbf{u}_t(x, t) + A\mathbf{u}_x(x, t) + \epsilon N[\mathbf{u}(x, t)] = \mathbf{0}, \tag{10.3.30}$$

where $\mathbf{u}(x, t)$ is an n-component vector, A is a constant $n \times n$ matrix, and N is a linear or nonlinear differential operator. The matrix A is assumed to have n linearly independent eigenvectors. With $C(\lambda) = |A - \lambda I|$ as the characteristic polynomial of A, we assume that it can be factored as $C(\lambda) = C_1(\lambda)C_2(\lambda)$, where $C_1(\lambda)$ and $C_2(\lambda)$ have no common eigenvalues as roots. Further, it is assumed that $0 < m < n$ independent eigenvectors are associated with the eigenvalues of $C_1(\lambda)$ and the remaining $n - m$ independent eigenvectors are associated with $C_2(\lambda)$. The eigenspaces spanned by these two sets of eigenvectors have the property that any n-component vector can be expressed as a sum of two vectors, each of which lies in one of the eigenspaces.

Proceeding as before, but replacing M by A in the equations, we can construct the projection operators P_1 and P_2 associated with the eigenspaces corresponding to $C_1(\lambda)$ and $C_2(\lambda)$, respectively. If $\mathbf{v}$ lies in the eigenspace related to $C_1(\lambda)$ and $\mathbf{w}$ lies in the eigenspace related to $C_2(\lambda)$, we have $P_1\mathbf{v} = \mathbf{v}$, $P_1\mathbf{w} = \mathbf{0}$, $P_2\mathbf{w} = \mathbf{w}$, and $P_2\mathbf{v} = \mathbf{0}$. With $\mathbf{v}(x, t)$ and $\mathbf{w}(x, t)$ as elements of the two eigenspaces defined previously, we look for a solution of (10.3.30) in the form

$$\mathbf{u}(x, t) = \mathbf{v}(x, t) + \epsilon\mathbf{w}(x, t). \tag{10.3.31}$$

In doing so it is assumed that we are interested in perturbing $\mathbf{u}(x, t)$ around a solution of the reduced system $\mathbf{u}_t(x, t) + A\mathbf{u}_x(x, t) = \mathbf{0}$, which lies in the eigenspace associated with $\mathbf{v}(x, t)$.

Inserting (10.3.31) into (10.3.30) yields

$$\mathbf{v}_t(x,t) + A\mathbf{v}_x(x,t) + \epsilon\{\mathbf{w}_t(x,t) + A\mathbf{w}_x(x,t) + N[\mathbf{v}(x,t) + \epsilon\mathbf{w}(x,t)]\} = \mathbf{0}. \tag{10.3.32}$$

We assume that $N[\mathbf{v}(x,t) + \epsilon\mathbf{w}(x,t)]$ can be expanded in powers of ϵ and that $N[\mathbf{v}(x,t) + \epsilon\mathbf{w}(x,t)] = N_0[\mathbf{v}(x,t)] + O(\epsilon)$. Multiplying (10.3.32) by the projection operator P_1 gives

$$\mathbf{v}_t(x,t) + A\mathbf{v}_x(x,t) + \epsilon P_1 N[\mathbf{v}(x,t) + \epsilon\mathbf{w}(x,t)] = \mathbf{0}. \tag{10.3.33}$$

To obtain (10.3.33) we have used the facts that $P_1 A = AP_1$ and that P_1 is a constant matrix. Thus, for example, $P_1 A\mathbf{v}_x(x,t) = AP_1\mathbf{v}_x(x,t) = A(\partial/\partial x)[P_1\mathbf{v}(x,t)] = A\mathbf{v}_x(x,t)$ and $P_1\mathbf{w}_t(x,t) = (\partial/\partial t)[P_1\mathbf{w}(x,t)] = \mathbf{0}$. Similarly, we obtain, on multiplying (10.3.32) by P_2,

$$\epsilon\{\mathbf{w}_t(x,t) + A\mathbf{w}_x(x,t) + P_2 N[\mathbf{v}(x,t) + \epsilon\mathbf{w}(x,t)]\} = \mathbf{0}. \tag{10.3.34}$$

Next, we expand $\mathbf{w}(x,t)$ as

$$\mathbf{w}(x,t) = \sum_{j=0}^{\infty} \mathbf{w}_j(x,t)\epsilon^j \tag{10.3.35}$$

and insert this series into (10.3.33)–(10.3.34). Using the expansion of $N[\mathbf{v}(x,t) + \epsilon\mathbf{w}(x,t)]$ in powers of ϵ, we equate like powers of ϵ in (10.3.34). This yields

$$\frac{\partial\mathbf{w}_0(x,t)}{\partial t} + A\frac{\partial\mathbf{w}_0(x,t)}{\partial x} + P_2 N_0[\mathbf{v}(x,t)] = \mathbf{0} \tag{10.3.36}$$

for $\mathbf{w}_0(x,t)$ and similar equations for the remaining $\mathbf{w}_j(x,t)$. The equations for $\mathbf{w}_j(x,t)$ are to be solved only in terms of $\mathbf{v}(x,t)$, and any solutions of the homogeneous versions of these equations [which do not depend on $\mathbf{v}(x,t)$] are to be discarded. The resulting expressions for the $\mathbf{w}_j(x,t)$ are then inserted into (10.3.33). This yields the *simplified system* for $\mathbf{v}(x,t)$ since $\mathbf{v}(x,t)$ belongs to a lower-dimensional space than $\mathbf{u}(x,t)$. Thus (10.3.33) effectively contains fewer equations in fewer unknowns than does (10.3.30).

The equation (10.3.33) for $\mathbf{v}(x,t)$ may have a very complicated form. However, we do not expand $\mathbf{v}(x,t)$ in powers of ϵ since this—in combination with (10.3.35)—would yield the conventional perturbation solution of (10.3.30). That solution may contain secular behavior that we are trying to avoid. Instead, we truncate (10.3.33) at some order of ϵ. This new equation can then be used as a basis for a perturbation result. For example, retaining only $O(\epsilon)$ terms in (10.3.33) gives $\mathbf{v}_t(x,t) + A\mathbf{v}_x(x,t) + \epsilon P_1 N_0[\mathbf{v}(x,t)] = \mathbf{0}$. This equation corresponds to (10.3.24) that was found above. Combined with (10.3.36), this yields a leading-order approximation to a solution of (10.3.30). The leading-order perturbation result can always be retrieved by expanding $\mathbf{v}(x,t)$ in a series in powers of ϵ. Thus (10.3.33) represents a more general result, and it can be constructed for each eigenspace associated with the matrix A.

For example, if $\lambda = \lambda_1$ is a simple eigenvalue of A and $\mathbf{r}_1$ is the corresponding eigenvector, we may consider $C_1(\lambda) = \lambda - \lambda_1$. The eigenspace for $C_1(\lambda)$ is then one-dimensional and is spanned by $\mathbf{r}_1$. Then we can set $\mathbf{v}(x,t) = V(x,t)\mathbf{r}_1$, where $V(x,t)$ is a scalar function, in (10.3.31). Since P_1 projects vectors into the eigenspace spanned by $\mathbf{r}_1$, we have $P_1 N[\mathbf{v}(x,t) + \epsilon \mathbf{w}(x,t)] = \alpha(V(x,t), \mathbf{w}(x,t), \epsilon)\mathbf{r}_1$, where α is a scalar function. Thus (10.3.33) becomes

$$[V_t(x,t) + \lambda_1 V_x(x,t) + \epsilon\alpha(V(x,t), \mathbf{w}(x,t), \epsilon)]\mathbf{r}_1 = \mathbf{0}. \tag{10.3.37}$$

This is scalar equation for $V(x,t)$ once the $\mathbf{w}_j(x,t)$ are specified in terms of $V(x,t)$.

In the following we show how systems of equations of the form (10.3.30) arise in the *theory of viscous fluid flow* and the *theory of water waves*. The asymptotic decompositon method is used to obtain simplified forms for these systems. We remark that the simplification method presented can also be applied if the matrix A in (10.3.30) is a slowly varying function of x and t. Also, a related simplification method can be used even if the given equation is not in the form (10.3.30), as we have shown for the dissipative wave equation. However, we restrict our discussion to the constant coefficient case. An extension of this method to higher space dimensions is not obvious since the reduced system would generally contain more than one coefficient matrix to be dealt with in the foregoing manner. However, if one of the spatial variables can be distinguished from the others for some reason, this procedure can be carried out, but this is not considered here.

Navier-Stokes Equations

The *Navier-Stokes equations* that describe *viscous fluid flow* in one dimension were given in Section 8.5 as

$$\rho_t + \rho u_x + u\rho_x = 0, \tag{10.3.38}$$

$$\rho u_t + \rho u u_x + p_x = \frac{4}{3}\mu u_{xx}, \tag{10.3.39}$$

$$\rho c_v[T_t + uT_x] + pu_x = kT_{xx} + \frac{4}{3}\mu(u_x)^2, \tag{10.3.40}$$

$$p = R\rho T. \tag{10.3.41}$$

Here $\rho = \rho(x,t)$ is the density, $u = u(x,t)$ is the velocity, $p = p(x,t)$ is the pressure, and $T = T(x,t)$ is the temperature. The parameter μ is the (constant) coefficient of viscosity, k is the coefficient of heat conduction, c_p and c_v are the specific heats at constant pressure and volume, respectively, and $R = c_p - c_v$ is the gas constant. We have assumed that there are no external forces or heat sources.

To apply the asymptotic decomposition method to these equations, we *perturb* $\rho(x,t)$ and $T(x,t)$ around a constant density and temperature and $u(x,t)$ around a zero velocity. Thus we set

$$u(x,t) = \epsilon\,\hat{u}(x,t),\ \rho(x,t) = \rho_0 + \epsilon\,\hat{\rho}(x,t),\ T(x,t) = T_0 + \epsilon\,\hat{T}(x,t), \tag{10.3.42}$$

where $\epsilon > 0$ is a small parameter. Also, we consider the solution in the far field and at large times and let $x = \sigma/\epsilon$, $t = \tau/\epsilon$.

We insert the above into (10.3.38)–(10.3.41) and use (10.3.41) to express p as a function of ρ and T, and only retain terms up to order ϵ in the resulting equations. If we set $\gamma = c_p/c_v$, the ratio of the specific heats, we find that with $c_p = 1/(\gamma - 1)$, we have $c_v = 1/\gamma(\gamma - 1)$, $R = 1/\gamma$, and $R/c_v = \gamma - 1$. By rescaling the variables and parameters in these equations, it is possible to eliminate the constants ρ_0 and T_0. Thus we may set $\rho_0 = T_0 = 1$. Also, for simplicity of notation we drop the carets and replace σ and τ by x and t.

As a result, the system can be written in the form

$$\mathbf{u}_t(x,t) + A\mathbf{u}_x(x,t) + \epsilon[B\mathbf{u}_x(x,t) + C\mathbf{u}_t(x,t) + D\mathbf{u}_{xx}(x,t)] = \mathbf{0}, \quad (10.3.43)$$

where ρ, u, T are functions of (x,t),

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1/\gamma & 0 & 1/\gamma \\ 0 & \gamma - 1 & 0 \end{bmatrix}, \; B = \begin{bmatrix} u & \rho & 0 \\ T/\gamma & u & \rho/\gamma \\ 0 & (\gamma-1)(T+\rho) & u \end{bmatrix},$$

$$C = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \rho & 0 \\ 0 & 0 & \rho \end{bmatrix}, \; D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -4\mu/3 & 0 \\ 0 & 0 & k(1-\gamma)\gamma \end{bmatrix}, \; \mathbf{u} = \begin{bmatrix} \rho \\ u \\ T \end{bmatrix}. \quad (10.3.44)$$

This system is of the form (10.3.30), and we note that the operator multiplying ϵ is nonlinear. The linear system $\mathbf{u}_t(x,t) + A\mathbf{u}_x(x,t) = \mathbf{0}$ obtained on setting $\epsilon = 0$ in (10.3.43) yields the *equations of linear acoustics*.

The characteristic polynomial for A is $C(\lambda) = |A - \lambda I| = \lambda(1-\lambda)(1+\lambda)$, so that the eigenvalues of A are $\lambda_0 = 0$, $\lambda_1 = 1$, and $\lambda_2 = -1$. The corresponding linearly independent eigenvectors are

$$\mathbf{r}_0 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \quad \mathbf{r}_1 = \begin{bmatrix} 1 \\ 1 \\ \gamma - 1 \end{bmatrix}, \quad \mathbf{r}_2 = \begin{bmatrix} 1 \\ -1 \\ \gamma - 1 \end{bmatrix}. \quad (10.3.45)$$

Using these results, we easily find that the reduced system

$$\mathbf{u}_t(x,t) + A\mathbf{u}_x(x,t) = \mathbf{0} \quad (10.3.46)$$

has the solutions

$$\mathbf{u}(x,t) = \alpha(x)\mathbf{r}_0, \qquad \mathbf{u}(x,t) = \beta(x-t)\mathbf{r}_1, \qquad \mathbf{u}(x,t) = \delta(x+t)\mathbf{r}_2, \quad (10.3.47)$$

where α, β, and δ are arbitrary functions. [Here $\delta(x)$ does not represent the Dirac delta function.] The general solution of (10.3.46) is a linear combination of these three solutions and it shows that two waves traveling to the right and to the left with unit speed occur in the solution of the acoustic equations.

Using the (conventional) *perturbation method* to solve (10.3.43), we set

$$\mathbf{u}(x,t) = \mathbf{u}_0(x,t) + \epsilon \mathbf{u}_1(x,t) + \cdots. \tag{10.3.48}$$

Then $\mathbf{u}_0(x,t)$ satisfies the reduced equation (10.3.46). If we choose $\mathbf{u}_0(x,t) = \beta(x-t)\mathbf{r}_1$, we readily find that $\mathbf{u}_1 = \mathbf{u}_1(x,t)$ satisfies

$$\frac{\partial \mathbf{u}_1}{\partial t} + A\frac{\partial \mathbf{u}_1}{\partial x} + \beta(x-t)\beta'(x-t)\hat{B}\mathbf{r}_1 - \beta(x-t)\beta'(x-t)\hat{C}\mathbf{r}_1 + \beta''(x-t)D\mathbf{r}_1 = \mathbf{0}, \tag{10.3.49}$$

where

$$\hat{B} = \begin{bmatrix} 1 & 1 & 0 \\ (\gamma-1)/\gamma & 1 & 1/\gamma \\ 0 & \gamma(\gamma-1) & 1 \end{bmatrix}, \qquad \hat{C} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{10.3.50}$$

Either by expressing $\mathbf{u}_1(x,t)$ and all the vectors in (10.3.49) as linear combinations of the eigenvectors (10.3.45) or by using the projection operators associated with the eigenspaces for A, the system (10.3.49) can be solved. It is then seen that *secular terms* result in the expression for $\mathbf{u}_1(x,t)$ and the expansion (10.3.48) becomes invalid for large t, specifically when $t = O(1/\epsilon)$. We do not carry out the solution of (10.3.49). Although we have shown peviously how secular effects can be removed, we do not apply those methods here. Instead, we show how the asymptotic simplification method enables us to avoid any problems with secular terms.

Asymptotic Simplification of the Navier-Stokes Equations: Burgers' and Heat Equations

The given perturbation result indicates that with increasing time, the solution (or wave) $\mathbf{u}_0 = \beta(x-t)\mathbf{r}_1$ of the reduced system loses its validity, and that nonlinear and higher derivative terms in (10.3.43) become significant. To determine their effect, we look for a solution of (10.3.43) in the form

$$\mathbf{u}(x,t) = \beta(x,t)\mathbf{r}_1 + \epsilon[\alpha(x,t)\mathbf{r}_0 + \delta(x,t)\mathbf{r}_2]. \tag{10.3.51}$$

This corresponds to (10.3.31) with $\mathbf{v}(x,t) = \beta(x,t)\mathbf{r}_1$ and $\mathbf{w}(x,t) = \alpha(x,t)\mathbf{r}_0 + \delta(x,t)\mathbf{r}_2$. The functions $\beta(x,t)$, $\alpha(x,t)$, and $\delta(x,t)$ are to be specified, but $\beta(x,t)\mathbf{r}_1$ lies in the eigenspace spanned by $\mathbf{r}_1$, whereas $\alpha(x,t)\mathbf{r}_0 + \delta(x,t)\mathbf{r}_2$ is a typical vector in the eigenspace spanned by $\mathbf{r}_0$ and $\mathbf{r}_2$. To obtain the projection operators for these eigenspaces, we set $C(\lambda) = |A - \lambda I| = C_1(\lambda)C_2(\lambda)$, where $C_1(\lambda) = 1 - \lambda$ and $C_2(\lambda) = \lambda(1+\lambda)$. The eigenvalue of A corresponding to $C_1(\lambda)$ is $\lambda_1 = 1$, and the remaining eigenvalues $\lambda_0 = 0$ and $\lambda_2 = -1$ correspond to the factor $C_2(\lambda)$. Since $1/(\lambda(1-\lambda)(1+\lambda)) = (1/2)/(1-\lambda) + (1+\lambda/2)/(\lambda(1+\lambda))$, we find that the projection operator P_1, which projects vectors into the eigenspace spanned by $\mathbf{r}_1$, is

$$P_1 = \frac{1}{2}A[I+A] = \frac{1}{2\gamma}\begin{bmatrix} 1 & \gamma & 1 \\ 1 & \gamma & 1 \\ \gamma-1 & \gamma(\gamma-1) & \gamma-1 \end{bmatrix}. \tag{10.3.52}$$

The projection operator P_2 corresponding to the eigenspace of $\mathbf{r}_0$ and $\mathbf{r}_2$ is

$$P_2 = I - P_1 = \frac{1}{2\gamma}\begin{bmatrix} 2\gamma-1 & -\gamma & -1 \\ -1 & \gamma & -1 \\ 1-\gamma & \gamma(1-\gamma) & 1+\gamma \end{bmatrix}. \tag{10.3.53}$$

Recalling the definition of $\mathbf{u}(x,t)$ given in (10.3.44), we see that (10.3.51) implies that $\rho(x,t) = \beta(x,t) + O(\epsilon)$, $u(x,t) = \beta(x,t) + O(\epsilon)$, $T(x,t) = (\gamma - 1)\beta(x,t) + O(\epsilon)$. If these expressions for $\rho(x,t)$, $u(x,t)$, and $T(x,t)$ are inserted into the matrices B and C as given in (10.3.44), we obtain $B = \beta(x,t)\hat{B} + O(\epsilon)$, $C = \beta(x,t)\hat{C} + O(\epsilon)$, with $\hat{B}$ and $\hat{C}$ given in (10.3.50). Thus on inserting (10.3.51) in (10.3.43), we have, correct to $O(\epsilon^2)$,

$$[\beta_t + \beta_x]\mathbf{r}_1 + \epsilon\{[\delta_t - \delta_x]\mathbf{r}_2 + \alpha_t\mathbf{r}_0 + \beta\beta_x\hat{B}\mathbf{r}_1 + \beta\beta_t\hat{C}\mathbf{r}_1 + \beta_{xx}D\mathbf{r}_1\} = 0. \tag{10.3.54}$$

On multiplying (10.3.54) with the projection operator P_1, we obtain

$$\beta_t + \beta_x + \epsilon\left\{\frac{\gamma^2+3\gamma-1}{2\gamma}\beta\beta_x + \frac{2\gamma-1}{2\gamma}\beta\beta_t - \left[\frac{2}{3}\mu + \frac{1}{2}k(\gamma-1)^2\right]\beta_{xx}\right\} = 0, \tag{10.3.55}$$

since $P_1\mathbf{r}_0 = P_1\mathbf{r}_2 = \mathbf{0}$. Now (10.3.55) implies that $\beta_t = -\beta_x + O(\epsilon)$, so we can replace (10.3.55) by

$$\beta_t + \beta_x + \epsilon\left\{\frac{\gamma+1}{2}\beta\beta_x - \left[\frac{2}{3}\mu + \frac{1}{2}k(\gamma-1)^2\right]\beta_{xx}\right\} = 0. \tag{10.3.56}$$

This is *Burgers' equation* for *right-traveling waves*. It contains both a nonlinear term, $\beta\beta_x$, and a second derivative term, β_{xx}.

Next we multiply (10.3.54) by P_2, and this yields

$$[\delta_t - \delta_x]\mathbf{r}_2 + \alpha_t\mathbf{r}_0 + \beta\beta_x P_2\hat{B}\mathbf{r}_1 + \beta\beta_t P_2\hat{C}\mathbf{r}_1 + \beta_{xx}P_2D\mathbf{r}_1 = 0, \tag{10.3.57}$$

since $P_2\mathbf{r}_1 = \mathbf{0}$. We solve (10.3.57) by expanding $\alpha(x,t)$ and $\delta(x,t)$ as $\alpha(x,t) = \sum_{j=0}^{\infty}\alpha_j(x,t)\epsilon^j$, $\delta(x,t) = \sum_{j=0}^{\infty}\delta_j(x,t)\epsilon^j$. It is easiest to solve (10.3.57) by constructing projection operators $\hat{P}_0$ and $\hat{P}_2$ that project vectors into the eigenspaces spanned by $\mathbf{r}_0$ and $\mathbf{r}_2$, respectively. Using this method and the fact that $P_1 + \hat{P}_0 + \hat{P}_2 = I$, we easily establish that

$$\hat{P}_0 = \frac{1}{\gamma}\begin{bmatrix} \gamma-1 & 0 & -1 \\ 0 & 0 & 0 \\ 1-\gamma & 0 & 1 \end{bmatrix}, \quad \hat{P}_2 = \frac{1}{2\gamma}\begin{bmatrix} 1 & -\gamma & 1 \\ -1 & \gamma & -1 \\ \gamma-1 & \gamma(\gamma-1) & \gamma-1 \end{bmatrix}. \tag{10.3.58}$$

Then $\hat{P}_0\mathbf{r}_0 = \mathbf{r}_0$, $\hat{P}_0\mathbf{r}_2 = \mathbf{0}$, $\hat{P}_2\mathbf{r}_2 = \mathbf{r}_2$, and $\hat{P}_2\mathbf{r}_0 = \mathbf{0}$. Since $\beta_t = -\beta_x + O(\epsilon)$, we may express (10.3.57) as

$$[\delta_t - \delta_x]\mathbf{r}_2 + \alpha_t\mathbf{r}_0 + \beta\beta_x[P_2\hat{B} - P_2\hat{C}]\mathbf{r}_1 + \beta_{xx}P_2D\mathbf{r}_1 = O(\epsilon). \tag{10.3.59}$$

Inserting the series into (10.3.59) and multiplying by $\hat{P}_0$ yields for $\alpha_0(x,t)$,

$$\frac{\partial \alpha_0}{\partial t} + (1-\gamma)\beta\beta_x + k(\gamma-1)^2\beta_{xx} = 0. \tag{10.3.60}$$

Multiplying (10.3.59) by $\hat{P}_2$, we obtain for $\delta_0(x,t)$,

$$\frac{\partial \delta_0}{\partial t} - \frac{\partial \delta_0}{\partial x} + \frac{\gamma^2 - 3\gamma + 4}{2\gamma}\beta\beta_x + \left(\frac{2}{3}\mu - \frac{1}{2}k(\gamma-1)^2\right)\beta_{xx} = 0. \tag{10.3.61}$$

We do not consider the equations for α_j and δ_j with $j \geq 1$.

If we set

$$\mathbf{u}(x,t) = \delta(x,t)\mathbf{r}_2 + \epsilon[\alpha(x,t)\mathbf{r}_0 + \beta(x,t)\mathbf{r}_1], \tag{10.3.62}$$

and proceed as before, it may be shown that $\delta(x,t)$ satisfies *Burgers' equation* for *left-traveling waves*. The derivation is carried out in the exercises. Finally, if we set

$$\mathbf{u}(x,t) = \alpha(x,t)\mathbf{r}_0 + \epsilon[\beta(x,t)\mathbf{r}_1 + \delta(x,t)\mathbf{r}_2], \tag{10.3.63}$$

we find that $\alpha(x,t)$ satisfies a *heat conduction equation* to the previously specified level of approximation, as we now show.

With $\mathbf{u}(x,t)$ given by (10.3.63), we have $\rho(x,t) = \alpha(x,t) + O(\epsilon)$, $u(x,t) = O(\epsilon)$, $T(x,t) = (\gamma-1)\alpha(x,t) + O(\epsilon)$, and the matrices B and C become

$$B = \alpha \begin{bmatrix} 0 & 1 & 0 \\ (\gamma-1)/\gamma & 0 & 1/\gamma \\ 0 & \gamma(\gamma-1) & 0 \end{bmatrix} + O(\epsilon) \equiv \alpha\tilde{B} + O(\epsilon),\ C = \alpha\hat{C} + O(\epsilon), \tag{10.3.64}$$

with $\hat{C}$ defined as in (10.3.50). Inserting (10.3.63) into (10.3.43) yields

$$\alpha_t\mathbf{r}_0 + \epsilon\{[\beta_t + \beta_x]\mathbf{r}_1 + (\delta_t - \delta_x)\mathbf{r}_2 + \alpha\alpha_x\tilde{B}\mathbf{r}_0 + \alpha\alpha_t\hat{C}\mathbf{r}_0 + \alpha_{xx}D\mathbf{r}_0\} = O(\epsilon^2). \tag{10.3.65}$$

Multiplying (10.3.65) by the projection operator $\hat{P}_0$ gives

$$\alpha_t(x,t) + \epsilon\left\{\frac{1}{\gamma}\alpha(x,t)\alpha_t(x,t) - k(\gamma-1)\alpha_{xx}(x,t)\right\} = O(\epsilon^2). \tag{10.3.66}$$

Since $\alpha_t(x,t) = O(\epsilon)$, we may replace (10.3.66) by

$$\alpha_t(x,t) - \epsilon k(\gamma-1)\alpha_{xx}(x,t) = O(\epsilon^2). \tag{10.3.67}$$

As $\gamma > 1$, (10.3.67) is an *equation of heat conduction* if we drop the $O(\epsilon^2)$ terms. On using the projection operators P_1 and $\hat{P}_2$ in (10.3.65), we may express $\beta(x,t)$ and $\delta(x,t)$ in terms of $\alpha(x,t)$, but this is not carried out here.

Burgers' Equation: Simple Waves and Shock Waves

We have shown that to $O(\epsilon^2)$, $\beta = \beta(x,t)$, satisfies *Burgers' equation*,

$$\beta_t + \beta_x + \epsilon\left\{\frac{\gamma+1}{2}\beta\beta_x - \left[\frac{2}{3}\mu + \frac{1}{2}k(\gamma-1)^2\right]\beta_{xx}\right\} = 0. \tag{10.3.68}$$

The leading terms $\alpha_0(x,t)$ and $\delta_0(x,t)$ in the expansions of $\alpha(x,t)$ and $\delta(x,t)$ are expressed in terms of $\beta(x,t)$ by means of (10.3.60)–(10.3.61). Thus, the perturbed density, velocity, and temperature as defined in (10.3.42) are $\hat{\rho}(x,t) = \beta(x,t) + \epsilon\alpha_0(x,t) + \epsilon\delta_0(x,t)$, $\hat{u}(x,t) = \beta(x,t) - \epsilon\delta_0(x,t)$, $\hat{T}(x,t) = (\gamma-1)\beta(x,t) - \epsilon\alpha_0(x,t) + \epsilon(\gamma-1)\delta_0(x,t)$.

To compare the present results with those obtained in Section 8.5, we discuss the foregoing approximations as they relate to the velocity $u(x,t) = \epsilon\hat{u}(x,t)$ [see (10.3.42)]. We have $u(x,t) = \epsilon\hat{u}(x,t) = \epsilon\beta(x,t) - \epsilon^2\delta_0(x,t)$, and to leading order we set $u(x,t) = \epsilon\beta(x,t)$. Multiplying (10.3.68) by ϵ gives

$$u_t(x,t) + \left[1 + \frac{\gamma+1}{2}u(x,t)\right]u_x(x,t) - \epsilon\left[\frac{2}{3}\mu + \frac{1}{2}k(\gamma-1)^2\right]u_{xx}(x,t) = 0. \tag{10.3.69}$$

The diffusion term $u_{xx}(x,t)$ in (10.3.69) disappears if $\epsilon = 0$ or if $\mu = k = 0$. Then (10.3.69) takes the form of the quasilinear one-dimensional wave equation (8.5.71) except that $c_0 = 1$ in our case. The equation (8.5.71) was obtained on the basis of Euler's equations of fluid motion, in which viscosity and conductivity are neglected (i.e., we have $\mu = k = 0$). It was also assumed that the gas is polytropic [i.e., (8.5.63) is valid]. Consequently, (10.3.69) is seen to represent a modification of the quasilinear equation (8.5.72) if the effects of viscosity and conductivity are included. The solutions of (8.5.72) and its generalizations were called *simple waves* in Section 8.5, and it was found that they can break down after a finite time. We expect that the presence of a second derivative term in Burgers' equation eliminates this problem. Since both (8.5.72) and (10.3.69) can be solved exactly, we can compare their solutions for identical initial values and determine the role of the diffusion term in smoothing out the solutions. This will be carried out below.

With $\nu = (2/3)\,\mu + (1/2)\,k(\gamma-1)^2$, (10.3.69) becomes

$$u_t(x,t) + \left[1 + \frac{\gamma+1}{2}u(x,t)\right]u_x(x,t) - \epsilon\nu u_{xx}(x,t) = 0. \tag{10.3.70}$$

On equating ϵ or ν to zero in (10.3.70), the solution of the initial value problem for the resulting first order equation with the initial condition

$$u(x,0) = h(x), \tag{10.3.71}$$

is given implicitly as

$$u = h\left[x - \left(1 + \frac{\gamma+1}{2}u\right)t\right]. \tag{10.3.72}$$

Cole and *Hopf* observed that with $1+[(\gamma+1)/2]u(x,t) = -2\epsilon\nu[v_x(x,t)/v(x,t)]$ in (10.3.70), $v(x,t)$ satisfies the *heat conduction equation*

$$v_t(x,t) = \epsilon\nu v_{xx}(x,t), \tag{10.3.73}$$

as is readily verified. The initial condition (10.3.71) is transformed into

$$v(x,0) = \exp\left\{-\frac{1}{2\epsilon\nu}\int_0^x \left[1+\frac{\gamma+1}{2}\,h(s)\right]\,ds\right\} \equiv H(x), \tag{10.3.74}$$

and the solution of the initial value problem (10.3.73)–(10.3.74) is given as

$$v(x,t) = \frac{1}{\sqrt{4\pi\epsilon\nu t}}\int_{-\infty}^{\infty} H(\xi)\exp\left[-\frac{(x-\xi)^2}{4\epsilon\nu t}\right]\,d\xi. \tag{10.3.75}$$

The solution $u(x,t)$ of (10.3.70)–(10.3.71) can then be obtained from

$$u(x,t) = \frac{2}{\gamma+1}\left\{\frac{\int_{-\infty}^{\infty}[(x-\xi)/t]\exp[-g(\xi,t)/2\epsilon\nu]\,d\xi}{\int_{-\infty}^{\infty}\exp[-g(\xi,t)/2\epsilon\nu]\,d\xi} - 1\right\}, \tag{10.3.76}$$

where

$$g(\xi,t) = \int_0^{\xi}\left[1+\frac{\gamma+1}{2}\,h(s)\right]\,ds + \frac{(x-\xi)^2}{2t}. \tag{10.3.77}$$

We examine the solution (10.3.76) for the special *initial condition*

$$u(x,0) = h(x) = \begin{cases} h_1, & x>0, \\ h_2, & x<0, \end{cases} \tag{10.3.78}$$

where $0 < h_1 < h_2$ and h_1, h_2 are constants. Then $g(\xi,t)$ can be expressed in terms of the functions $g_i(\xi,t)$, defined as

$$g_i(\xi,t) = a_i\xi + \frac{(x-\xi)^2}{2t} = \frac{(\xi - x + a_i t)^2}{2t} + a_i\left(x - \frac{a_1+a_2}{2}\,t\right) + \frac{a_1a_2}{2}\,t, \tag{10.3.79}$$

where $a_i = 1+[(\gamma+1)/2]h_i$ $(i=1,2)$, in the form $g(\xi,t) = \begin{cases} g_1(\xi,t), & \xi>0, \\ g_2(\xi,t), & \xi<0. \end{cases}$

We recall the *fundamental solution* $G(x,t)$ of the heat equation $G(x,t) = \exp\left(-x^2/4t\right)/\sqrt{4\pi t}$. Then we can express (10.3.76) as

$$1+\frac{\gamma+1}{2}\,u(x,t) = \frac{\exp[\hat{\lambda}(x-\tilde{\lambda}t)/\epsilon\nu]\int_{-\infty}^{0}[(x-\xi)/t]G_2(\xi)\,d\xi + \int_0^{\infty}[(x-\xi)/t]G_1(\xi)\,d\xi}{\exp[\hat{\lambda}(x-\tilde{\lambda}t)/\epsilon\nu]\int_{-\infty}^{0}G_2(\xi)\,d\xi + \int_0^{\infty}G_1(\xi)\,d\xi}, \tag{10.3.80}$$

where $\tilde{\lambda} = (a_1+a_2)/2$, $\hat{\lambda} = (a_1-a_2)/2$ and $G_i(\xi) \equiv G(x-a_it-\xi, \epsilon\nu t)$ $(i = 1, 2)$, as is easily verified.

Now as $\epsilon\nu \to 0$, we have $\lim_{\epsilon\nu\to 0} G_i(\xi) = \lim_{\epsilon\nu\to 0} G(x - a_it - \xi, \epsilon\nu t) = \delta(x - a_it - \xi)$, as follows from the property of the fundamental solution $G(x, t)$ established previously. Also, since $a_1 < a_2$, the exponentials in (10.3.80) tend to infinity as $\epsilon\nu \to 0$ if $x - \tilde{\lambda}t < 0$, and to zero if $x - \tilde{\lambda}t > 0$. Noting these results, we readily find that

$$\lim_{\epsilon\nu\to 0} u(x,t) = \begin{cases} h_1, & x - \left[1 + \frac{\gamma+1}{4}(h_1+h_2)\right] t > 0, \\ h_2, & x - \left[1 + \frac{\gamma+1}{4}(h_1+h_2)\right] t < 0, \end{cases} \tag{10.3.81}$$

since

$$\tilde{\lambda} = \frac{a_1+a_2}{2} = 1 + \frac{\gamma+1}{4}(h_1+h_2). \tag{10.3.82}$$

We expect that (10.3.81) satisfies the reduced equation for (10.3.70),

$$u_t(x,t) + \left[1 + \left(\frac{\gamma+1}{2}\right) u(x,t)\right] u_x(x,t) = 0, \tag{10.3.83}$$

with the initial condition (10.3.78). [Since (10.3.83) results on setting the parameter $\nu = 0$ in *Burgers' equation* (10.3.70), and $\nu = 0$ if viscosity and heat conduction effects are neglected, it is referred to as the *inviscid Burgers' equation*.] However, the implicit solution (10.3.72) of (10.3.83) shows that u is a constant on the lines $x - [1 + (\gamma+1)u/2]t = \xi$ = constant. Also, at $t = 0$ we have $u = h_2$ if $x < 0$ and $u = h_1$ if $x > 0$, so that

$$u(x,t) = \begin{cases} h_1, & x - a_1t > 0, \\ h_2, & x - a_2t < 0. \end{cases} \tag{10.3.84}$$

Since $a_1 < a_2$, these regions in the (x, t)-plane are given as shown in Figure 10.8. In the sector between $x = a_1t$ and $x = a_2t$, the solution (10.3.84) is multivalued, and since this multivaluedness occurs as soon as t exceeds zero, this solution is not valid for any time $t > 0$.

However, (10.3.81) yields a (discontinuous) solution of (10.3.83) that satisfies the initial condition (10.3.78) and has a jump across the line $x = \tilde{\lambda}t$ that lies between the lines $x = a_1t$ and $x = a_2t$ as pictured in Figure 10.8. The jump in $u(x, t)$ travels to the right with the speed $\tilde{\lambda}$, and this does not correspond to either of the two characteristic speeds a_1 and a_2 for this problem. In fact, the traveling wave solution (10.3.81) of the quasilinear equation (10.3.83) is a *shock wave* and it has been obtained as the limit of the smooth solution (10.3.80) of Burgers' equation as $\epsilon\nu \to 0$.

Burgers' Equation: Shock Structure

In Section 2.3 we constructed *shock waves* for first order quasilinear equations directly, without having to solve a higher-order equation and evaluating a limit, as we have

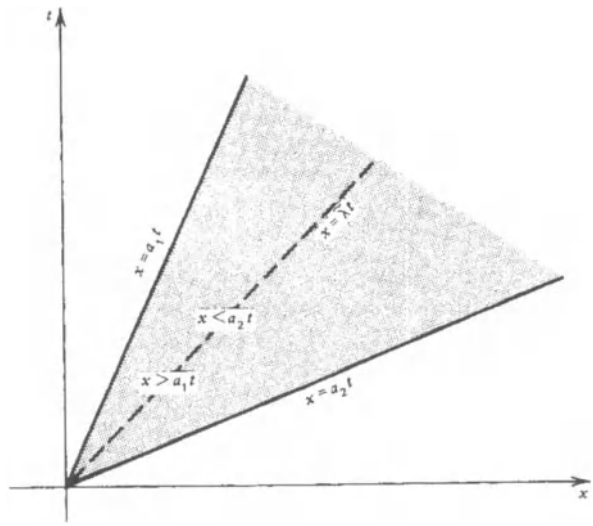

Figure 10.8 The multivalued region and shock line.

done here. [The direct construction of the shock wave (10.3.81) is considered in the exercises.] We recognize, however, that the shock wave is an idealized and simplified form of the solution of the given problem for Burgers' equation, in which $\epsilon\nu$ is small but nonzero. Consequently, the solution of Burgers' equation for small but nonzero $\epsilon\nu$ is valuable in that it yields a solution of the problem that varies rapidly near the shock front but is nevertheless smooth. This more detailed description of the solution near the shock gives what is called the *shock structure*.

There is an alternative approach to determining the shock wave solution (10.3.81) of (10.3.83), which is also useful for more general (quasilinear) equations that do not have an exact solution as is the case for Burgers' equations. We look for a *traveling wave* solution of the partial differential equation (10.3.70) in the form

$$u(x,t) = v(x - \lambda t), \tag{10.3.85}$$

where the function $v(x - \lambda t)$ and the constant λ are to be specified. They are determined by the requirement that

$$v(x - \lambda t) \to \begin{cases} h_1, & x - \lambda t \to \infty, \\ h_2, & x - \lambda t \to -\infty, \end{cases} \tag{10.3.86}$$

where the constants h_1 and h_2 have the values given in the initial condition (10.3.78).

Inserting (10.3.85) into (10.3.70) yields a second order ordinary differential equation for $v(x - \lambda t)$,

$$(1 - \lambda)v' + \frac{\gamma + 1}{2} vv' = \epsilon\nu v''. \tag{10.3.87}$$

Integrating once in (10.3.87) and noting that $v' \to 0$ as $|x - \lambda t| \to \infty$, in view of (10.3.86), we have

$$(1 - \lambda)v + \frac{\gamma + 1}{4} v^2 + v_0 = \epsilon\nu v', \tag{10.3.88}$$

where v_0 is an arbitrary constant of integration. On letting $x - \lambda t \to \infty$, we obtain

$$(1 - \lambda)h_1 + \frac{\gamma + 1}{4} h_1^2 + v_0 = 0, \tag{10.3.89}$$

and as $x - \lambda t \to -\infty$, we have

$$(1 - \lambda)h_2 + \frac{\gamma + 1}{4} h_2^2 + v_0 = 0. \tag{10.3.90}$$

On solving the simultaneous system (10.3.89)–(10.3.90) for λ and v_0, we obtain

$$\lambda = 1 + \frac{\gamma + 1}{4} (h_1 + h_2) = \tilde{\lambda} \tag{10.3.91}$$

for the value of λ as given in (10.3.82), and the value of the constant v_0 is determined to be

$$v_0 = \frac{\gamma + 1}{4} h_1 h_2. \tag{10.3.92}$$

On inserting these values into (10.3.88), we find that the equation can be written as

$$v' = \frac{\gamma + 1}{4\epsilon\nu} (v - h_1)(v - h_2). \tag{10.3.93}$$

This first order ODE is easily solved using a partial fraction decomposition. We find that

$$u(x,t) = v(x - \tilde{\lambda}t) = h_1 + \frac{h_1 - h_2}{1 + \exp\{[(\gamma + 1)/4\epsilon\nu](h_2 - h_1)(x - \tilde{\lambda}t)\}}, \tag{10.3.94}$$

and this satisfies (10.3.86). Furthermore, as $\epsilon\nu \to 0$, since $h_2 > h_1$, (10.3.94) tends to (10.3.81), as is readily checked. Thus for small but nonzero $\epsilon\nu$, this solution yields the *shock structure* and describes how the initial discontinuity (10.3.78) propagates as a smooth step whose transition from the value h_2 to the value h_1 moves at the shock speed $\tilde{\lambda}$.

We have assumed that $0 < h_1 < h_2$ in the initial condition (10.3.78). This gives rise to a *compression wave* solution of the inviscid Burgers' equation (10.3.83), which breaks immediately and requires the introduction of a *shock wave*. However, if $0 < h_2 < h_1$ in (10.3.78), we have an *expansion wave* for which continuous solutions can be found as shown in Section 2.3. Now if we look for a *traveling wave* solution of the viscous Burgers' equation with $h_2 < h_1$ in (10.3.86), we see from (10.3.93) that no traveling wave solution exists that increases smoothly from h_2 at $-\infty$ to h_1 at $+\infty$, since for $h_2 < v < h_1$ we have $v' < 0$. Thus, v must be a decreasing rather than an increasing function, and this is not possible. As a result, we find that

expansion shocks do not result as limits of traveling wave solutions of the viscous Burgers' equation.

More general initial data for (10.3.70) and the behavior of the corresponding solution (10.3.76) as $\epsilon\nu \to 0$ may be discussed by evaluating the integrals in (10.3.76) asymptotically using the method of steepest descents, the saddle point method. However, we do not pursue this matter.

Shallow Water Theory: Boussinesq Equations

It can be shown that the *theory of water waves* for the case of *shallow water* and waves of *small amplitude* can be (approximately) described by the *Boussinesq equations*

$$v_t(x,t) + w_x(x,t) + \epsilon[v(x,t)w_x(x,t) + w(x,t)v_x(x,t)] = 0, \tag{10.3.95}$$

$$w_t(x,t) + v_x(x,t) + \epsilon[w(x,t)w_x(x,t) - \frac{1}{3}w_{xxt}(x,t)] = 0. \tag{10.3.96}$$

If we set $gh(x,t) = 1 + \epsilon v(x,t)$ and $u(x,t) = \epsilon w(x,t)$, where g is the gravitational constant, $h(x,t)$ represents the height of the water above a horizontal bottom and $u(x,t)$ is the velocity of the water, at a point x and the time t. (The problem is one-dimensional.) If we drop the third derivative term in (10.3.96) and express the resulting first order system in terms of $h(x,t)$ and $u(x,t)$, we obtain the system of first order equations of Exercise 3.3.21. They are known as the equations of *shallow water theory*. The small parameter ϵ in (10.3.95)–(10.3.96) signifies that we are considering small wave amplitudes and shallow water, or, equivalently, long waves.

With

$$\mathbf{u} = \begin{bmatrix} v(x,t) \\ w(x,t) \end{bmatrix}, \quad A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} w(x,t) & v(x,t) \\ 0 & w(x,t) \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 0 \\ 0 & -\frac{1}{3} \end{bmatrix}, \tag{10.3.97}$$

the system (10.3.95)–(10.3.96) can be written as

$$\mathbf{u}_t(x,t) + A\mathbf{u}_x(x,t) + \epsilon[B\mathbf{u}_x(x,t) + C\mathbf{u}_{xxt}(x,t)] = \mathbf{0}. \tag{10.3.98}$$

The system (10.3.98) has the form (10.3.30). Putting $\epsilon = 0$ in (10.3.98) yields the linear *reduced system* $\mathbf{u}_t(x,t) + A\mathbf{u}_x(x,t) = \mathbf{0}$, equivalent to the one-dimensional *wave equation* with unit wave speed. In fact, the characteristic polynomial for A is $C(\lambda) = |A - \lambda I| = (\lambda - 1)(\lambda + 1)$. With $\lambda_1 = 1$ and $\lambda_2 = -1$ as the eigenvalues of A, the corresponding eigenvectors are

$$\mathbf{r}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad \mathbf{r}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}. \tag{10.3.99}$$

When $\mathbf{u}(x,t) = \alpha(x,t)\mathbf{r}_1$, we have $\mathbf{u}_t(x,t) + A\mathbf{u}_x(x,t) = (\alpha_t(x,t) + \alpha_x(x,t))\mathbf{r}_1 = \mathbf{0}$, and when $\mathbf{u}(x,t) = \beta(x,t)\mathbf{r}_2$, we have $\mathbf{u}_t(x,t) + A\mathbf{u}_x(x,t) = (\beta_t(x,t) - \beta_x(x,t))\mathbf{r}_2 = \mathbf{0}$. Thus, $\alpha(x,t) = \alpha(x-t)$ and $\beta(x,t) = \beta(x+t)$.

Asymptotic Simplification of the Shallow Water Equations: Korteweg-deVries Equation

We now carry out the asymptotic simplification of (10.3.98). Since A has only two (orthogonal) eigenvectors, the simplification method does not require the full machinery developed in our general discussion. The reduced system was shown to give rise to a right- and a left-traveling wave. To study the effect of the nonlinear and higher derivative terms on the *right-traveling wave*, we set

$$\mathbf{u}(x,t) = \alpha(x,t)\mathbf{r}_1 + \epsilon\beta(x,t)\mathbf{r}_2, \tag{10.3.100}$$

where $\alpha(x,t)$ and $\beta(x,t)$ are to specified. We insert (10.3.100) into (10.3.98) and obtain

$$[\alpha_t + \alpha_x]\mathbf{r}_1 + \epsilon\left\{[\beta_t - \beta_x]\mathbf{r}_2 + \alpha\alpha_x\begin{bmatrix} 2 \\ 1 \end{bmatrix} - \frac{1}{3}\alpha_{xxt}\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\} = O(\epsilon^2), \tag{10.3.101}$$

since $v(x,t) = \alpha(x,t) + O(\epsilon)$ and $w(x,t) = \alpha(x,t) + O(\epsilon)$. The vectors $\mathbf{r}_1$ and $\mathbf{r}_2$ are orthogonal, so we need not construct projection matrices corresponding to their eigenspaces. Instead, we simply take dot products with $\mathbf{r}_1$ and then with $\mathbf{r}_2$ in (10.3.101).

Dotting with $\mathbf{r}_1$ in (10.3.101) gives

$$\alpha_t + \alpha_x + \frac{3}{2}\epsilon\alpha\alpha_x - \frac{1}{6}\epsilon\alpha_{xxt} = O(\epsilon^2). \tag{10.3.102}$$

Dotting with $\mathbf{r}_2$ gives

$$\beta_t - \beta_x + \frac{1}{2}\alpha\alpha_x + \frac{1}{6}\alpha_{xxt} = O(\epsilon). \tag{10.3.103}$$

We solve (10.3.103) by expanding $\beta(x,t)$ in a series in powers of ϵ with $\beta(x,t) = \beta_0(x,t) + O(\epsilon)$. Thus

$$\frac{\partial\beta_0(x,t)}{\partial t} - \frac{\partial\beta_0(x,t)}{\partial x} + \frac{1}{2}\alpha(x,t)\alpha_x(x,t) + \frac{1}{6}\alpha_{xxt}(x,t) = 0. \tag{10.3.104}$$

Since $\alpha_t(x,t) = -\alpha_x(x,t) + O(\epsilon)$ in view of (10.3.102), we may replace that equation by

$$\alpha_t(x,t) + \alpha_x(x,t) + \frac{3}{2}\epsilon\alpha(x,t)\alpha_x(x,t) + \frac{1}{6}\epsilon\alpha_{xxx}(x,t) = 0, \tag{10.3.105}$$

on dropping the $O(\epsilon^2)$ terms. This is the *Korteweg-deVries equation.* As was indicated in Example 9.7, it is a *nonlinear dispersive equation.* In this regard it differs from the Burgers' equation, which is of dissipative type. However, the Korteweg-deVries equation is also exactly solvable; but the solution method is substantially more complicated than that for Burgers' equation, and it is not presented here.

We note that if (10.3.105) is solved by expanding $\alpha(x,t)$ in powers of ϵ (i.e., by using the perturbation method), secular terms result in that expansion. Thus we leave (10.3.105) as it stands, and in terms of its solution we obtain $v(x,t) = \alpha(x,t) + \epsilon\beta_0(x,t)$, $w(x,t) = \alpha(x,t) - \epsilon\beta_0(x,t)$, in view of (10.3.100) and (10.3.104), where $\beta_0(x,t)$ is solved for in terms of $\alpha(x,t)$. If we replace (10.3.100) by $\mathbf{u}(x,t) = \beta(x,t)\mathbf{r}_2 + \epsilon\alpha(x,t)\mathbf{r}_1$, we find on proceeding as earlier that $\beta(x,t)$ satisfies a Korteweg-deVries equation for left-traveling waves, as shown in the exercises.

Solitary Wave Solution of the Korteweg-deVries Equation

We have already constructed approximate periodic solutions of the Korteweg-deVries equation in Example 9.7. We now obtain a special traveling wave solution of (10.3.105) that represents what is known as a *solitary wave*. This wave consists of a single hump of constant shape that moves at constant speed.

Let

$$\alpha(x,t) = af(x - \lambda t), \tag{10.3.106}$$

where the constants a and λ and the function f are to be specified. We assume that $f(x-\lambda t)$ vanishes together with its derivatives as $|x-\lambda t| \to \infty$. Inserting (10.3.106) into (10.3.105) gives the third order ODE $(1-\lambda)f' + (3/2)\epsilon a f f' + (1/6)\epsilon f''' = 0$. Integrating once gives $(1-\lambda)f + (3/4)\epsilon a f^2 + (1/6)\epsilon f'' = 0$, with no constant of integration since f and f'' vanish at infinity. Next we multiply the equation by f' and integrate once more to obtain $((1-\lambda)/2)\, f^2 + (1/4)\epsilon a f^3 + (1/12)\epsilon f'^2 = 0$. Again, there is no constant of integration. This equation can be written as $(1/3a)f'^2 = f^2\,[2(\lambda-1)/\epsilon a - f]$.

The solution of this equation that vanishes at infinity is readily verified to be

$$f(x-\lambda t) = \frac{2(\lambda-1)}{\epsilon a}\,\operatorname{sech}^2\left\{\left[\frac{3(\lambda-1)}{2\epsilon}\right]^{1/2}(x-\lambda t)\right\}. \tag{10.3.107}$$

If we set $\hat{a} = 2(\lambda-1)/\epsilon$, the solution $\alpha(x,t) = af(x-\lambda t)$ can be written as

$$\alpha(x,t) = \hat{a}\,\operatorname{sech}^2\left\{\left(\frac{3\hat{a}}{4}\right)^{1/2}\left[x - \left(1 + \frac{\epsilon\hat{a}}{2}\right)t\right]\right\}, \tag{10.3.108}$$

with the solitary wave speed given as $\lambda = 1 + \epsilon\hat{a}/2$.

At the time $t = 0$, the solitary wave $\alpha(x,t)$ has the form of a single hump. The maximum value of $\alpha(x,0)$ occurs at $x = 0$ and equals $\hat{a}$. The waveform $\alpha(x,0)$ is symmetric with respect to $x = 0$. This form travels to the right without change in shape at the speed $\lambda = 1 + \epsilon\hat{a}/2$. Thus as $\hat{a}$ increases, so does the wave speed, and higher solitary waves move more rapidly than lower ones. This leads to some interesting problems regarding the interaction of two or more solitary waves moving at different speeds. Also, by modifying the foregoing discussion, it is possible to obtain exact periodic traveling wave solutions of the Korteweg-deVries equation. We do not discuss this here.

The Burgers, Korteweg-deVries, and nonlinear Schrödinger equations that were discussed in this and the preceding sections arise as the relevant approximate equations for a number of important problems of physical interest. Each of these equations can be solved exactly, and the study of the solution of these and related nonlinear PDEs has been the subject of intense investigation in recent years.

Exercises 10.3

10.3.1. Consider the equation $\epsilon(u_{ttt}(x,t) - a^2 u_{xxt}(x,t)) + u_{tt}(x,t) - c^2 u_{xx}(x,t) = 0$, where $a > c$. Let $v_1(x,t) = u_{tt}(x,t)$, $v_2(x,t) = u_{xt}(x,t)$, and $v_3(x,t) = u_{xx}(x,t)$ and express the equation for $u(x,t)$ as the system $\epsilon[\mathbf{u}_t(x,t) + A\mathbf{u}_x(x,t)] + B\mathbf{u}(x,t) = \mathbf{0}$, where $\mathbf{u}(x,t)$ has the components $v_1(x,t)$, $v_2(x,t)$, and $v_3(x,t)$. Construct a decomposition of this system based on two linearly independent null eigenvectors of the matrix B (i.e., the eigenvectors that correspond to the eigenvalue $\lambda = 0$). *Hint*: Each of these equations is a parabolic equation.

10.3.2. Express the equation $u_{tt}(x,t) - c^2 u_{xx}(x,t) + \epsilon u_{xxt}(x,t)) = 0$, as a system by setting $v(x,t) = u_t(x,t)$ and $w(x,t) = u_x(x,t)$ and obtain $\mathbf{u}_t(x,t) + A\mathbf{u}_x(x,t) + \epsilon B\mathbf{u}_{xx}(x,t) = \mathbf{0}$, where $\mathbf{u}(x,t)$ has the components $v(x,t)$ and $w(x,t)$. Obtain a decomposition of the system based on the eigenvalues and eigenvectors of the matrix A.

10.3.3. Obtain a Burgers' equation for left-traveling waves by using (10.3.62) and proceeding as in the text.

10.3.4. Determine the shock wave (10.3.81) using the method of Section 2.3.

10.3.5. Obtain a Korteweg-deVries equation for left-traveling waves by setting $\mathbf{u}(x,t) = \alpha(x,t)\mathbf{r}_2 + \epsilon\beta(x,t)\mathbf{r}_1$ in the system (10.3.98).

10.3.6. Show that (10.3.94) is the solution of (10.3.87) with λ and v_0 given as in (10.3.91)–(10.3.92) and verify that it tends to (10.3.81) as $\epsilon\nu \to 0$.

10.3.7. The nonlinear hyperbolic system $V_\tau(\sigma,\tau) + V_\sigma(\sigma,\tau) = W^2(\sigma,\tau) - V^2(\sigma,\tau)$, $W_\tau(\sigma,\tau) - W_\sigma(\sigma,\tau) = V^2(\sigma,\tau) - W^2(\sigma,\tau)$ represents a model for *Boltzmann's equation* in the *kinetic theory of gases*. Let $V(\sigma,\tau) = 1 + \epsilon v(x,t)$, $W(\sigma,\tau) = 1 + \epsilon w(x,t)$, $t = \epsilon\tau$, $x = \epsilon\sigma$, and obtain the system $C\mathbf{u}(x,t) + \epsilon[\mathbf{u}_t(x,t) + A\mathbf{u}_x(x,t) + B\mathbf{u}(x,t)] = \mathbf{0}$, where $\mathbf{u}(x,t)$ has the components $v(x,t)$ and $w(x,t)$, and the matrices A, B, and C are defined appropriately. Show that the matrix C has a null vector $\mathbf{r}^T = [1,1]$. Carry out a decomposition of the system corresponding to this vector and obtain a nonlinear parabolic equation as the approximating equation.

10.3.8. The dispersive equation $u_{tt}(x,t) - u_{xx}(x,t) + \epsilon[u_x(x,t)u_{xx}(x,t) - u_{xxxx}(x,t)] = 0$ models the *vibration of a nonlinear string*. Let $v(x,t) = u_t(x,t)$ and $w(x,t) = u_x(x,t)$, and express the equation in the form of a nonlinear system $\mathbf{u}_t(x,t) + A\mathbf{u}_x(x,t) + \epsilon[B\mathbf{u}_x(x,t) + C\mathbf{u}_{xxx}(x,t)] = \mathbf{0}$, where $\mathbf{u}^T(x,t) = [v(x,t), w(x,t)]$ and the matrices A, B, and C are defined appropriately. Show that the matrix A has eigenvalues $\lambda = +1$ and $\lambda = -1$ with orthogonal eigenvectors. Construct a decomposition of the system corresponding to the eigenvalue $\lambda = 1$ and obtain a *Korteweg-deVries equation* as an approximation.

10.4 MAPLE METHODS

The procedures *RegPerturb*, *BoundLayer*, and *ParabolicEqMeth* introduced in Chapter 9 to analyze regular and singular perturbation problems can also be used in dealing with the problems of Section 10.1. A number of new procedures relevant to the material in Sections 10.2 and 10.3 are presented below.

Equations with a Large Parameter

We consider the reduced wave equation $\nabla^2 u(\mathbf{x}) + k^2 n^2(\mathbf{x}) u(\mathbf{x}) = 0$, in the region G, with $k >> 1$. If the solution $u(\mathbf{x})$ is expressed as $u(\mathbf{x}) = v(\mathbf{x})\, e^{ik\phi(\mathbf{x})}$, and the *phase term* $\phi(\mathbf{x})$ satisfies the eiconal equation $(\nabla\phi(\mathbf{x}))^2 = n^2(\mathbf{x})$, then the *amplitude term* $v(\mathbf{x})$ is a solution of $ik[2\nabla\phi(\mathbf{x}) \cdot \nabla v(\mathbf{x}) + v(\mathbf{x})\nabla^2\phi(\mathbf{x})] + \nabla^2 v(\mathbf{x}) = 0$. The transformation $ik = 1/\epsilon$ converts this PDE into $AmplEq = 2\nabla\phi(\mathbf{x}) \cdot \nabla v(\mathbf{x}) + v(\mathbf{x})\nabla^2\phi(\mathbf{x}) + \epsilon\, \nabla^2 v(\mathbf{x}) = 0$. If $v(\mathbf{x}) = r(\mathbf{x})$ on the boundary S of G, the procedure $RegPerturb(AmplEq, [\,], [v = r(\mathbf{x})], v, [\mathbf{x}], [\mathbf{x}|_S], [\,], \epsilon, 1)$ yields the transport equations (10.1.12)–(10.1.13) for the first two terms $v_0(\mathbf{x})$ and $v_1(\mathbf{x})$ in the perturbation series $v(\mathbf{x}) = \sum_{j=0}^{\infty} v_j(\mathbf{x})\epsilon^j$, where $\epsilon = 1/ik$, with the boundary conditions $v_0(\mathbf{x})|_S = r(\mathbf{x})$ and $v_1(\mathbf{x})|_S = 0$.

The problems of Section 10.1 that require the introduction of boundary layers for their solution can be treated by the use the *BoundLayer* procedure. If the parabolic equation method can be applied, the *ParabolicEqMeth* procedure can be used. We do not consider any examples here.

Propagation of Discontinuities and Singularities for Hyperbolic Equations

We begin by considering representations of functions with jump discontinuities as given in Example 7.2. The procedure *DiscFunct* expresses piecewise smooth functions with single jumps in the terms of expressions and series that involve Heaviside and piecewise-defined functions. On invoking $DiscFunct([f(x), g(x)], x, \alpha, n)$, we obtain a representation of the discontinuous function $F(x) = \begin{cases} f(x), & x < \alpha, \\ g(x) & \alpha < x, \end{cases}$ with a jump at $x = \alpha$, in the form of the partial sum (if $n \geq 1$), $F(x) = f(x) + \sum_{j=1}^{n} [g^{(j)}(\alpha) - f^{(j)}(\alpha)] H_j(x - \alpha)$. The $H_j(x - \alpha)$ are defined as in (7.2.41) and the sum corresponds to (7.2.42). If $n = 0$, the output is $F(x) = f(x) + (g(x) - f(x))H(x - \alpha)$, which represents the piecewise-defined function $F(x)$ in terms of the Heaviside function. As a specific example we consider the procedure $DiscFunct([x, x^2], x, \alpha, 2)$. The discontinuous function is $F(x) = \begin{cases} x, & x < \alpha, \\ x^2 & \alpha < x, \end{cases}$ as defined in (7.2.43). The output is $F(x) = x + (\alpha^2 - \alpha)H(x - \alpha) + (2\alpha - 1)(x - \alpha)H(x - \alpha) + (x - \alpha)^2 H(x - \alpha)$, which agrees with (7.2.44).

$DiscFunctMult([f_1(x), f_2(x), \ldots, f_n(x)], x, [\alpha_1, \alpha_2, \ldots, \alpha_{n-1}])$ determines a representation of the piecewise continuous function $F(x)$, defined as $F(x) = f_1(x)$

for $x < \alpha_1$, $F(x) = f_j(x)$ for $\alpha_{j-1} < x < \alpha_j$, $j = 2, 3, \ldots, n-1$ and $F(x) = f_n(x)$ for $x > \alpha_{n-1}$, in terms of Heaviside functions.

Next we consider the $PDE = u_{tt}(\mathbf{x}, t) - \nabla \cdot (p(\mathbf{x})\nabla u(\mathbf{x}, t)) + 2\lambda u_t(\mathbf{x}, t) + q(\mathbf{x})u(\mathbf{x}, t) = 0$, and the singularity expansion $u = \sum_{j=0}^{\infty} v_j(\mathbf{x}, t) S_j[\phi(\mathbf{x}, t)]$, with $S_j'[\phi] = S_{j-1}[\phi]$. $\phi(\mathbf{x}, t)$ and $v_j(\mathbf{x}, t)$ satisfy the characteristic and transport equations (10.2.37) and (10.2.38), respectively. The Maple procedure $SingularExp(PDE, depvar, [indvar], S_0[\phi], \phi, n)$ exhibits $n+1$ terms in the singularity expansion. The dependent variable *depvar* and a list of independent variables [*indvar*] must be prescribed. If an arbitrary $\phi(\mathbf{x}, t)$ is given, the characteristic equation for the PDE is exhibited. If a solution $\phi(\mathbf{x}, t)$ of the characteristic equation is specified, the output is $0 = 0$. The transport equations for the first $n+1$ terms of the singularity expansion are also displayed.

As an example we consider the *telegrapher's equation* $TE = u_{tt}(x, t) - \gamma^2 u_{xx}(x, t) + 2\lambda u_t(x, t) = 0$. Then $SingularExp(TE, u, [x, t], H[\phi(x, t)], \phi(x, t), 1)$ yields the expansion $u(x, t) \approx v_0(x, t)H[\phi(x, t)] + v_1(x, t)H[\phi(x, t)]\phi(x, t)$. The phase term $\phi(x, t)$ must satisfy the characteristic equation and the terms $v_0(x, t)$ and $v_1(x, t)$ are solutions of the transport equations. If we set $\phi(x, t) = x \pm \gamma t \pm 1$ in the procedure, $\phi(x, t)$ is a solution of the characteristic equation, and the appropriate transport equations are given.

Asymptotic Simplification of Equations

To carry out the asymptotic simplification of the system of PDEs $\mathbf{u}_t(x, t) + A\mathbf{u}_x(x, t) + \epsilon N[\mathbf{u}(x, t)] = \mathbf{0}$, as given in (10.3.30), we must determine projection matrices associated with the eigenspaces of the matrix A. The procedure $ProjMatrix$ determines these projection matrices, but it requires that the $n \times n$ matrix A has n linearly independent eigenvalues and that one of the eigenspaces projected into be one-dimensional. (This is the case for the problems considered in Section 10.3.) In $ProjMatrix(A)$, the matrix A must be expressed (for technical reasons) in the form used in the *linalg* package rather than the *LinearAlgebra* package. That is, it must be expressed as a *matrix* or *array* rather than as a *Matrix* or *Array*. The output exhibits the matrix A, its characteristic polynomial, and a list of its eigenvalues and eigenvectors. The global variable $EVlist$ represents this list. The list may contain simple or multiple eigenvalues, but the procedure can only construct projection matrices if one of the eigenspaces projected into corresponds to a simple eigenvalue. If the simple eigenvalue is the nth element in $EVlist$, the procedure $ProjMatrix(A, EVlist, n)$ determines two projection matrices P_1 and P_2, where P_1 projects into the eigenspace associated with the simple eigenvalue and P_2 projects into the complementary space. (The projection matrices are global variables.)

As an example, we apply the procedure to $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, as given in (10.3.97) in connection with the equations of *shallow water theory*. Then $ProjMatrix(A)$ gives the characteristic polynomial $\lambda^2 - 1$ and the list of eigenvalues and eigenvectors of A as $EVlist = [[1, 1, \{[1, 1]\}], [-1, 1, \{[1, -1]\}]]$. This states that each eigenvalue $\lambda_1 = 1$ and $\lambda_2 = -1$ has multiplicity 1 and gives the eigenvectors $\mathbf{r}_1$ and $\mathbf{r}_2$ as in

(10.3.99). If we express the solution $\mathbf{u}(x,t)$ of the system (10.3.98) as $\mathbf{u}(x,t) = \alpha(x,t)\mathbf{r_1} + \epsilon\beta(x,t)\mathbf{r_2}$, we must project into the eigenspace spanned by $\mathbf{r_1}$. We use $ProjMatrix(A, EVlist, 1)$ since the first element of $EVlist$ corresponds to λ_1 and $\mathbf{r_1}$. The output exhibits λ_1 and $\mathbf{r_1}$ and finds that $P_1 = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}$, $P_2 = \begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{bmatrix}$.

In the general case we put $\mathbf{u}(x,t) = \mathbf{v}(x,t) + \epsilon\mathbf{w}(x,t)$, as in (10.3.31). $\mathbf{v}(x,t)$ is an element of the eigenspace spanned by the eigenvector $\mathbf{r_1}$ associated with a simple eigenvalue λ_1 of A whose projection matrix is P_1. $\mathbf{w}(x,t)$ is an arbitrary vector in the complementary space spanned by the eigenvectors $\mathbf{r}_2, \ldots, \mathbf{r}_n$, whose projection matrix is P_2. Once the projection matrices have been found, we find the simplified systems associated with the vectors $\mathbf{v}(x,t)$ and $\mathbf{w}(x,t)$ as in Section 10.3. $AsymptSimp(System, [depvar], [indvar], \epsilon, \mathbf{r_1}, P_1, [\mathbf{r}_2, \ldots, \mathbf{r}_n], P_2, n)$ determines the simplified equations (10.3.37) and (10.3.34) with $\mathbf{w}$ expanded as in (10.3.35). The number of terms in the expansion is prescribed by the last argument n in the procedure.

Next, we use $AsymptSimp$ to determine the PDE's satisfied by $\alpha(x,t)$ and $\beta(x,t)$ in the shallow water equations. We express them as $S = \partial\mathbf{u}/\partial t + \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \partial\mathbf{u}/\partial x + \epsilon\left(\begin{bmatrix} W & V \\ 0 & W \end{bmatrix}\partial\mathbf{u}/\partial x + \begin{bmatrix} 0 & 0 \\ 0 & -1/3 \end{bmatrix}\partial^3\mathbf{u}/\partial x^2\,\partial t\right) = [0,0]$, with $\mathbf{u} = [V,W]$. (For technical reasons we use capital letters for the arguments of $\mathbf{u}$.) Then, using the eigenvectors and projection matrices found above, $AsymptSimp(S, [V,W], [x,t], \epsilon, [1,1], P_1, [[1,-1]], P_2, 0)$ yields results that agree with those obtained in Section 10.3.

Exercises 10.4

10.4.1. Apply $RegPerturb$ to the amplitude equation $AmplEq$ given above to obtain the transport equations (10.1.12)–(10.1.13).

10.4.2. Use the procedure $DiscFunct$ with $n = 3$ to obtain a Heaviside function representation for the function $F(x)$ with a jump at $x = \alpha$. Then specialize the result to the function $F(x)$ defined above.

10.4.3. Apply $DiscFunctMult$ to the function given in (10.2.22)–(10.2.23).

10.4.4. Consider the telegrapher's equation and use $SingularExp$ to obtain the result above and then specialize to the case(s) with $\phi(x,t) = x \pm \gamma t \pm 1$.

10.4.5. Apply the procedure $ProjMatrix$ to the matrix A that arises in shallow water theory and obtain the results given in the text.

10.4.6. Apply the procedure $AsymptSimp$ to the shallow water equations and obtain results that correspond to those given in Section 10.3.

10.4.7. Carry out the asymptotic simplification of the Navier-Stokes equations, as given in (10.3.43), using $ProjMatrix$ and $AsymptSimp$. Obtain the Burgers' and heat equation approximations.

CHAPTER 11

FINITE DIFFERENCE METHODS

The *method of finite differences* is a numerical method for the solution of well- posed *initial value, initial and boundary value,* and *boundary value problems* for linear and nonlinear PDEs. Suppose that a problem is formulated for a function $u(\mathbf{x}, t)$. The difference method is based on the process of replacing the partial derivatives of $u(\mathbf{x}, t)$ by approximating difference quotients. A grid of discrete points in the coordinate space of the independent variables—appropriate to the region G in which the problem is formulated—is introduced. The PDE and the auxiliary conditions (should they involve partial derivatives) are represented in terms of difference quotients. At grid points where the auxiliary conditions specify the values of $u(\mathbf{x}, t)$, these values are used. (If the region does not have a simple form, special methods are needed to assign solution values at various grid points near initial and boundary regions.) This process yields a system of algebraic equations for the (approximate) solution $\hat{u}(\mathbf{x}, t)$ of the given problem, determined at the grid points. The finite difference solution $\hat{u}(\mathbf{x}, t)$ approximates the solution $u(\mathbf{x}, t)$ at the grid points of G. Now, the solution methods presented in Chapters 2–10 (generally) yielded exact or approximate solutions at all points in G. By refining the grid and increasing the number of grid points, the finite difference result can be improved, so that $\hat{u}(\mathbf{x}, t)$ approximates $u(\mathbf{x}, t)$ to a high level

Partial Differential Equations of Appplied Mathematics, Third Edition. By Erich Zauderer

of accuracy at any grid point in G. Then, approximate solution values at points outside the grid but within G can be obtained by interpolation.

Whereas it is formally possible to replace any partial derivative by an approximating difference quotient, it turns out that not all algebraic systems obtained thereby yield accurate descriptions of the solution at the grid points. More precisely, the numerical solution does not approximate the exact solution of the problem (evaluated at the grid points) to an increasingly improved level of accuracy, as the grid is increasingly refined. That is, the approximate solution fails to *converge* to the exact solution. Conditions must be placed on the finite difference approximations, such as *consistency* of the difference equations with the given PDE and *stability* of the difference scheme, for *convergence* to take place. Whereas it is usually a straightforward matter to show that a finite difference scheme is consistent with a PDE, the conditions for the stability of the scheme are harder to obtain and verify in general, but, as will be seen below, they can be determined for a number of problems. Proof of the convergence of the finite difference result to the exact solution is the hardest of all. Fortunately, there is a general result due to *Lax*, known as the *Lax equivalence theorem*, that states that *consistency* and *stability* are equivalent to *convergence*, for linear difference schemes for well-posed initial and initial and boundary value problems. As a result, once a linear difference scheme is proven to be consistent and stable, the convergence of its (approximate) solution to the exact solution of the problem is guaranteed. The verification of the consistency of difference schemes and the determination of conditions for their stability are emphasized in our discussion below.

The random walk models of Chapter 1 represent finite difference approximations to a large class of PDEs as was seen. The approach used in that chapter differs from that presented here, in that we obtained PDEs as limits of difference equations rather than constructing difference equations from PDEs, as will be done here. An important consequence of this is that questions relating to consistency, and even stability, did not arise for the problems considered there. The finite difference approximations introduced in this chapter greatly increase the number of difference equations that are related to a given PDE, but they also include the difference equations introduced on the basis of random walk models in Chapter 1. It will be seen that useful insights relating to consistency and stability considerations for finite difference approximations follow from random walk analogies.

The replacement of derivatives by difference quotients in finite difference methods is referred to as a *discretization process*. In general, all derivatives are replaced by differences so that the resulting finite difference method is *fully discrete*. However, when dealing with time-dependent problems, it is often useful to employ a numerical method in which only the spatial partial derivatives are approximated by difference quotients. The time derivatives in the PDEs are not discretized. Thus, the numerical method is *semi-discrete*. This method gives rise to a system of coupled ODEs in the time variable, with prescribed initial conditions. The resulting system can either be solved exactly or can again be solved numerically, using any of numerical methods available for ODEs. (In the latter case, the method becomes fully discrete.) This approach is referred to as the *method of lines* and will be introduced below for the numerical treatment of parabolic and hyperbolic equations.

The use of a computer is indispensable for the numerical solution of problems for PDEs if a reasonable number of grid points is to be considered. To this end we have constructed a number of *Maple* procedures that solve the finite difference problems, enable graphical representations of the solutions, and deal with related questions, such as the stability of these problems. Procedures that deal with parabolic, elliptic, and hyperbolic PDEs, as well as hyperbolic systems of PDEs, are presented. We note that the built-in Maple procedure *pdsolve* for the solution of PDEs or systems of PDEs contains an option that permits a large class of time-dependent problems with one space variable to be solved numerically using finite difference methods. Many of the difference schemes presented in this chapter can be used, and the output can be presented in various forms.

11.1 FINITE DIFFERENCE OPERATORS

The basic finite difference operators or procedures that have been constructed, yield first order *forward*, *backward* and *centered finite difference quotients*. By iterating these operators, higher-order differences can be generated. We use these operators to approximate partial derivatives by difference quotients in the given PDEs or systems of PDEs.

Forward, Backward, and Centered Differences: Maple Procedures

The Maple procedures *ForwDiff*, *BackwDiff*, *CentDiff* and *ShiftOp* generate *forward*, *backward*, and *centered differences* and *shifts*, respectively, for functions in any number of variables. The difference and shift operators can be used, in combination, to approximate derivatives of any order.

First order forward, backward and centered difference quotients for a function $f(x,t)$ with a positive increment (or step size) h in the x-variable are

$$ForwDiff(f(x,t),[x,t],[x,t],[h,0],[x]) = \frac{f(x+h,t)-f(x,t)}{h}, \qquad (11.1.1)$$

$$BackwDiff(f(x,t),[x,t],[x,t],[h,0],[x]) = \frac{f(x,t)-f(x-h,t)}{h}, \qquad (11.1.2)$$

$$CentDiff(f(x,t),[x,t],[x,t],[h,0],[x]) = \frac{f(x+h,t)-f(x-h,t)}{2h}. \qquad (11.1.3)$$

The first order shift operator is

$$ShiftOp(f(x,t),[x,t],[x,t],[\pm h,0]) = f(x \pm h,t). \qquad (11.1.4)$$

If $h = 0$, the shift operator is an identity operator. Clearly, the shift operator can be used to construct the three types of difference quotients.

To obtain the forward difference quotient for a function $f(x,t)$ with an increment k in the t-variable we invoke

$$ForwDiff(f(x,t),[x,t],[x,t],[0,k],[t]) = \frac{f(x,t+k)-f(x,t)}{k}. \qquad (11.1.5)$$

For functions $f(\mathbf{x}) = f(x_1, \ldots, x_j, \ldots, x_n)$ of n variables (where $n \geq 2$) we obtain first order forward difference quotients, with increment h_j in x_j as,

$$\textit{ForwDiff}(f(\mathbf{x}), [\mathbf{x}], [\mathbf{x}], [0, \ldots, h_j, \ldots, 0], [x_j])$$

$$= \frac{f(x_1, \ldots, x_j + h_j, \ldots, x_n) - f(x_1, \ldots, x_j, \ldots, x_n)}{h_j}. \quad (11.1.6)$$

The same is true for the other first order difference operators.

Next, we consider the accuracy of the approximation of the first order derivatives by the difference quotients. For the expressions in (11.1.1)–(11.1.3), this is done by applying the Taylor formula to each difference quotient and expanding around $h = 0$. This yields an expansion in powers of h with derivatives of $f(x, t)$ taken with respect to x. The (built-in) Maple procedure *mtaylor*, (which finds multi-dimensional Taylor series) with m as the number of terms to be exhibited in the output, truncates the Taylor series at the term of $O(h^m)$. The terms in the expansion involve $f(x, t)$ and its derivatives evaluated at (x, t). To obtain the *Taylor formula* we must replace x, in the argument of the $O(h^m)$ term, by ξ, a variable that is unspecified but must satisfy $x < \xi < x + h$, for the case of a forward difference quotient. For a backward difference we have $x - h < \xi < x$, and for a centered difference $x - h < \xi < x + h$. The results are

$$\text{Forward difference: } f_x(x, t) - \frac{f(x + h, t) - f(x, t)}{h} = -\frac{1}{2} f_{xx}(\xi, t)h, \quad (11.1.7)$$

$$\text{Backward difference: } f_x(x, t) - \frac{f(x, t) - f(x - h, t)}{h} = \frac{1}{2} f_{xx}(\xi, t)h, \quad (11.1.8)$$

$$\text{Centered difference: } f_x(x, t) - \frac{f(x + h, t) - f(x - h, t)}{2h} = -\frac{1}{6} f_{xxx}(\xi, t)h^2. \quad (11.1.9)$$

The term on the right side of each equation (11.1.7)–(11.1.9) is called the *truncation error* for each of the difference formulas. As we are concerned with small values of h, we see that the centered difference quotient yields the best approximation to the first derivative. The truncation errors are $O(h)$ for the forward and backward differences, so we say that these approximations are *first order accurate*. The centered difference approximation whose truncation error is $O(h^2)$ is *second order accurate*.

Second and higher-order difference quotients are obtained by iterating the first order difference operators. However, second order difference quotients can also be obtained by a single application of each one of the three procedures. Thus, to obtain the second order centered difference quotient for $f(x, t)$ with a positive increment (or step size) h in the x-variable, we can invoke either of the following procedures:

$$\textit{CentDiff}(f(x,t),[x,t],[x,t],[h,0],[x,x]) = \frac{f(x+h,t)-2f(x,t)+f(x-h,t)}{h^2}, \tag{11.1.10}$$

$$\textit{CentDiff}(\textit{CentDiff}(f(x,t),[x,t],[x,t],[h/2,0],[x]),[x,t],[x,t],[h/2,0],[x])) = \frac{f(x+h,t)-2f(x,t)+f(x-h,t)}{h^2}. \tag{11.1.11}$$

[To obtain the form for the difference quotient given in (11.1.11), an increment $h/2$ was used in each *CentDiff* procedure.] The difference quotients approximate the second order x-derivative, and on using the Maple procedure *mtaylor*, we find that the truncation error is given as

$$f_{xx}(x,t) - \frac{f(x+h,t)-2f(x,t)+f(x-h,t)}{h^2} = O(h^2). \tag{11.1.12}$$

Thus the centered difference approximation is *second order accurate.*

The second order centered difference quotient for $f(x,t)$ with positive increments h and k in the x- and t-variables, respectively, approximates the mixed partial derivative $f_{xt}(x,t)$. The use of

$$\textit{CentDiff}(f(x,t),[x,t],[x,t],[h,k],[x,t]) = \frac{f(x+h,t+k)-f(x-h,t+k)-f(x+h,t-k)+f(x-h,t-k)}{4hk} \tag{11.1.13}$$

yields the required result. Applying the Maple procedure *mtaylor* shows that the truncation error is given as

$$f_{xt}(x,t) - \frac{f(x+h,t+k)-f(x-h,t+k)-f(x+h,t-k)+f(x-h,t-k)}{4hk} = O(h^2+k^2), \tag{11.1.14}$$

so that the centered difference approximation is *second order accurate.*

Example 11.1. Difference Approximations for a Prescribed Function. To compare the various levels of accuracy obtained from the variety of difference quotient approximations to first and second derivatives of a function, we apply them to the function $f(x,y,z) = e^{xy-z^2}$ at the point $(x,y,z) = (3,2,-1)$ with the increment $h = 0.0001$, in all cases. [With regard to mixed partial derivatives, only $f_{xy}(3,2,-1)$ is determined.] The centered difference approximation to $f_x(3,2,-1)$ is found from the procedure $\textit{CentDiff}(e^{xy-z^2},[x,y,z],[3,2,-1],[0.0001,0,0],[x])$ to be 296.83. (The third argument in the procedure prescribes the point at which the approximation is to be found.) All other approximations are determined by using

similar arguments in the appropriate procedures. The exact values of the partial derivatives are also found and the results are listed in tabular form.

e^{xy-z^2}	$\frac{\partial}{\partial x}$	$\frac{\partial}{\partial y}$	$\frac{\partial}{\partial z}$	$\frac{\partial^2}{\partial x^2}$	$\frac{\partial^2}{\partial y^2}$	$\frac{\partial^2}{\partial z^2}$	$\frac{\partial^2}{\partial x\,\partial y}$
Exact	296.82	445.23	296.82	593.64	1335.7	296.82	1038.9
ForwDiff	296.86	445.31	296.84	590.0	1340.0	300.0	1040.0
BackwDiff	296.8	445.17	296.81	600.0	1340.0	290.0	1050.0
CentDiff	296.83	445.24	296.83	590.0	1330.0	300.0	1040.0

An inspection of the values given in the table shows that best overall results are obtained by using centered differences.

While it appears that the smallest truncation errors with a corresponding improved accuracy are achieved by using centered difference quotients, we will see that on replacing derivatives by differences in PDEs, stability considerations preclude the use of certain difference approximations for specific derivatives, even if they yield the smallest truncation error.

Exercises 11.1

11.1.1 Use the procedures *ForwDiff, BackwDiff,* and *CentDiff* to obtain first order differences of $f(x,t)$ in x and in t.

11.1.2. Reproduce the results of Exercise 11.1.1 by using *ShiftOp*.

11.1.3. Use the procedures *ForwDiff, BackwDiff,* and *CentDiff* to obtain all second order differences of $f(x,t)$ in x and in t.

11.1.4. Verify the truncation errors given in (11.1.7)–(11.1.9).

11.1.5. Given the function $f(x,y,z) = \sin(\cos(xyz))$, the point $(x,y,z) = (1,2,3)$, and the increment $h = 0.001$, construct a table as given in Example 11.1.

11.2 FINITE DIFFERENCE METHODS FOR THE ONE-DIMENSIONAL HEAT EQUATION

The one-dimensional nonhomogeneous *heat* or *diffusion equation* for the temperature or concentration $u(x,t)$ of a substance is given as

$$\frac{\partial u(x,t)}{\partial t} - c^2 \frac{\partial^2 u(x,t)}{\partial x^2} = F(x,t), \tag{11.2.1}$$

with a constant $c > 0$. Let $h > 0$ be an increment in x and $k > 0$ be an increment in t. We replace the partial derivatives by various (approximating) difference quotients, thereby creating a number of difference schemes. They can be used to solve Cauchy and initial and boundary value problems for (11.2.1), and a number of them are considered below.

Explicit Forward Difference Method for the One-Dimensional Heat Equation

The most straightforward difference scheme for (11.2.1) is obtained by approximating $u_t(x,t)$ by a *forward first order difference quotient* in t and $u_{xx}(x,t)$ by a *centered second order difference quotient* in x. This yields

$$\frac{u(x,t+k)-u(x,t)}{k}=\frac{c^2\left(u(x+h,t)-2u(x,t)+u(x-h,t)\right)}{h^2}+F(x,t). \tag{11.2.2}$$

We put $r=kc^2/h^2$ and express (11.2.2) as the *difference scheme*

$$u(x,t+k)=(1-2r)u(x,t)+ru(x+h,t)+ru(x-h,t)+kF(x,t). \tag{11.2.3}$$

As will be seen, r represents a *stability parameter* for the difference scheme (11.2.3). The scheme (11.2.3) shows that if the values of the solution $u(x,t)$ are known at the time t at the points $x-h,\ x,\ x+h$, they can be determined at the point x at the time $t+k$. In particular, if we consider a *Cauchy problem* and $u(x,0)=f(x)$ is prescribed, $u(x,t)$ at all points $x=ih,\ i=0,\pm1,\pm2,\ldots$ at the times $t=jk,\ j=1,2,\ldots$ can be found. The points (ih,jk) with $i=0,\pm1,\pm2,\ldots$ and $j=0,1,2,\ldots$ represent the *grid points* for the difference scheme. The *computational stencil* is exhibited in Figure 11.1.

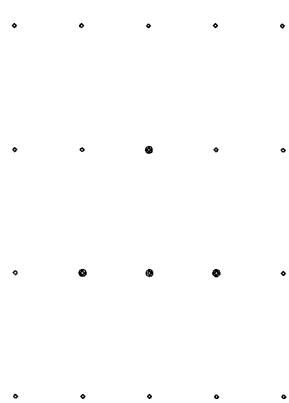

Figure 11.1 The computational stencil.

Next we determine the *truncation error* for the difference equation (11.2.2). Applying the Maple procedure *mtaylor* to (11.2.2) gives

$$u_t(x,t)-c^2u_{xx}(x,t)-F(x,t)\approx-\frac{1}{2}u_{tt}k-\frac{1}{6}u_{ttt}k^2+\frac{c^2}{12}u_{xxxx}h^2+\frac{c^2}{360}u_{xxxxxx}h^4. \tag{11.2.4}$$

The difference between the heat equation (11.2.1) and the difference equation (11.2.2) is of $O(k+h^2)$. This is the order of the *truncation error* for the difference approx-

imation. As $h, k \to 0$, so does the truncation error, so the difference approximation (11.2.2) is *consistent* with the heat equation (11.2.1).

A useful conclusion can be drawn from the representation (11.2.4). If $F(x,t) = 0$ in (11.2.1), and we differentiate the equation with respect to t, it follows that $u_{tt} = c^2 u_{xxt} = c^4 u_{xxxx}$. Then, if we set $k = h^2/6c^2$ in (11.2.4) [where we put $F(x,t) = 0$] the truncation error is reduced to $O(k^2 + h^4)$ with a concomitant improvement in the difference approximation. In terms of the stability parameter $r = kc^2/h^2$, this means that $r = 1/6$.

The difference scheme (11.2.3) is referred to as an *explicit forward difference scheme*. The solution at the time $t + k$ is determined *explicitly* in terms of known solution values at the earlier time t. It is a *forward scheme* in that all the solution values can be determined by moving forward in time, step by step, starting from the initial time. Now, it may appear that the difference scheme yields a valid approximation to the solution without having to place a restriction on the parameter $r = kc^2/h^2$ or on h and k. To see that this is not so, in general, we put $c^2 = 10$ and $F(x,t) = 0$ in (11.2.1), and take as the initial condition $u(x,0) = f(x) = 1/(1+x^2)$. Then with $h = k = 1/10$, so that $r = 100$, the solution of difference equation at $(x,t) = (0, 1/10)$ is $u(0, 1/10) = -199 + 100/[1 + (1/10)^2] + 100/[1 + (1/10)^2] = -99/101$. Now, it has been shown that the solution of the Cauchy problem with a bounded, positive initial value can never be negative. Thus, the negative solution value of $u(0, 1/10)$ has no meaning for the given problem.

In the context of the *random walk problem* considered in Section 1.1, the equation (1.1.16) corresponds to (11.2.3) if we put $F(x,t) = 0$ in (11.2.3), and $\tau = k$, $\delta = h$, and $r = p = q$ in (1.1.16), where p and q are probabilities. Since $p + q \leq 1$ and $p+q = 2r$, we conclude that we must have $r \leq 1/2$ if the probabilistic interpretation is to remain valid. Indeed, this restriction on r will be obtained below when the stability of the difference scheme is analyzed. Therefore, the random walk interpretation plays a role in the determination of the stability condition for the explicit forward difference scheme.

If we are concerned with an *initial and boundary value problem* for (11.2.1), the x-interval in which the problem is to be solved is of finite or semi-infinite extent. We restrict the following discussion to a finite interval $0 \leq x \leq l$, with the initial condition given at $t = 0$ and the boundary conditions at $x = 0$ and $x = l$. The x-interval is divided into n equal parts. The grid in (x,t)-space on which the solution is to be found is given as $(x_i, t_j) = (il/n, jk)$ with $i = 0, 1, 2, \ldots, n$ and $j = 0, 1, 2, \ldots, m$. The x step size is $h = l/n$ and the time step or increment is k. The solution is to be determined at the time $t_m = mk$, $m \geq 1$. The difference scheme (11.2.3) can be written as

$$u_{i,j+1} = (1 - 2r)u_{i,j} + ru_{i+1,j} + ru_{i-1,j} + kF_{i,j}, \tag{11.2.5}$$

where $u_{i,j}$ represents the approximation to the exact solution $u(x,t)$ of the given problem at the point $(x_i, t_j) = (il/n, jk)$, and where $F_{i,j} = F(il/n, jk)$. We again have an *explicit forward difference scheme*.

With the *initial condition* given as $u(x,0) = f(x)$, we have $u_{i,0} = f(il/n)$ for $i = 0, 1, 2, \ldots, n$. The boundary lines $x = 0$ and $x = l$ correspond to the values $i = 0$ and $i = n$. If *first kind conditions* $u(0,t) = g(t)$ and $u(l,t) = s(t)$ are assigned on the boundaries, the values of $u_{0,j} = g(jk)$ and $u_{n,j} = s(jk)$ are prescribed. Thus the solution of the difference equations has to be determined only at the points $(x_i, t_j) = (il/n, jk)$, with $i = 1, 2, \ldots, n-1$ and $j \geq 1$. [Compatibility questions for the initial and boundary data at the points $(0,0)$ and $(l,0)$ need to be addressed. In general, the values at these points are taken as the averages of the initial and boundary values at the points.]

If *second* or *third kind* boundary conditions are prescribed, an x-derivative term occurs in the boundary conditions and it must be replaced by a difference quotient for the difference scheme to be solvable. To accomplish this, we replace the first derivative term by a centered difference quotient in order to achieve the same level of accuracy attained for the centered difference approximation of the second derivative term in the PDE. We use (11.1.8), which yields $u_x(x,t) \approx (u(x+h,t) - u(x-h,t))/2h$. At the boundary $x = 0$, this gives $u_x(0,t) \approx (u(h,t) - u(-h,t))/2h$, and at the boundary $x = l$, we have $u_x(l,t) \approx (u(l+h,t) - u(l-h,t))/2h$. Now the values of $u(x,t)$ are not known at $x = -h$ and at $x = l + h$, as these points lie outside the given x-interval. Additionally, in contrast to the Dirichlet problem, the values of the solution must be determined at $x = 0$ and $x = l$. As a result, the solution of the difference equations must now be found at the points $(x_i, t_j) = (il/n, jk)$ with $i = 0, 1, 2, \ldots, n-1, n$ and $j \geq 1$.

To resolve this problem, we introduce a set of *ghost points* (x_{-1}, t_j) and (x_{n+1}, t_j) that are not elements of the given set of grid points. Now the difference equation (11.2.5) evaluated at $i = 0$ takes the form $u_{0,j+1} = (1-2r)u_{0,j} + ru_{1,j} + ru_{-1,j} + kF_{0,j}$, and at $x = l$ we have $u_{n,j+1} = (1-2r)u_{n,j} + ru_{n+1,j} + ru_{n-1,j} + kF_{n,j}$. The terms $u_{-1,j}$ and $u_{n+1,j}$ are evaluated at ghost points and must be eliminated. This is done by using the centered difference approximations to the first derivative (that occurs in the boundary condition) given above. For example, if there is a Neumann boundary condition at $x = l$ of the form $u_x(l,t) = s(t)$, the difference approximation for $t = jk$ is $u_x(l,jk) \approx u_{n+1,j} - u_{n-1,j} = 2hs(jk)$. Thus $u_{n+1,j} = u_{n-1,j} + 2hs(jk)$, so that $u_{n+1,j}$ (i.e., u evaluated at a ghost point) is expressed in terms of a prescribed value and a value of u at a grid point. The same approach works for a boundary condition of the third kind and for boundary conditions at $x = 0$. In all cases, a solvable system of difference equations for $u_{i,j}$ at all grid points is obtained. (The ghost points no longer play a role in the difference scheme.) The approximate solution is determined at the boundary points as well. If there are mixed boundary conditions and there are Dirichlet conditions at $x = 0$, say, there is no need for ghost points to the left of the boundary at $x = 0$ and the solution $u_{i,j}$ is obtained for $i = 1, 2, \ldots, n$.

The use of the *explicit forward difference scheme* to determine approximate solutions of initial and boundary value problems for the one-dimensional heat or diffusion equation (11.2.1) over the interval $a < x < b$ with boundary conditions of the first, second, third, or mixed kind, involves a significant number of steps and calculations if the grid is of a reasonable size. We have constructed the procedure $NumHeatForw$

to automate this process. It is given as $NumHeatForw(c^2, F(x,t), t = t_0..t_f, f(x), x = a..b, bcl, g(t), bcr, s(t), n, k)$. The initial condition is $u(x,t_0) = f(x)$, and the solution is to be found at the time $t = t_f$. The given interval (a,b) is divided into n equal parts, and the x step size is $h = (b-a)/n$. The time increment is specified to be a positive number k, the last argument in the procedure. (If $t_f - t_0$ is not an integral multiple of k, the procedure determines the solution at the largest value of t that is an integral multiple of k but lies below t_f.) The arguments bcl and bcr characterize the boundary conditions at $x = a$ and $x = b$, respectively. If $bcl = dirichlet$ and $bcr = dirichlet$, Dirichlet boundary conditions $u(a,t) = g(t)$ and $u(b,t) = s(t)$ are specified at the left and right end points. If bcl and bcr are both zero, Neumann boundary conditions $u_x(a,t) = -g(t)$ and $u_x(b,t) = s(t)$ are assigned at the left and right endpoints. Finally, if $bcl = \alpha$ and $bcr = \beta$, Robin boundary conditions $u_x(a,t) - \alpha u(a,t) = -g(t)$ and $u_x(b,t) + \beta u(b,t) = s(t)$ are assigned at the left or right endpoints. The boundary conditions can be mixed. Additionally, the procedure contains global variables that can be used to plot the finite difference solutions as curves or surfaces. This is demonstrated in the examples that are presented below.

The initial and boundary value problem that is to be solved is displayed as part of the output of the procedure, as are the values of the increments h and k. The value of the stability parameter $r = c^2k/h^2$ is also displayed. As will be shown below, using the von Neumann stability analysis, we must have $r \leq 1/2$, otherwise, the method is unstable. If the parameter r exceeds $1/2$, the output states that the difference method is unstable. (However, the results may still be valid.) The procedure represents derivative boundary conditions in the appropriate forms and displays the finite difference solution values at the points $x_i = a + i(b-a)/n,\ i = 0, 1, \ldots, n$ at the time $t = t_f$.

Example 11.2. A Dirichlet Problem. The procedure $NumHeatForw(1, 0, t = 0..0.5, \sin(\pi x), x = 0..1, dirichlet, 0, dirichlet, 0, 10, .005)$ obtains the numerical solution of the initial and boundary value problem described in the output of the procedure as given below.

$$u_t(x,t) - u_{xx}(x,t) = 0, \qquad x = 0..1,\ 0 < t, \tag{11.2.6}$$

$$u(x,0) = \sin(\pi x),\ x = 0..1, \qquad u(0,t) = 0,\ u(1,t) = 0, \qquad 0 < t, \tag{11.2.7}$$

$$k = 0.005,\ h = 1/10,\ N = 10,\ r = 0.500,\ t = 0.5. \tag{11.2.8}$$

The tenth argument of the procedure indicates that the interval $0 < x < 1$ is divided into 10 equal subintervals so that $h = 1/10$ and the eleventh argument specifies the time increment as $k = 0.005$. The parameter r equals 0.5, so that the scheme is stable. (If the parameter $r > 0.5$, the output of the procedure states that the difference method is unstable.) The last entry, $t = 0.5$, in (11.2.8) indicates that the solution is to be found at $t = 0.5$. [We note that the initial and boundary values are compatible at the points (0,0) and (1,0).]

The procedure displays the solution in the following tabular form:

$$\begin{bmatrix} x & u(x,0.5) \\ 0.0 & 0.0 \\ 0.10 & 0.002044630895 \\ 0.20 & 0.003889119061 \\ 0.30 & 0.005352913177 \\ 0.40 & 0.006292726828 \\ 0.50 & 0.006616564564 \\ 0.60 & 0.006292726828 \\ 0.70 & 0.005352913177 \\ 0.80 & 0.003889119061 \\ 0.90 & 0.002044630895 \\ 1.0 & 0.0 \end{bmatrix}. \tag{11.2.9}$$

The global variable $PList$ lists the grid values at $t = 0.5$. They are given as $[[0,0],[0.1,.00204],[0.2,.00389],[.3,.00535],[.4,.00629],[.5,.00662],[.6,.00629],[.7,.00535],[.8,.00389],[.9,.00204],[1,0]]$. The exact solution of the problem is $u(x,t) = \sin(\pi x)\exp(-\pi^2 t)$. The values of the exact solution at the grid points with $t = 0.5$ is $[0,0],[.1,.00222],[.2,.00423],[.3,.00582],[.4,.00684],[.5,.00719],[.6,.00684],[.7,.00582],[.8,.00423],[.9,.00222],[1,0]$.

The difference approximation to $u(x,0.5)$ can be plotted with $plot(PList)$. A plot of a sequence of finite difference approximations for $u(x,t_j)$, as t_j ranges through a set of times from $t = 0$ to $t = 0.5$, can be generated by invoking $NumHeatForw$ with each of the t_j as the final time. Each application of the procedure yields a global variable $PList$. If each of these is assigned a name $PList(j)$, the solution curves are displayed with $plot(\{PList(1), PList(2), \ldots\})$.

A plot of the finite difference solution as a function of x and t can be generated by using the built-in Maple procedure *surfdata*$(PList3dMod)$, where $PList3dMod$ is a global variable for $NumHeatForw$ that lists $[x_i, t_j, u(x_i,t_j)]$ as the (x_i,t_j) range through all the grid values for the problem. The *surfdata* procedure is in Maple's *plots* package and generates a surface in (x,t,u)-space from a list of points given as its argument. We denote this plot by $P1$. The plot of the exact solution $u(x,t)$ for $0 \le t \le 0.5$ is given by $P2 = plot3d(\sin(\pi x)\exp(-\pi^2 t), x = 0..1, t = 0..0.5)$. Then, $plots[display](\{P1, P2\})$ yields the plot shown in Figure 11.2. We observe that $u(x,t)$ decays rapidly from its initial value $u(x,0) = \sin(\pi x)$.

The agreement between the exact and the finite difference solution values could stand improvement. A more accurate finite difference result can be expected if the h and k in the procedure are chosen so that the stability parameter $r = 1/6$, as was indicated above. If we replace the last argument 0.005 by $1/600$, so that $k = 1/600$,

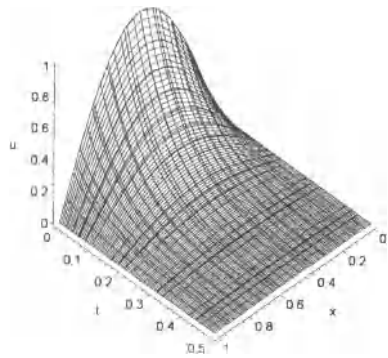

Figure 11.2 The exact and finite difference solutions.

the parameter r equals $1/6$. In that case, the global variable $PList$ yields the same values as for the exact solution.

To exhibit the erratic behavior of the forward difference scheme in the case of *instability*, we reapply the foregoing procedure and replace the last argument by 0.05 so that $k = 0.05$. This results in $r = 5$ as the value of the stability parameter, so that the scheme is unstable. Carrying out the procedure yields the solution values at $t = 0.5$ as given by $PList$, $[[0, 0], [.1, -90.3], [.2, 184.63], [.30, -290.81], [.4, 418.6], [.50, -567.65], [.6, 708.52], [.70, -774.33], [.80, 686.02], [.9, -409.12], [1, 0]]$. They have no relevance to the solution of the problem.

Example 11.3. A Mixed Problem. The procedure $NumHeatForw(1, 2tx^2 - 2t^2, t = 0..1, 0, x = 0..1, 0, 0, 1, 3t^2, 4, .001)$ solves the following initial and boundary value problem for the nonhomogeneous heat equation

$$u_t(x,t) - u_{xx}(x,t) = 2tx^2 - 2t^2, \qquad x = 0..1,\ 0 < t, \tag{11.2.10}$$

with the conditions $u(x,0) = 0$, $x = 0..1$, $u_x(0,t) = 0$, $u_x(1,t) + u(1,t) = 3t^2$, $0 < t$. The explicit forward difference method has the increments $k = 0.001$, $h = 1/4$, with $r = 0.016$, and the solution is found at $t = 1$.

The solution values at $t = 1$ given by $PList$ are $[[0, -.001765], [.25, .06069], [.50, .24794], [.75, .55998], [1.0, .99682]]$. Since Neumann and Robin boundary conditions are prescribed at the endpoints, the finite difference scheme determines approximate solution values not only at interior points but also at the endpoints. The exact solution of the problem is $u(x,t) = x^2t^2$, and its values at $t = 1$ are $[[0., 0.], [.25, .062500], [.50, .25000], [.75, .56250], [1., 1.]]$. The accuracy of the approximation can be improved if the size of the time step k is reduced from 0.001 to 0.0001, say.

This yields a significant decrease in the stability parameter r. $plot(\{x^2, PList\}, x = 0..1)$ plots the finite difference and exact solutions at $t = 1$. The plot is exhibited in Figure 11.3.

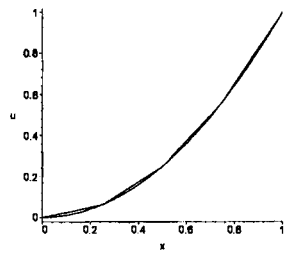

Figure 11.3 The exact and finite difference solutions at $t = 1$.

It is of interest to express the explicit forward difference scheme in *matrix form*. It exhibits directly how the solution values of $u_{i,j+1}$ for all the (relevant) values of i in the grid (given in vector form) are expressed in terms of products of matrices and vectors that involve prescribed initial and boundary values. This representation is used in the matrix stability analysis of the difference scheme to be presented below.

For *Dirichlet boundary conditions* with $u(0,t) = g(t)$ and $u(l,t) = s(t)$, the system of linear equations for the $u_{i,j}$ has the matrix form

$$\mathbf{u}_{j+1} = B\mathbf{u}_j + \mathbf{w}_j, \qquad j \geq 0, \tag{11.2.11}$$

where

$$B = \begin{bmatrix} 1-2r & r & & & \mathbf{0} \\ r & 1-2r & r & & \\ & \ddots & \ddots & \ddots & \\ & & & & r \\ \mathbf{0} & & & r & 1-2r \end{bmatrix}, \tag{11.2.12}$$

$$\mathbf{u}_j = \begin{bmatrix} u_{1,j} \\ u_{2,j} \\ \vdots \\ u_{n-1,j} \end{bmatrix}, \qquad \mathbf{w}_j = \begin{bmatrix} kF_{1,j} + rg_j \\ kF_{2,j} \\ \vdots \\ kF_{n-1,j} + rs_j \end{bmatrix}, \tag{11.2.13}$$

with $g_j = g(jk)$ and $s_j = s(jk)$. The matrix B in (11.2.12) is in *tridiagonal form*. That is, all the elements of the matrix below the subdiagonal and above the superdiagonal are zero.

With *Neumann boundary conditions* $u_x(0,t) = g(t)$ and $u_x(l,t) = s(t)$, the system of linear equations for the $u_{i,j}$ again has the form (11.2.11) but with

$$B = \begin{bmatrix} 1-2r & 2r & & & \mathbf{0} \\ r & 1-2r & r & & \\ & \ddots & \ddots & \ddots & \\ & & & & r \\ \mathbf{0} & & & 2r & 1-2r \end{bmatrix}, \tag{11.2.14}$$

$$\mathbf{u}_j = \begin{bmatrix} u_{0,j} \\ u_{1,j} \\ \vdots \\ u_{n,j} \end{bmatrix}, \qquad \mathbf{w}_j = \begin{bmatrix} kF_{0,j} - 2hrg_j \\ kF_{2,j} \\ \vdots \\ kF_{n,j} + 2hrs_j \end{bmatrix}. \tag{11.2.15}$$

The dimension of the matrix B is adjusted accordingly. The forms taken by the matrix B and the vectors $\mathbf{u}_j$ and $\mathbf{w}_j$ if the boundary conditions are of the third kind or are mixed are easily obtained. In each case, the $\mathbf{u}_j$ can be solved for recursively since $\mathbf{u}_0$ is prescribed.

This concludes our discussion of the *explicit forward difference scheme* for the heat or diffusion equation. Although the procedure *NumHeatForw* is designed to deal with initial and boundary value problems, it can be used for initial value problems as well if the x-interval is chosen to be sufficiently large so that the boundary conditions (which can then be assigned arbitrarily) do not play a role in the solution at the time t_f when the solution is to be found. The stability criterion, however, cannot be neglected. The stability condition $r = kc^2/h^2 \leq 1/2$ that must be satisfied for the explicit scheme to be stable implies that if we want to choose a small step size h to improve the accuracy of the result, it becomes necessary to decrease the time increment k correspondingly. This can give rise to an excessive number of time steps if the solution is to be found at a time that is not close to the initial time. The use of a large number of steps in carrying out the scheme can lead to significant round-off errors and can adversely affect the accuracy of the numerical results.

This problem can be avoided by replacing the *explicit forward difference scheme* by an *implicit backward difference scheme*, which we present below. It has the advantage of being stable for all values of the stability index $r = kc^2/h^2$. However, because it is a backward scheme, the difference equations at each time step for all x values in the grid must be solved simultaneously rather than individually for each x value, as in the forward scheme. For this reason it can only be used for initial and (finite) boundary value problems. Nevertheless, because stability considerations do not restrict the size of the time increment—the scheme is stable for all values of the stability parameter r—and despite the need to solve a possibly large number of simultaneous equations (which can be carried out efficiently), the backward difference scheme is preferred over the forward difference scheme.

Implicit Backward Difference Method for the One-Dimensional Heat Equation

We again consider the nonhomogeneous heat or diffusion equation (11.2.1) for the temperature or concentration $u(x,t)$ in one dimension and let $h > 0$ be an increment in x and $k > 0$ be an increment in t. Now, however, we approximate $u_t(x,t)$ by a *backward first order difference quotient* in t and $u_{xx}(x,t)$ by a *centered second order difference quotient* in x. This yields

$$\frac{u(x,t) - u(x,t-k)}{k} = \frac{c^2(u(x+h,t) - 2\,u(x,t) + u(x-h,t))}{h^2} + F(x,t). \tag{11.2.16}$$

To determine the *truncation error* that results on replacing the heat equation (11.2.1) by the difference equation (11.2.16), we apply the Maple procedure *mtaylor* to (11.2.16) and obtain

$$u_t(x,t) - c^2 u_{xx}(x,t) - F(x,t) \approx \frac{1}{2}u_{tt}k - \frac{c^2}{12}u_{xxxx}h^2. \tag{11.2.17}$$

The difference between the heat equation (11.2.1) and the difference equation (11.2.11) is of $O(k+h^2)$, which is the *truncation error* for the difference approximation. The truncation error tends to zero as $h, k \to 0$, so that the difference approximation (11.2.16) is *consistent* with the heat equation (11.2.1).

We again introduce the stability parameter $r = kc^2/h^2$ and express (11.2.16) as the difference scheme

$$u(x,t-k) = (1+2r)u(x,t) - ru(x+h,t) - ru(x-h,t) - kF(x,t). \tag{11.2.18}$$

If $u(x,t-k)$ is known, we cannot determine $u(x-h,t)$, $u(x,t)$, and $u(x+h,t)$ at the later time t from a single equation. Instead, (11.2.18) must be applied over the entire set of values of x in the grid for the problem.

As indicated, we are concerned with an *initial and boundary value problem* for (11.2.1) over the finite interval $0 < x < l$. The initial condition is given at $t = 0$ and the boundary conditions at $x = 0$ and $x = l$. We divide the interval into n equal parts. The grid in (x,t)-space, on which the solution is to be found, is given as $(x_i, t_j) = (il/n, jk)$ with $i = 0,1,2,\ldots,n$ and $j = 0,1,2,\ldots,m$. The x step size is $h = l/n$ and the time increment is k. The solution is to be determined at the time $t_m = mk$. The difference scheme (11.2.18) can then be written in the form

$$u_{i,j-1} = (1+2r)u_{i,j} - ru_{i+1,j} - ru_{i-1,j} - kF_{i,j}, \tag{11.2.19}$$

where $u_{i,j}$ represents the approximation to the exact solution $u(x,t)$ of the given problem at the point $u(x_i,t_j) = u(il/n, jk)$ and $F_{i,j} = F(il/n, jk)$. This is an *implicit backward difference scheme*. It is *unconditionally stable*, as will be shown.

The *initial condition* $u(x,0) = f(x)$ yields $u_{i,0} = f(il/n)$ with $i = 0,1,2,\ldots,n$. The boundary lines $x = 0$ and $x = l$ correspond to the values $i = 0$ and $i = n$. If

Dirichlet conditions $u(0,t) = g(t)$ and $u(l,t) = s(t)$ are assigned on the boundaries, the values of $u_{0,j} = g(jk)$ and $u_{n,j} = s(jk)$ are prescribed. The solution of the difference equations must be determined at the points $(x_i, t_j) = (il/n, jk)$ with $i = 1, 2, \ldots, n-1$ and $j \geq 1$. If *Neumann* or *Robin conditions* are prescribed, a derivative term occurs in the boundary conditions and must be replaced by a difference quotient for the difference scheme to be solvable. Then the approach presented above in our discussion of the explicit forward difference scheme is used, and the details will not be repeated here.

The main problem is how to proceed from the initial time $t = 0$ where the solution $u(x,0)$ (i.e., $u_{i,0}$) is prescribed, to the time $t = k$ and onward. The first step is to determine $u_{i,1}$ for all i values for which it is unspecified and we do so by letting i range through all these values in (11.2.19). For example, if Dirichlet boundary conditions are given at both end points, the index i ranges from 1 to $n-1$, while if Neumann or Robin boundary conditions are specified at the end points, the range of i extends from 0 to n. Thus, in the former case (11.2.13) yields $n-1$ simultaneous linear equations for the unknowns $u_{i,1}$ $(i = 1, 2, \ldots, n-1)$, while in the latter case there are $n+1$ simultaneous linear equations for the unknowns $u_{i,1}$ $(i = 0, 1, 2, \ldots, n)$. Once the $u_{i,1}$ are determined, this process can be repeated to determine the $u_{i,2}$, $u_{i,3}, \ldots, u_{i,m}$.

For the case of *Dirichlet boundary conditions* with $u(0,t) = g(t)$ and $u(l,t) = s(t)$, the system of linear equations for the $u_{i,j}$ is represented in matrix form as

$$A\mathbf{u}_j = \mathbf{u}_{j-1} + \mathbf{w}_j, \qquad j \geq 1, \tag{11.2.20}$$

where

$$A = \begin{bmatrix} 1+2r & -r & & & \mathbf{0} \\ -r & 1+2r & -r & & \\ & \ddots & \ddots & \ddots & \\ & & & & -r \\ \mathbf{0} & & & -r & 1+2r \end{bmatrix}, \tag{11.2.21}$$

$$\mathbf{u}_j = \begin{bmatrix} u_{1,j} \\ u_{2,j} \\ \vdots \\ u_{n-1,j} \end{bmatrix}, \qquad \mathbf{w}_j = \begin{bmatrix} kF_{1,j} + rg_j \\ kF_{2,j} \\ \vdots \\ kF_{n-1,j} + rs_j \end{bmatrix}, \tag{11.2.22}$$

with $g_j = g(jk)$ and $s_j = s(jk)$.

With *Neumann boundary conditions* $u_x(0,t) = g(t)$ and $u_x(l,t) = s(t)$, the system of linear equations for the $u_{i,j}$ again has the form (11.2.20) but with

$$A = \begin{bmatrix} 1+2r & -2r & & & \mathbf{0} \\ -r & 1+2r & -r & & \\ & \ddots & \ddots & \ddots & \\ & & & & -r \\ \mathbf{0} & & & -2r & 1+2r \end{bmatrix}, \tag{11.2.23}$$

$$\mathbf{u}_j = \begin{bmatrix} u_{0,j} \\ u_{1,j} \\ \vdots \\ u_{n,j} \end{bmatrix}, \qquad \mathbf{w}_j = \begin{bmatrix} kF_{0,j} - 2hrg_j \\ kF_{2,j} \\ \vdots \\ kF_{n,j} + 2hrs_j \end{bmatrix}. \tag{11.2.24}$$

In the case of *Robin boundary conditions* $u_x(0,t) - \alpha u(0,t) = g(t)$ and $u_x(l,t) + \beta u(l,t) = s(t)$ with $\alpha > 0$, $\beta > 0$, the system of linear equations for the $u_{i,j}$ again has the form (11.2.20), but with

$$A = \begin{bmatrix} 1+2r+2hr\alpha & -2r & & & \mathbf{0} \\ -r & 1+2r & -r & & \\ & \ddots & \ddots & \ddots & \\ & & & & -r \\ \mathbf{0} & & & -2r & 1+2r+2hr\beta \end{bmatrix} \tag{11.2.25}$$

and $\mathbf{u}_j$ and $\mathbf{w}_j$ defined as in (11.2.24). The modifications needed for the case of mixed boundary conditions are easily found and are not presented here.

All the coefficient matrices A in (11.2.20) are of *tridiagonal form*. In addition, they are (strictly) *diagonally dominant*. This means that if A is any $m \times m$ matrix with elements a_{ij}, then

$$|a_{ii}| > \sum_{j=1, j\neq i}^{m} |a_{ij}|, \qquad i = 1, 2, \ldots, m. \tag{11.2.26}$$

For the matrices A exhibited above, the diagonal dominance is apparent, and the same holds true for all matrices that occur in the implicit backward difference scheme. An important consequence of this property is that each matrix A has an inverse, so that the systems that arise at each step of the backward difference scheme are guaranteed to have a unique solution. If A^{-1} is the inverse of A, the equation for $\mathbf{u}_j$ can be written as

$$\mathbf{u}_j = A^{-1}\mathbf{u}_{j-1} + A^{-1}\mathbf{w}_j, \qquad j \geq 1, \tag{11.2.27}$$

and the $\mathbf{u}_j$ can be solved for recursively since $\mathbf{u}_0$ is prescribed. Additionally, when solving the systems using *Gaussian elimination*, the diagonal dominance implies that no row interchanges are required. In combination with the fact that A is tridiagonal, the solution of the linear systems that arise at each step of the scheme becomes fairly easy. Nevertheless, much work is involved in finding approximate solutions via the implicit backward difference scheme.

To automate the numerical solution process, we invoke the Maple procedure *NumHeatBackw*$(c^2, F(x,t), t = t_0..t_f, f(x), x = a..b, bcl, g(t), bcr, s(t), n, k, \theta)$. Each of the first 11 arguments in the procedure has the same meaning as in the procedure *NumHeatForw* presented above. The twelfth argument θ is a parameter that can be assigned a value in the range $0 \leq \theta \leq 1$. It corresponds to the values that can be prescribed for the general θ *difference scheme* that will be presented below. For the present purposes, we put $\theta = 1$ and the procedure solves the problem using the implicit backward difference scheme. The output of the procedure has the same form as in *NumHeatForw*. The value of the stability parameter r is exhibited, but the scheme is stable (as will be shown) for all values of r. The procedure contains global variables that can be used to plot the finite difference solutions as curves or surfaces.

We apply the procedure to the problem considered in Example 11.2.

Example 11.4. A Dirichlet Problem. The procedure $NumHeatBackw(1, 0, t = 0..0.5, \sin(\pi x), x = 0..1, dirichlet, 0, dirichlet, 0, 10, .005, 1)$ obtains the numerical solution of the initial and boundary value problem described in the output of the procedure: $u_t(x,t) - u_{xx}(x,t) = 0$, $x = 0..1$, $0 < t$, $u(x,0) = \sin(\pi x)$, $x = 0..1$, $u(0,t) = 0$, $u(1,t) = 0$, $0 < t$, $k = 0.005$, $h = 1/10$, $N = 10$, $r = 0.500$, $t = 0.5$. The last argument $\theta = 1$ in the procedure causes Maple to use the *implicit backward difference scheme* to solve the problem.

We invoke the global variable $PList$ to generate a list of the solution values at the time $t = 0.5$ in the form $[[0,0], [.1, .0026], [.2, .00494], [.3, .0068], [.4, .008], [.5, .00841], [.6, .008], [.7, .0068], [.8, .00494], [.9, .00260], [1, 0]]$. The exact solution is $u(x,t) = \sin(\pi x)\exp(-\pi^2 t)$. A list of exact solution values at the grid points with $t = .5$ is $[[0,0], [.1, .0022], [.2, .0042], [.3, .0058], [.4, .0068], [.5, .0072], [.6, .00684], [.7, .00582], [.8, .00423], [.9, .00222], [1, 0]]$. Again, the agreement between the exact and the finite difference solution values could stand improvement. As in Example 11.2, the global variable $PList3dMod$ and the procedure $surfdata(PList3dMod)$ can be used to plot the finite difference solution as a function of x and t. The result is similar to that given in Figure 11.1 and is not displayed here.

To show that the implicit backward difference scheme does not exhibit instability, we reapply the foregoing procedure and replace the next-to-last argument by 0.05 so that $k = 0.05$. This results in $r = 5$ as the value of the stability parameter. The use of this value in Example 11.2 yields an unstable forward difference scheme, as was shown in Example 11.2. Carrying out the $NumHeatBackw$ procedure yields the solution values at $t = 0.5$ as given by $PList$, $[[0., 0.], [.10, .00575], [.20, .01090], [.30, .01510], [.40, .01770]], [[.50, .0186], [.60, .01770], [.70, .01510], [.80, .01090], [.90, .00575], [1.0, 0.]]$. These values are certainly not as close to the exact solution values as

before, but they are far superior to the values obtained in Example 11.2 from the forward difference scheme.

Additional Difference Methods for the One-Dimensional Heat Equation

Much effort has been expended to create finite difference schemes that improve on the forward and backward difference schemes presented above. The most obvious method for improving the forward difference scheme is to replace the forward difference quotient approximation to $u_t(x,t)$ in the heat equation by the more accurate centered difference approximation. This yields

$$\frac{u(x,t+k)-u(x,t-k)}{2k}=\frac{c^2\left(u(x+h,t)-2u(x,t)+u(x-h,t)\right)}{h^2}+F(x,t). \tag{11.2.28}$$

On applying the Maple procedure *mtaylor* to (11.2.28), we determine that the truncation error is $O(k^2+h^2)$, which improves on the truncation error $O(k+h^2)$ for the explicit forward scheme.

Using the notation introduced above, the difference scheme becomes

$$u_{i,j+1}=u_{i,j-1}+2r\left(u_{i+1,j}-2u_{i,j}+u_{i-1,j}\right)+2kF_{i,j}. \tag{11.2.29}$$

Although this is an explicit forward scheme, in contrast to the above, it is a *two-step scheme*. To obtain $u_{i,j+1}$ we must know $u_{i,j}$ and $u_{i,j-1}$. But only $u_{i,0}$ is prescribed for a given problem, so that the scheme cannot be used unless the relevant $u_{i,1}$ are completely specified. One method for finding the $u_{i,1}$ is to use the explicit forward difference scheme for a single (time) step and then revert to the foregoing two-step method. Unfortunately, these efforts are in vain, since this two-step scheme is *unstable* for every value of the stability parameter r, so that it should never be used. We conclude that the consistency of a difference scheme with a PDE does not guarantee its validity in approximating the solution of a given problem for the PDE, as it may turn out to be unconditionally unstable.

On the other hand, a stable difference scheme may not be consistent with a given PDE, yet its solution can yield a useful approximation to the solution of a problem for the PDE. An example of this is afforded by the *DuFort-Frankel scheme* for approximating solutions of the heat equation. This is an explicit difference scheme that is unconditionally stable for all values of the stability parameter $r=c^2k/h^2$. It is given as

$$\frac{u(x,t+k)-u(x,t-k)}{2k}$$

$$=\frac{c^2\left(u(x+h,t)-u(x,t-k)-u(x,t+k)+u(x-h,t)\right)}{h^2}+F(x,t). \tag{11.2.30}$$

As was done in (11.2.28), the $u_t(x,t)$ term in the heat equation is approximated by a centered difference quotient. However, instead of approximating the $u_{xx}(x,t)$ term by a centered difference quotient as in (11.2.28), the term $2u(x,t)$ in that difference quotient is replaced by an averaged representation $u(x,t-k)+u(x,t+k)$. The procedure *mtaylor* applied to (11.2.30) yields $u_t(x,t)-c^2u_{xx}(x,t)-F(x,t)\approx -c^2k^2/h^2u_{tt}-k^2/6u_{ttt}+c^2h^2/12u_{xxxx}-c^2k^4/12h^2u_{tttt}$ and shows that the truncation error is $O(k^2+h^2+k^2/h^2)$. Thus, the difference equation is consistent with the heat equation (11.2.1) only if $k=o(h)$ as $k,h\to 0$. But if $k=O(h)$ as $k,h\to 0$, and we put $k=\lambda h$, the difference equation is consistent with the nonhomogeneous telegrapher's equation

$$u_t(x,t)-c^2u_{xx}(x,t)+c^2\lambda u_{tt}(x,t)=F(x,t). \tag{11.2.31}$$

The *DuFort-Frankel scheme* can be written as

$$u_{i,j+1}=u_{i,j-1}+2r\left(u_{i+1,j}-u_{i,j-1}-u_{i,j+1}+u_{i-1,j}\right)+2kF_{i,j}. \tag{11.2.32}$$

It is an *explicit forward two step scheme*—as was (11.2.29)—but it unconditionally stable for all values of the parameter r. The relevant $u_{i,1}$ can be specified in the manner indicated above for (11.2.29). Even though the difference scheme is not consistent with the heat equation for all choices of h and k, it can nevertheless be used to construct approximate solutions. In a similar fashion, it was shown in Section 1.2 that the correlated random walk model, which is related to the telegrapher's equation, approximates the random walk and its related diffusion equation. We have not constructed a procedure that can be applied directly to the DuFort-Frankel scheme, but the procedure $NumCorrRandomWalkConst$ of Section 1.5 can be adapted to deal with this scheme.

An improvement in the *implicit backward difference scheme* [see (11.2.16)] can be achieved by averaging the right-hand side of (11.2.16) over its values at t and $t-k$. This yields the difference equation

$$\frac{u(x,t)-u(x,t-k)}{k}=\frac{c^2(u(x+h,t)-2\,u(x,t)+u(x-h,t))}{2h^2} \tag{11.2.33}$$

$$+\frac{c^2(u(x+h,t-k)-2\,u(x,t-k)+u(x-h,t-k))}{2h^2}+\frac{F(x,t)+F(x,t-k)}{2}.$$

mtaylor yields $u_t(x,t)-c^2u_{xx}(x,t)-F(x,t)=O(k^2+h^2)$. The *truncation error* is $O(k^2+h^2)$ compared to the truncation error $O(k+h^2)$ for the implicit backward scheme, so that (11.2.33) is consistent with (11.2.1).

Using the notation above, the difference equation for $u_{i,j}$ can be written as

$$(1+r)u_{i,j} - \frac{r}{2}(u_{i+1,j} + u_{i-1,j}) - \frac{k}{2}F_{i,j}$$

$$= (1-r)u_{i,j-1} + \frac{r}{2}(u_{i+1,j-1} + u_{i-1,j-1}) + \frac{k}{2}F_{i,j-1}. \tag{11.2.34}$$

This is the *Crank-Nicolson scheme.* It is an *implicit backward difference scheme* and will be shown to be *unconditionally stable.*

The forward, backward, and Crank-Nicolson one-step difference schemes can all be subsumed under what is referred to as the θ *scheme.* It is obtained by replacing $u_t(x,t)$ in the heat equation (11.2.1) by a forward difference and $u_{xx}(x,t)$ and $F(x,t)$ by a weighted average of centered differences evaluated at t and $t+k$. The difference equation is given as

$$\frac{u(x,t+k) - u(x,t)}{k} = \theta \frac{c^2(u(x+h,t+k) - 2u(x,t+k) + u(x-h,t+k))}{h^2}$$

$$+(1-\theta)\frac{c^2(u(x+h,t) - 2u(x,t) + u(x-h,t))}{h^2} + \theta F(x,t+k) + (1-\theta)F(x,t). \tag{11.2.35}$$

To guarantee nonnegative weights, we require that $0 \le \theta \le 1$. With $\theta = 0$, this reduces to the difference equation for the explicit forward scheme. For all other θ values this yields an implicit scheme. With $\theta = 1$ and $\theta = 1/2$, it reduces to the equations for the backward and Crank-Nicolson difference schemes, respectively. The truncation error is $O(k + h^2)$, except when $\theta = 1/2$ (the Crank-Nicolson case). The difference equation is consistent with (11.2.1).

Again retaining the notation introduced above, we obtain

$$(1+2r\theta)u_{i,j+1} - r\theta u_{i+1,j+1} - r\theta u_{i-1,j+1} - k\theta F_{i,j+1}$$

$$= (1-2r(1-\theta))u_{i,j} + r(1-\theta)u_{i+1,j} + r(1-\theta)u_{i-1,j} + k(1-\theta)F_{i,j} \tag{11.2.36}$$

as the θ *scheme.* It reduces to the *explicit forward, implicit backward,* and *implicit Crank-Nicolson schemes* for $\theta = 0$, $\theta = 1$, and $\theta = 1/2$, respectively. It will be shown that if $1/2 \le \theta \le 1$, the scheme is unconditionally stable, whereas if $0 \le \theta < 1/2$, the stability condition is $r \le 1/(2-4\theta)$. Thus, with $\theta = 0$, when the θ scheme reduces to the explicit forward scheme, the stability condition becomes $r \le 1/2$.

The construction of grids and the prescription of initial and boundary values for the $u_{i,j}$ on using the θ scheme to approximate solutions to initial and boundary value problems for the heat or diffusion equation proceeds as for the explicit and implicit methods considered above and is not be repeated here.

For the case of Dirichlet boundary conditions with $u(0,t) = g(t)$ and $u(l,t) = s(t)$, the system of linear equations (11.2.36) for the $u_{i,j}$ is represented in matrix form as

$$A\mathbf{u}_{j+1} = B\mathbf{u}_j + \mathbf{w}_{(j,j+1)}, \qquad j \ge 0, \tag{11.2.37}$$

where

$$A = \begin{bmatrix} 1+2r\theta & -r\theta & & & \mathbf{0} \\ -r\theta & 1+2r\theta & -r\theta & & \\ & \ddots & \ddots & \ddots & \\ & & & & -r\theta \\ \mathbf{0} & & & -r\theta & 1+2r\theta \end{bmatrix}, \quad \mathbf{u}_j = \begin{bmatrix} u_{1,j} \\ u_{2,j} \\ \vdots \\ u_{n-1,j} \end{bmatrix}, \tag{11.2.38}$$

$$B = \begin{bmatrix} 1+2r\theta-2r & -r\theta+r & & & \mathbf{0} \\ -r\theta+r & 1+2r\theta-2r & -r\theta+r & & \\ & \ddots & \ddots & \ddots & \\ & & & & -r\theta+r \\ \mathbf{0} & & & -r\theta+r & 1+2r\theta-2r \end{bmatrix}, \tag{11.2.39}$$

$$\mathbf{w}_{(j,j+1)} = \begin{bmatrix} k\theta F_{1,j+1} + k(1-\theta)F_{1,j} + r\theta g_{j+1} + r(1-\theta)g_j \\ k\theta F_{2,j+1} + k(1-\theta)F_{2,j} \\ \vdots \\ k\theta F_{n-1,j+1} + k(1-\theta)F_{n-1,j} + r\theta s_{j+1} + r(1-\theta)s_j \end{bmatrix}, \tag{11.2.40}$$

with $g_j = g(jk)$ and $s_j = s(jk)$. With $\theta = 0$ and $\theta = 1$, the matrices A and B, respectively, reduce to identity matrices. Then, the matrix formulation agrees with that obtained previously in the cases of explicit forward and implicit backward schemes with Dirichlet boundary conditions.

The matrix A is *tridiagonal* and diagonally dominant, so it has an inverse. If A^{-1} is the inverse of A, the equation for $\mathbf{u}_j$ can be written as

$$\mathbf{u}_{j+1} = A^{-1}B\mathbf{u}_j + A^{-1}\mathbf{w}_{(j,j+1)}, \qquad j \geq 0, \tag{11.2.41}$$

and the $\mathbf{u}_j$ can be solved for recursively since $\mathbf{u}_0$ is prescribed.

To automate the numerical solution process for the θ scheme, we use *NumHeat Backw*$(c^2, F(x,t), t = t_0..t_f, f(x), x = a..b, bcl, g(t), bcr, s(t), n, k, \theta)$. The first 11 arguments are as described above. The twelfth argument θ must be assigned a value in the range $0 \leq \theta \leq 1$ and corresponds to the value prescribed for the θ *difference scheme*. If $\theta = 0$, $\theta = 1/2$ and $\theta = 1$, the explicit forward, Crank-Nicolson, and implicit backward difference schemes, respectively, are used. Again, the procedure contains global variables that can be used to plot the finite difference solutions as curves or surfaces.

Example 11.5. A Mixed Problem. We reconsider the problem of Example 11.3. The procedure $NumHeatBackw(1, 2tx^2 - 2t^2, t = 0..1, 0, x = 0..1, 0, 0, 1, 3t^2, 4, .001, 0)$ uses the *explicit forward difference scheme* as in Example 11.3. The solution values at $t = 1$ as given by $PList$ are $[[0, -.000204], [.25, .062291], [.50, .24978], [.75, .56228], [1.0, .99980]]$. Even though the values of h and k are identical with those used in Example 11.3, the solution values differ because the solution is obtained in a slightly different manner. The exact solution of the problem is $u(x, t) = x^2t^2$, and the exact solution values at the time $t = 1$ are $[[0., 0.], [.25, .06250], [.50, .25000], [.75, .56250], [1., 1.]]$. The approximate solution is close to the exact solution.

If the last argument in $NumHeatBackw$ is replaced by the number $1/2$, the *Crank-Nicolson scheme* is used. Then, the solution values at $t = 1$ as given by $PList$ agree closely with the exact solution values. When the last argument in $NumHeatBackw$ is replaced by the number 1, the *implicit backward scheme* is used. Then, the solution values at $t = 1$ as given by $PList$ are $[[0, .000203], [.25, .06271], [.50, .25022], [.75, .56272], [1, 1.0002]]$. The agreement with the exact result is not as good as for the Crank-Nicolson scheme. The use of $NumHeatBackw$ with the last argument equal to $1/4$ invokes a stable implicit θ scheme whose stability condition is $r \leq 1$ and whose stability parameter is $r = 0.016$. The solution values at $t = 1$ as given by *PList* are $[[0, -.000102], [.25, .06240], [.50, .24989], [.75, .56239], [1, .99990]]$. Finally, the use of $NumHeatBackw$ with the last argument equal to $3/4$ invokes an unconditionally stable implicit θ scheme. The solution values at $t = 1$ as given by $PList$ are $[[0, .000102], [.25, .0626], [.5, .25011], [.75, .5626], [1, 1.0001]]$.

It is apparent that implicit forms of the θ schemes yield improved numerical results, but the choice of the parameter θ plays a significant role. Plots of the finite difference solutions can be obtained as before.

Method of Lines for the One-Dimensional Heat Equation

We consider initial and boundary value problems for the nonhomogeneous heat or diffusion equation (11.2.1) and replace $u_{xx}(x, t)$ by a *centered second order difference quotient* in x. We retain t as a continuous variable and the term $u_t(x, t)$ is not discretized. This yields

$$\frac{\partial u(x,t)}{\partial t} = \frac{c^2\left(u\left(x+h,t\right) - 2\,u\left(x,t\right) + u\left(x-h,t\right)\right)}{h^2} + F(x,t). \qquad (11.2.42)$$

The truncation error for this approximation (11.2.1) is clearly $O(h^2)$, so that the *semidiscrete difference scheme* (11.2.42) is consistent with (11.2.1).

We restrict our discussion to a finite interval $0 \leq x \leq l$. The initial condition is given at $t = 0$ and the boundary conditions at $x = 0$ and $x = l$. The interval is divided into n equal parts. Since t is not discretized, the grid comprises the x values at which the solution is to be found, and is given as $x_i = il/n$ with $i = 0, 1, 2, \ldots, n$. The x step size or increment is $h = l/n$. The approximate solution is to be determined at a

specified time $t > 0$. The scheme (11.2.42) can be written in the form

$$\frac{du_i(t)}{dt} = \frac{c^2}{h^2}\left(u_{i+1}(t) - 2u_i(t) + u_{i-1}(t)\right) + F_i(t), \tag{11.2.43}$$

where $u_i(t)$ approximates the exact solution $u(x,t)$ of the problem at the point $x_i = il/n$ at the time t, and where $F_i(t) = F(il/n, t)$.

The *initial condition* $u(x,0) = f(x)$ for the problem yields $u_i(0) = f(il/n)$ for $i = 0, 1, 2, \ldots, n$. The boundary lines $x = 0$ and $x = l$ correspond to the values $i = 0$ and $i = n$. If *Dirichlet conditions* are prescribed, then $u_0(t)$ and $u_n(t)$ are known for all t. Then (11.2.43) is a system of $n - 1$ coupled first order ODEs for $u_i(t)$ at $x_i = il/n$ with $i = 1, 2, \ldots, n-1$.

If *Neumann or Robin boundary conditions* are prescribed, an x-derivative term occurs in the boundary conditions, and it must be replaced by a difference quotient if the difference scheme is to be solvable. This is done exactly as before, when we dealt with explicit and implicit difference schemes. For those cases, (11.2.43) is a system of $n + 1$ coupled first order differential equations for $u_i(t)$ at the values of $x_i = il/n$ with $i = 0, 1, 2, \ldots, n-1, n$. The approximate solution is then found on the boundary as well. If there are mixed boundary conditions, with Dirichlet conditions at $x = 0$, say, the system contains n differential equations for $u_i(t)$ at $x_i = il/n$ with $i = 1, 2, \ldots, n$.

This scheme is known as the *method of lines* because the solution at each grid point x_i is determined along the line $x = x_i$ with $t > 0$ in the (x,t)-plane, subject to initial conditions given at $t = 0$. Ideally, one would like to solve this initial value problem for (11.2.43) exactly, retaining t as a continuous variable. In general, this is not feasible and the system must be solved numerically. As a result, the time variable t is discretized. Thereby, the method of lines is converted to a fully discretized difference scheme, and stability questions for this scheme must be addressed. These are resolved by employing stability theories for systems of ODEs, in both the continuous and discrete cases.

For the case of Dirichlet boundary conditions with $u(0,t) = g(t)$ and $u(l,t) = s(t)$, (11.2.43) for the $u_i(t)$ is represented in matrix form as $d\mathbf{u}(t)/dt = A\mathbf{u}(t) + \mathbf{w}(t)$, where

$$A = \frac{c^2}{h^2}\begin{bmatrix} -2 & 1 & & & \mathbf{0} \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & & & 1 \\ \mathbf{0} & & & 1 & -2 \end{bmatrix}, \tag{11.2.44}$$

$$\mathbf{u}(t) = \begin{bmatrix} u_1(t) \\ u_2(t) \\ \vdots \\ u_{n-1}(t) \end{bmatrix}, \qquad \mathbf{w} = \begin{bmatrix} F_1(t) + c^2 g(t)/h^2 \\ F_2(t) \\ \vdots \\ F_{n-1}(t) + c^2 s(t)/h^2 \end{bmatrix}. \tag{11.2.45}$$

The matrix A in (11.2.44) is in *tridiagonal form*. The eigenvalues of the $(n-1) \times (n-1)$ matrix A are given as $\lambda_j = 2\cos(j\pi/n) - 2$, $j = 1, 2, \ldots, n-1$, so that they are all negative. This implies the stability of the system of ODEs. If the time variable is discretized, further analysis is required, and the results depend on the method of discretization that is employed.

Maple's built-in *dsolve* procedure can solve initial value problems for systems of ODEs exactly and numerically. The procedure $NumHeatLines$ that we have created for solving initial and boundary value problems for (11.2.1) with general boundary conditions using the *method of lines* uses *dsolve*. But rather than ask Maple to try to find an exact solution of the system of ODEs generated by the method of lines, the procedure invokes Maple's default numerical method $rkf45$, the *Runge-Kutta-Fehlberg method*. The system is solved by Maple to a high level of accuracy, if possible.

An optional argument may be entered that forces Maple to use a number of other numerical methods for solving ODEs that are available within Maple, such as Euler's method. The step sizes can be controlled by the user when these methods are used. The global variable $SysODE$ exhibits the full system of ODEs together with the initial conditions. These can then be solved explicitly using *dsolve* if so desired and if possible. When derivative boundary conditions are given at the endpoints, the derivatives are expressed by the procedure as difference quotients in the manner presented above, and the solution is determined at these points as well.

$NumHeatLines$ is given as $NumHeatLines(c^2, F(x,t), t = t_0, f(x), x = a..b, bcl, g(t), bcr, s(t), n)$. The first 10 arguments in the procedure are exactly as described above for $NumHeatForw$, the only difference being that the third argument $t = t_0$ specifies only the initial time and not the final time. Boundary conditions of all three types can be treated by the procedure. The output is given as a Maple procedure *proc(rkf45_x)... end proc*. The global variable SH can be used to exhibit the numerical solution values $u_i(t_f)$ (at the time $t = t_f$) at all the x grid points, by invoking $SH(t_f)$.

If an optional eleventh argument $t = t_f$ is added in the procedure, the numerical solution $u_i(t_f)$ at each x grid point is exhibited in tabular form. The global variable $PList$ can also be employed to display a list of these solution values. If the optional eleventh argument is *type = numeric*, a number of classical numerical solution methods for ODEs such as the *forward Euler, Heun, lower-order Runge-Kutta*, or *predictor-corrector methods* can be invoked for the solution of the system of ODEs. The step sizes used in these methods can be controlled, as can the form of the output. (To accomplish this, it is necessary to include additional optional arguments.) As was the case before, the procedure contains global variables that can be used to plot the finite difference solutions as curves or surfaces.

Example 11.6. A Problem with Robin Boundary Conditions. The initial and mixed boundary value problem for the nonhomogeneous heat equation considered in Example 11.3 has the exact solution $u(x,t) = x^2t^2$ as was indicated above. We apply the *method of lines* to a problem with Robin boundary conditions that yield the same solution and obtain the numerical solution at $t = 1$.

On entering $NumHeatLines(1, 2tx^2 - 2t^2, t = 0, 0, x = 0..1, 1, 0, 1, 3t^2, 4)$, the output exhibits the PDE (11.2.10), the initial condition $u(x, 0) = 0$, and the Robin boundary conditions $-u_x(0, t) + u(0, t) = 0$, $u_x(1, t) + u(1, t) = 3t^2$. The interval $0 < x < 1$ for the problem is subdivided into four equal parts, so that the x increment $h = 1/4$. The final output of the $NumHeatLines$ procedure is the Maple procedure *proc(rkf45_x)... end proc*. The global variable $SH(t)$ exhibits the numerical solution values at the time t.

If an eleventh argument $t = 1$ is added to the procedure, the numerical solution at $t = 1$ is exhibited in tabular form as

$$\begin{bmatrix} x & u(x,1) \\ 0.0 & 0.41725\,10^{-8} \\ 0.25 & 0.06250 \\ 0.50 & 0.25000 \\ 0.75 & 0.56250 \\ 1.0 & 1.00000 \end{bmatrix}. \tag{11.2.46}$$

PList yields the same values: $[[0, 0.41725\ 10^{-8}], [.25, .0625], [.5, .25], [.75, .5625], [1, 1]]$. Alternatively, the numerical solution at $t = 1$ is given by $SH(1)$ as $[t = 1., u_0(t) = .41725\,10^{-8}, u_1(t) = .0625, u_2(t) = .25, u_3(t) = .5625, u_4(t) = 1.]$. $SH(0)$ yields $[t = 0, u_0(t) = 0, u_1(t) = 0, u_2(t) = 0, u_3(t) = 0, u_4(t) = 0]$, which represents the (exact) initial values at the points [0, 1/4, 1/2. 3/4, 1].

The numerical solution at $t = 1$ can be plotted by invoking $plot(PList)$. The global variable $PList3d$ modifies the output of $PList$ given above, by inserting a second argument 1, (it corresponds to $t = 1$), in each element of the list. (In the general case, the value $t = t_f$ of the procedure is entered in each term of $PList$.) To obtain a time dependent plot of the finite difference solution, we consider a sequence of values t_j that extend from t_0 to t_f. Typically, we put $t_j = t_0 + (t_f - t_0)j/m$, $j = 0, \ldots, m$ for some positive integer m. (In this example, $t_0 = 0$ and $t_f = 1$.) For each t_j we carry out the procedure $NumHeatLines$, with the eleventh argument given as $t = t_j$. Then we assign the name $PList3d(j)$ to the $PList3d$ output of each procedure in the sequence. Finally, the procedure *surfdata*($[PList3d(0), \ldots, PList3d(m)]$) yields a time-dependent plot of the finite difference solution, as before.

The global variable $SysODE$ exhibits the coupled system of ODEs (11.2.43) appropriate for this problem together with the initial conditions. It has the form $\{u_0'(t) = 32u_1(t) - 40u_0(t) - 2t^2,\ u_1'(t) = 16u_0(t) - 32u_1(t) + 16u_2(t) + 3t/25 - 2t^2, u_2'(t) = 16u_1(t) - 32u_2(t) + 16u_3(t) + t/2 - 2t^2,\ u_4'(t) = 32u_3(t) - 40u_4(t) + 22t^2 + 2t, u_3'(t) = 16u_2(t) - 32u_3(t) + 16u_4(t) + 11t/10 - 2t^2,\ u_0(0) = 0, u_1(0) = 0, u_2(0) = 0, u_3(0) = 0, u_4(0) = 0\}$. Then, *dsolve* applied to this system, yields the exact solution $u_0(t) = 0, u_1(t) = t^2/16, u_2(t) = t^2/4, u_3(t) = 9t^2/16, u_4(t) = t^2$.

If we replace the eleventh argument $t = 1$ by $type = numeric, method = classical[foreuler], stepsize = .00005, output = array([1])$, we obtain

$$\left[\begin{array}{c} [t,\ u_0(t),\ u_1(t),\ u_2(t),\ u_3(t),\ u_4(t)] \\ [\ 1.0 \quad -.55357\,10^{-5} \quad 0.06249 \quad 0.24999 \quad 0.56249 \quad 0.99999\ ] \end{array}\right]. \tag{11.2.47}$$

Maple uses the *forward Euler method* to approximate the solution of the system (11.2.43). The t step size is $k = 0.0005$ and the solution is given at $t = 1$. If the eleventh argument is $type = numeric, method = classical[rk4], stepsize = .001, output = array([1])$, we obtain

$$\left[\begin{array}{c} [t,\ u_0(t),\ u_1(t),\ u_2(t),\ u_3(t),\ u_4(t)] \\ [\ 1.0 \quad -.34342\,10^{-12} \quad 0.06250 \quad 0.25000 \quad 0.56250 \quad 1.0000\ ] \end{array}\right]. \tag{11.2.48}$$

Maple uses the *fourth order Runge-Kutta method* to approximate the solution of the system (11.2.43). The step size in t is $k = 0.001$ and the solution is given at $t = 1$. Even though the step size is considerably larger than for the Euler method, the results are much better.

Exercises 11.2

11.2.1. Use *mtaylor* to verify the truncation error for the forward difference scheme given in (11.2.4).

11.2.2. Show that if $k = h^2/6c^2$, the truncation error in (11.2.4) is $O(k^2 + h^4)$.

11.2.3. Apply the procedure $NumHeatForw(c^2, F(x,t), t = t_0..t_f, f(x), x = a..b, bcl, g(t), bcr, s(t), n, k)$ in the general case with the following boundary conditions at the endpoints. (a) Dirichlet conditions; (b) Neumann conditions; (c) Robin boundary conditions; (d) Mixed boundary conditions: Dirichlet conditions at $x = a$ and Neumann conditions at $x = b$.

11.2.4. Use *pdetest* to verify that $u(x,t) = \sin(\pi x)\exp(-\pi^2 t)$ is the exact solution of the initial and boundary value problem (11.2.6)–(11.2.7).

11.2.5. Use the procedure $NumHeatForw$ to solve the initial and boundary value problem (11.2.6)–(11.2.7) subject to the conditions given in (11.2.8). Generate the plot displayed in Figure 11.2 using the Maple procedures referred to in Example 11.2.

11.2.6. Carry out the procedure $NumHeatForw$ to solve the problem of Example 11.2 with the increments chosen so that the stability parameter $r = 1/6$.

11.2.7. Choose values of h and k for the problem of Example 11.2 so that instability results and display the output given by $NumHeatForw$.

11.2.8. Reproduce the results of Example 11.3.

11.2.9. Use *mtaylor* to verify the truncation error for the backward difference scheme given in (11.2.17).

11.2.10. Verify the truncation relation of the text for the DuFort-Frankel difference scheme.

11.2.11. Use the procedure *NumHeatBackw* to establish the results obtained in Example 11.4.

11.2.12. Solve the mixed problem of Example 11.5 using *NumHeatBackw* with the five θ values given in the example. Plot the five results.

11.2.13. Reproduce the results given in Example 11.6 using *NumHeatLines*.

11.2.14. Solve the problem of Example 11.6 using *NumHeatLines* if the Robin boundary conditions are replaced by Dirichlet conditions that yield the same exact solution.

11.2.15. Solve the problem of Example 11.6 using *NumHeatLines* if the Robin boundary conditions are replaced by homogeneous Neumann conditions.

11.3 FINITE DIFFERENCE METHODS FOR THE ONE-DIMENSIONAL WAVE EQUATION

The one-dimensional nonhomogeneous wave equation is given as

$$\frac{\partial^2 u(x,t)}{\partial t^2} - c^2 \frac{\partial^2 u(x,t)}{\partial x^2} = F(x,t), \tag{11.3.1}$$

with $c > 0$. Let $h > 0$ be an increment in x and $k > 0$ be an increment in t. As was done for the heat equation, we replace the partial derivatives by various (approximating) difference quotients as introduced in Section 11.1, thereby creating a number of difference schemes. They can be used to solve initial and boundary value problems for (11.3.1), and a number of them are presented below. The initial value problem for the homogeneous and nonhomogeneous wave equation has been shown to have an explicit solution, so that there is no need to consider difference methods for its approximate solution.

Explicit Forward Difference Method for the One-Dimensional Wave Equation

We begin our construction of difference schemes for (11.3.1) by approximating $u_{tt}(x,t)$ and $u_{xx}(x,t)$ by *centered second order difference quotients* in t and x, respectively. This yields

$$\frac{u(x,t+k) - 2u(x,t) + u(x,t-k)}{k^2}$$

$$= \frac{c^2\left(u\left(x+h,t\right) - 2\,u\left(x,t\right) + u\left(x-h,t\right)\right)}{h^2} + F(x,t). \tag{11.3.2}$$

We put $r = ck/h$ and express (11.3.2) as the *difference scheme*

$$u(x,t+k) = 2(1-r^2)u(x,t)+r^2u(x+h,t)+r^2u(x-h,t)-u(x,t-k)+k^2F(x,t). \tag{11.3.3}$$

As demonstrated below, r represents a *stability parameter* for the difference scheme (11.3.3). The scheme (11.3.3) shows that if the values of the solution $u(x,t)$ are known at the time t at the points $x-h$, x, $x+h$, and at the time $t-k$ at the point x, they can be determined at the point x at the time $t+k$. Solution values must be known at two previous times t and $t-k$ to advance the solution to the time $t+k$, so that this is a *two-step method*. The computational stencil is exhibited in Figure 11.4.

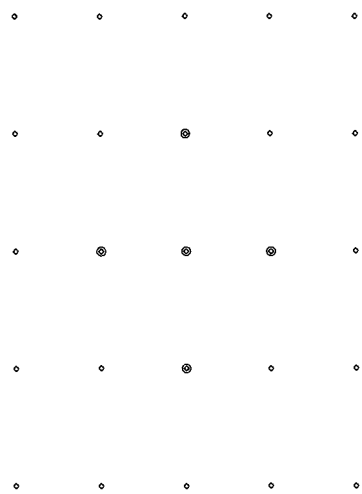

Figure 11.4 The computational stencil.

This immediately poses the following problem. Suppose that initial conditions for (11.3.1) are prescribed at $t = 0$ as $u(x,0) = f(x)$ and $u_t(x,0) = g(x)$. To find the values of $u(x,2k)$ at all points $x = ih$, $i = 0, \pm1, \pm2, \ldots$, we must know the values of $u(x,t)$ at $t = k$ and at $t = 0$. However, the initial conditions supply the values of $u(x,0)$ but not $u(x,k)$. To resolve this problem and determine the $u(x,k)$ values, we replace the derivative $u_t(x,0)$ by a difference quotient. The easiest approach is to replace the derivative by a forward difference quotient, so that $u_t(x,0) = g(x)$ will give $u(x,k)$ in terms of $u(x,0)$, whose values are known. This approximation is rejected because it is only first order accurate, while the centered difference approximations used for (11.3.1) are second order accurate. As a result, we use a centered difference quotient to approximate $u_t(x,0)$ and obtain $u(x,k) = u(x,-k) + 2kg(x)$, which gives $u(x,k)$ in terms of $u(x,-k)$ whose value is not known. However, the difference equation (11.3.3) also yields an (approximate) value for $u(x,k)$. Putting $t = 0$ in (11.3.3) gives $u(x,-k) = 2(1-r^2)u(x,0) + r^2u(x+h,0) + r^2u(x-h,0) - u(x,k) + k^2F(x,0)$. From the initial condition $u(x,0) = f(x)$ and $u(x,k) = u(x,-k) + 2kg(x)$, we obtain the required result $u(x,k) = (1-r^2)f(x) + r^2f(x+h)/2 + r^2f(x-h)/2 + k^2F(x,0)/2 + kg(x)$.

Next we determine the *truncation error* that results on replacing the wave equation (11.3.1) by the difference equation (11.3.2). Applying the Maple procedure *mtaylor* to (11.3.2) gives

$$u_{tt}(x,t) - c^2 u_{xx}(x,t) - F(x,t) \approx -\frac{1}{12}u_{tttt}k^2 + \frac{c^2}{12}u_{xxxx}h^2. \qquad (11.3.4)$$

Thus, the difference between the wave equation (11.3.1) and the difference equation (11.3.2) is of $O(k^2 + h^2)$, and this is the order of the *truncation error*. As $h, k \to 0$ the truncation error tends to zero, so that the difference approximation (11.3.2) is *consistent* with the wave equation (11.3.1).

We use the difference scheme to solve initial and boundary value problems for (11.3.1) over the finite interval $0 < x < l$. The initial conditions are prescribed at $t = 0$ and were given above. The boundary conditions are given at $x = 0$ and $x = l$. We divide the interval into n equal parts. The grid in (x,t)-space, on which the solution is to be found, is given as $(x_i, t_j) = (il/n, jk)$ with $i = 0, 1, 2, \ldots, n$ and $j = 0, 1, 2, \ldots, m$. The x step size is $h = l/n$ and the time increment is k. The solution is to be found at the time $t_m = mk$. The difference scheme (11.3.3) can then be written in the form

$$u_{i,j+1} = 2(1 - r^2)u_{i,j} + r^2 u_{i+1,j} + r^2 u_{i-1,j} - u_{i,j-1} + k^2 F_{i,j}, \qquad (11.3.5)$$

where $u_{i,j}$ represents the approximation to the exact solution $u(x,t)$ of the given problem at the point $u(x_i, t_j) = u(il/n, jk)$ and $F_{i,j} = F(il/n, jk)$. This is a two-step *explicit forward difference scheme.*

With the initial conditions $u(x,0) = f(x)$ and $u_t(x,0) = g(x)$ we have $u_{i,0} = f(il/n)$ and (as shown above) $u_{i,1} = (1-r^2)f_i + r^2 f_{i+1}/2 + r^2 f_{i-1}/2 + k^2 F_{i,0}/2 + kg_i$ for $i = 0, 1, 2, \ldots, n$, with $f_i = f(il/n)$ and $g_i = g(il/n)$. The boundary lines $x = 0$ and $x = l$ correspond to the values $i = 0$ and $i = n$. The range of values $x_i = il/n$ at which the solution of the difference scheme is found depends, as before, on the type of boundary condition assigned at the endpoints. The determination of the x grid and the relevant boundary values is identical to that given for the heat equation and is not repeated here. We note that if a Neumann or Robin boundary condition is given at an endpoint, the values of f_{-1} and/or f_{n+1} require adjustment via the use of ghost points.

The forward difference scheme (11.3.5) is *stable* if $r = ck/h \leq 1$ and *unstable* if $r = ck/h > 1$, as we demonstrate below. The stability condition can be given the following geometric interpretation. It was shown in Section 4.5 that the *domain of dependence* of the (exact) solution of the wave equation (11.3.1) at a grid point $(il/n, jk)$ is a triangle with one vertex at the given point and the other two vertices at the points $(il/n - cjk, 0)$ and $(il/n + cjk, 0)$ on the x-axis. The corresponding *domain of dependence* of the (numerical) solution of the difference scheme (11.3.5) at a point $(il/n, jk)$ is readily found to be a triangle with one vertex at the given point and the other two vertices at the points $(il/n - hj, 0)$ and $(il/n + hj, 0)$ on the x-axis. We see that the interval $il/n - hj \leq x \leq il/n + hj$ on the x-axis contains or is identical with the interval $il/n - cjk \leq x \leq il/n + cjk$ if $ck \leq h$ or $r \leq 1$. This

signifies that the numerical domain of dependence includes the domain of dependence for the PDE. If $r > 1$, the reverse is true. Thus, the stability condition asserts that the numerical scheme must contain all the data that is required for the determination of the exact solution at the given point; otherwise, the numerical solution will fail to converge to the exact solution. This formulation of the stability condition was given by *Courant, Friedrichs,* and *Lewy* in the paper on difference methods for PDEs cited in Section 1.3 in connection with random walk problems. For this reason the stability condition $r \leq 1$ is often referred to as the *CFL stability condition* or the *Courant stability condition.*

We have created the procedure $NumWaveForw$ to automate the process of carrying out the *explicit forward difference scheme* for approximating solutions of initial and boundary value problems for the wave equation. It is given in the form $NumWaveForw(c, F(x,t), t = t_0..t_f, f(x), g(x), x = a..b, bcl, l(t), bcr, s(t), n, k)$. Apart from the addition of a fifth argument $g(x)$, which specifies the initial condition $u_t(x, t_0) = g(x)$, the arguments are defined as in the procedure $NumHeatForw$ presented above. [The first argument c is the speed of wave propagation in the wave equation (11.3.1).] The output of the procedure has the general form of that for $NumHeatForw$.

Example 11.7. A Problem with a Variety of Boundary Conditions. We consider the nonhomogeneous wave equation

$$u_{tt}(x,t) - 4u_{xx}(x,t) = 3xt(x^2 - 4t^2)(2\cos(x^3t^3) - 3\sin(x^3t^3)t^3x^3) \quad (11.3.6)$$

over the interval $0 < x < 1$ with the initial conditions $u(x,0) = 0$, $u_t(x,0) = 0$. The function $u(x,t) = \sin(x^3t^3)$ satisfies (11.3.6) and the given initial conditions. Various boundary conditions that the given function satisfies at $x = 0$ and $x = 1$ are introduced, and the resulting problems are solved numerically using the *explicit forward difference scheme.* The x interval is subdivided into 10 equal parts and we apply the procedure $NumWaveForw$ to find the numerical solution at $t = 1$. The exact solution evaluated at the x grid points $i/10$, $i = 0, 1, 2, \ldots, 10$, has the values [0, .001, .008, .027, .06396, .12467, .2143, .33631, .48992, .66612, .84147].

We begin with the *Dirichlet problem* with $u(0,t) = 0$ and $u(1,t) = \sin(t^3)$. $NumWaveForw(2, 3xt(x^2-4t^2)(2\cos(x^3t^3)-3\sin(x^3t^3)t^3x^3), t = 0..1, 0, 0, x = 0..1., dirichlet, 0, dirichlet, \sin(t^3), 10, .1)$ determines the numerical solution at $t = 1$. The stability parameter r has the value $r = 4$, so that the difference scheme is *unstable.* The instability is confirmed by the output of *PList,* [[0, 0], [.1, −.00283], [.2, 146.8], [.3, -1476.5], [.4, 72044], [.5, -22328],[.6, 48397], [.7, -75784], [.8, 83942], [.9, -56542], [1, .84147]]. If we replace the last argument in the procedure for the Dirichlet problem (i.e., the t increment, by 0.01), we have $r = 0.04$, so that the difference scheme is *stable.*

For the *Neumann problem,* the boundary values are given as $u_x(0,t) = 0$ and $u_x(1,t) = 3t^3\cos(t^3)$. $NumWaveForw(2, 3xt(x^2-4t^2)(2\cos(x^3t^3)-3\sin(x^3t^3)$ $t^3x^3), t = 0..1, 0, 0, x = 0..1., 0, 0, 0, 3t^3\cos(t^3), 10, .001)$ finds the numerical solution at $t = 1$ with the stability parameter $r = 0.0004$.

Finally, we consider the *mixed problem* with the boundary conditions $u_x(0,t) = 0$ and $u_x(1,t) + u(1,t) = 3t^3\cos(t^3) + \sin(t^3)$. $NumWaveForw(2, 3xt(x^2 - 4t^2)(2\cos(x^3t^3) - 3\sin(x^3t^3)t^3x^3), t = 0..1, 0, 0, x = 0..1., 0, 0, 1, 3t^3\cos(t^3) + \sin(t^3), 10, .001)$ finds the numerical solution at $t = 1$ with the stability parameter $r = 0.0004$.

The global variable *PList* lists the solution values at the grid points for each of the three foregoing procedures as *Dirichlet* =[[0., 0.], [.1, .00093], [.2, .0078], [.3, .02666], [.4, .06335],[.5, .12367],[.6, .2182], [.7, .3342], [.8, .4875], [.9, .66416], [1., .84147]]; *Neumann* =[[0.0, .001], [.1, .00099], [.2, .0068], [.3, .02448], [.4, .06], [.5, .119],[.6, .208], [.7, .329], [.8, .483], [.9, .662], [1., .8437]]; *Mixed* =[[0., .0018], [.1, .0018], [.2, .0077], [.3, .02562], [.4, .061458],[.5, .1211],[.6, .20987], [.7, .3315], [.8, .48590], [.9, .66496], [1., .84648]].

The piecewise linear curves determined by the foregoing point lists can be plotted by invoking $plot(PList)$ after applying $NumWaveForw$ for each case. A time-dependent plot that yields an approximate finite difference solution surface for $0 \le x \le 1$, $0 \le t \le 1$ in each case can be generated from the output of global variable $PlotList3dMod$. The exact and approximate solutions can be plotted by entering $P1 = plot3d(\sin(x^3t^3), x = 0..1, t = 0..1)$. $P2 = plots[surfdata](PList3dMod)$, and $plots[display](\{P1, P2\})$.

Implicit Backward Difference Methods for the One-Dimensional Wave Equation

As for the heat equation, the introduction of backward implicit difference schemes for the solution of initial and boundary value problems for (11.3.1) has the effect of removing the stability restriction found for the forward difference scheme. We do this by introducing the ω *difference scheme*, whose construction parallels that used for the θ *difference scheme* for (11.2.1).

We again approximate $u_{tt}(x,t)$ in (11.3.1) by a centered difference quotient but approximate $u_{xx}(x,t)$ by a weighted average of centered difference quotients evaluated at $t - k$, t and $t + k$, with the respective *nonnegative weights* ω, $1 - 2\omega$, and ω. The nonhomogeneous term $F(x,t)$ is also replaced by a corresponding weighted average. (We must have $0 \le \omega \le 1/2$ for the weights to be nonnegative.) It is readily shown that the resulting difference approximation is *consistent* with (11.3.1) with a *truncation error* of $O(k^2 + h^2)$. The ω *difference scheme* has the form

$$\begin{aligned} u_{i,j+1} = {} & 2u_{i,j} - u_{i,j-1} - 2r^2\omega u_{i,j+1} + r^2\omega u_{i+1,j+1} + r^2\omega u_{i-1,j+1} + k^2\omega F_{i,j+1} \\ & - 2r^2(1-2\omega)u_{i,j} + r^2(1-2\omega)u_{i+1,j} + r^2(1-2\omega)u_{i-1,j} + k^2(1-2\omega)F_{i,j} \\ & - 2r^2\omega u_{i,j-1} + r^2\omega u_{i+1,j-1} + r^2\omega u_{i-1,j-1} + k^2\omega F_{i,j-1}. \quad (11.3.7) \end{aligned}$$

We have retained the notation used for the explicit difference scheme, and the initial and boundary conditions are dealt with as before. When $\omega = 0$, the ω *difference*

scheme reduces to the *explicit forward scheme* presented above. For all other admissible values of ω, the scheme is *implicit*. The case with $\omega = 1/2$ is referred to as the *implicit backward scheme*. As will be shown, if $0 \leq \omega < 1/4$ the stability condition for the ω *difference scheme* is $r^2 \leq 1/(1-4\omega)$. If $1/4 \leq \omega \leq 1/2$, the scheme is unconditionally stable.

For the case of *Dirichlet boundary conditions* with $u(0,t) = l(t)$ and $u(l,t) = s(t)$, the system of linear equations (11.3.7) for the $u_{i,j}$ is represented in matrix form as

$$A\mathbf{u}_{j+1} = B\mathbf{u}_j + A\mathbf{u}_{j-1} + \mathbf{w}_{(j+1,j,j-1)}, \qquad j \geq 1, \tag{11.3.8}$$

where

$$A = \begin{bmatrix} 1+2r^2\omega & -r^2\omega & & & \mathbf{0} \\ -r^2\omega & 1+2r^2\omega & -r^2\omega & & \\ & \ddots & \ddots & \ddots & \\ & & & & -r^2\omega \\ \mathbf{0} & & & -r^2\omega & 1+2r^2\omega \end{bmatrix}, \quad \mathbf{u}_j = \begin{bmatrix} u_{1,j} \\ u_{2,j} \\ \vdots \\ u_{n-1,j} \end{bmatrix}, \tag{11.3.9}$$

$$B = \begin{bmatrix} 2r^2(1-2\omega)-2 & -r^2(1-2\omega) & & & \mathbf{0} \\ -r^2(1-2\omega) & 2r^2(1-2\omega)-2 & & & \\ & \ddots & \ddots & \ddots & \\ & & & & -r^2(1-2\omega) \\ \mathbf{0} & & & -r^2(1-2\omega) & 2r^2(1-2\omega)-2 \end{bmatrix}, \tag{11.3.10}$$

$$\mathbf{w}_{(j+1,j,j-1)} = \tag{11.3.11}$$

$$\begin{bmatrix} \omega(k^2F_{1,j+1} + k^2F_{1,j-1} + r^2l_{j+1} + r^2l_{j-1}) + (1-2\omega)(k^2F_{1,j} + r^2l_j) \\ \omega(k^2F_{2,j+1} + k^2F_{2,j-1}) + (1-2\omega)(k^2F_{2,j}) \\ \vdots \\ \omega(k^2F_{n-1,j+1} + k^2F_{n-1,j-1} + r^2s_{j+1} + r^2s_{j-1}) + (1-2\omega)(k^2F_{n-1,j} + r^2s_j) \end{bmatrix},$$

with $l_j = l(jk)$ and $s_j = s(jk)$. (We do not exhibit the matrix forms that correspond to other boundary conditions.) The coefficient matrix of A of $\mathbf{u}_{j+1}$ in (11.3.8) is tridiagonal and diagonally dominant. As a result, A is invertible and (11.3.8) has a unique solution. In addition, if $\omega = 0$, the matrix A reduces to the identity matrix and the ω scheme becomes an explicit scheme. For all other admissible values of ω, the scheme is implicit. If A^{-1} is the inverse of the matrix A, we can write (11.3.8) as

$$\mathbf{u}_{j+1} = A^{-1}B\mathbf{u}_j + \mathbf{u}_{j-1} + A^{-1}\mathbf{w}_{(j+1,j,j-1)}, \qquad j \geq 1. \tag{11.3.12}$$

The procedure $NumWaveBackw$ carries out the ω *difference scheme* for approximating solutions of initial and boundary value problems for the wave equation. It is given as $NumWaveBackw(c, F(x,t), t = t_0..t_f, f(x), g(x), x = a..b, bcl, l(t), bcr, s(t), n, k, \omega)$. The first 12 arguments are as in $NumWaveForw$ given above. The thirteenth argument specifies the value of the parameter ω. If $\omega = 0$, the scheme is explicit. If $0 < \omega \leq 1/2$, the scheme is implicit. The output of the procedure has the form of that for $NumWaveForw$. Plots of the results can be generated by using the global variables $PList$ and $PList3dMod$, as was done for the explicit forward scheme.

Example 11.8. A Problem with a Variety of Boundary Conditions. We reconsider the initial and boundary value problem for the nonhomogeneous wave equation (11.3.6) over the interval $0 < x < 1$ with the initial conditions $u(x,0) = 0$ and $u_t(x,0) = 0$. The function $u(x,t) = \sin(x^3t^3)$ satisfies (11.3.6) and the given initial conditions. Various boundary conditions that the given function satisfies at $x = 0$ and $x = 1$ are introduced, and the resulting problems are solved numerically using the ω *difference scheme*. The x interval is subdivided into 10 equal parts and we apply the procedure $NumWaveBackw$ to find the numerical solution at $t = 1$. We recall that the exact solution evaluated at the x grid points $i/10$, $i = 0, 1, 2, \ldots, 10$ has the values [0, .001, .008, .027, .06396, .12467, .21432, .33631, .48992, .66612, .84147].

We begin with the *Dirichlet problem*. The boundary values are $u(0,t) = 0$ and $u(1,t) = \sin(t^3)$. $NumWaveBackw(2, 3xt(x^2 - 4t^2)(2\cos(x^3t^3) - 3\sin(x^3t^3)t^3x^3), t = 0..1, 0, 0, x = 0..1., dirichlet, 0, dirichlet, \sin(t^3), 10, .01, 0)$ determines the numerical solution at $t = 1$. The last argument states that $\omega = 0$, so that this corresponds to the explicit difference scheme. The stability parameter $r = 0.04$ so that the scheme is stable. The global variable $PList$ lists the solution values at the grid points as, [[0, 0], [.1, .00095], [.2, .008], [.3, .02743], [.4, .06523],[.5, .12731], [.6, .219], [.7, .34378], [.8, .50075], [.9, .68], [1., .84147]].

On replacing the penultimate argument in the procedure, the t increment, by 0.1, we find that the stability parameter $r = 4$. Then, if the last argument is replaced by $1/4$, the scheme is implicit and unconditionally stable. The output of *PList* is [[0, 0], [.1, .0008], [.2, .0096], [.3, .0342], [.4, .0823],[.5, .1615], [.6, .278], [.7, .43451], [.8, .62296], [.90, .81989], [1.0, .84147]]. The results are not as good as those found above but the wild instability exhibited by *NumWaveForw* in Example 11.7 for the same value of the stability parameter is not present here. For the *Neumann problem*, the boundary values are $u_x(0,t) = 0$ and $u_x(1,t) = 3t^3\cos(t^3)$. Then $NumWaveBackw(2, 3xt(x^2 - 4t^2)(2\cos(x^3t^3) - 3\sin(x^3t^3)t^3x^3), t = 0..1, 0, 0, x = 0..1., 0, 0, 0, 3t^3\cos(t^3), 10, 0.01, 1/2)$ finds the numerical solution at $t = 1$ with $r = .04$. The global variable *PList* lists the solution values at the grid points as [[0, 0.0044], [.1, .0045], [.2, .001], [.3, .029], [.4, .06667], [.5, .12854],[.6, .22046], [.7, .3461], [.8, .5051], [.9, .6881], [1, .87086]].

Finally, we examine the *mixed problem* with boundary conditions $u_x(0,t) = 0$ and $u_x(1,t) + u(1,t) = 3t^3\cos(t^3) + \sin(t^3)$. $NumWaveBackw(2, 3xt(x^2 - 4t^2)(2\cos(x^3t^3) - \sin(x^3t^3)t^3x^3), t = 0..1, 0, 0, x = 0..1., 0, 0, 0, 3t^3\cos(t^3) + \sin(t^3), 10, .001, .35)$ finds the numerical solution at $t = 1$ with $r = 0.0004$. The

scheme is implicit. The output of $PList$ is [[0, 0.00180], [.1, .00179], [.2, .00773], [.3, .0256], [.4, .0615],[.5, .121],[.6, .2099], [.7, .332], [.8, .486], [.9, .665], [1, .847]].

Method of Lines for the One-Dimensional Wave Equation

To construct semi-discrete numerical approximations to the solutions of initial and boundary value problems for the nonhomogeneous wave equation (11.3.1), we replace $u_{xx}(x,t)$ by a *centered second order difference quotient* in x but retain t as a continuous variable, so that $u_{tt}(x,t)$ is not discretized. Then

$$\frac{\partial^2 u(x,t)}{\partial t^2} = \frac{c^2\left(u\left(x+h,t\right) - 2\,u\left(x,t\right) + u\left(x-h,t\right)\right)}{h^2} + F(x,t) \qquad (11.3.13)$$

and gives rise to the *method of lines* for the wave equation. The truncation error for this approximation to the wave equation is clearly $O(h^2)$, so that the *semidiscrete difference scheme* (11.3.13) is consistent with the wave equation.

As was done for the heat equation, we restrict our discussion to a finite interval given as $0 \le x \le l$. The initial conditions are given at $t = 0$ and the boundary conditions at $x = 0$ and $x = l$. The interval is divided into n equal parts. Since t is not discretized, the grid comprises the x values at which the solution is to be found and is given as $x_i = il/n$ with $i = 0, 1, 2, \ldots, n$. The x step size or increment is $h = l/n$. The approximate solution is to be determined at a specified time $t > 0$. The scheme (11.3.13) can be written as

$$\frac{d^2 u_i(t)}{dt^2} = \frac{c^2}{h^2}\left(u_{i+1}(t) - 2u_i(t) + u_{i-1}(t)\right) + F_i(t), \qquad (11.3.14)$$

where $u_i(t)$ represents the approximation to the exact solution $u(x,t)$ of the problem at the point $x_i = il/n$ at the time t and where $F_i(t) = F(il/n, t)$.

The only difference between the formulation of this method for the wave and heat equations lies in the fact that two *initial conditions* $u(x,0) = f(x)$ and $u_t(x,0) = g(x)$ must be prescribed for the wave equation, whereas only $u(x,0) = f(x)$ is specified for the heat equation. As a result, the initial conditions for the system (11.3.14) are $u_i(0) = f(il/n)$ and $u_i'(0) = g(il/n)$ for $i = 0, 1, 2, \ldots, n$. Because t is not discretized, it is not necessary to approximate $u_t(x,0)$ by a difference quotient, as was the case for the difference schemes considered above. The boundary lines $x = 0$ and $x = l$ correspond to the values $i = 0$ and $i = n$. Boundary conditions of all three kinds and their effect on the number of equations to be solved are dealt with as was done for the heat equation. The discussion is not repeated here.

The resulting approximation scheme is the *method of lines* for the wave equation with $t > 0$ in the (x,t)-plane. Again, it may be possible to solve the initial value problem for the system of ODEs (11.3.14) exactly, retaining t as a continuous variable. But this is not feasible in general and the system must be solved by using numerical methods, such as finite difference schemes. As a result, the time variable t is discretized. Thereby, the method of lines is converted to a fully discretized difference

scheme. Questions of stability for this scheme must also be addressed, and they are resolved by employing stability theories for systems of ODEs.

For the case of *Dirichlet boundary conditions* with $u(0,t) = g(t)$ and $u(l,t) = s(t)$, the system of ODEs (11.3.14) for the $u_i(t)$ has the matrix representation

$$\frac{d^2\mathbf{u}(t)}{dt^2} = A\mathbf{u}(t) + \mathbf{w}(t), \tag{11.3.15}$$

where the matrix A is defined as in (11.2.44), and $\mathbf{u}(t)$ and $\mathbf{w}(t)$ are given as in (11.2.45). The negativity of the eigenvalues of the matrix A implies the stability of the system (11.3.15). If the time variable is discretized, further analysis is required, based on the method of discretization that is employed.

The procedure $NumWaveLines$ solves initial and boundary value problem for the one-dimensional wave equation with general boundary conditions using the *method of lines*. As in the procedure $NumHeatLines$, rather than ask Maple to attempt to find an exact solution of the problem generated by the method of lines, the procedure invokes Maple's default numerical method $rkf45$, the *Runge-Kutta-Fehlberg method*. Maple solves the problem to a high level of accuracy, if possible. An optional argument may be entered that forces Maple to use various other numerical methods for solving ODEs that are available within Maple, such as Euler's method. The step sizes can be controlled by the user when these methods are used. The global variable $SysODE$ exhibits the full system of ODEs together with the initial conditions. These can then be solved explicitly using $dsolve$ if so desired and if possible. When derivative boundary conditions are given at the endpoints, the derivatives are expressed as differences in the manner presented above, and the solution is determined at these points as well.

$NumWaveLines$ is given as $NumWaveLines(c, F(x,t), t = t_0, f(x), g(x), x = a..b, bcl, g(t), bcr, s(t), n)$. The first 11 arguments in the procedure are as described above for $NumWaveForw$. The only difference is that the third argument $t = t_0$ specifies only the initial time. Boundary conditions of all three types can be treated. The output is given as a Maple procedure *proc(*$rkf45\_x$*)... end proc*. Then the global variable SW can be used to exhibit the numerical solution values $u_i(t_f)$ and $u_i'(t_f)$ at all the x grid points, by invoking $SW(t_f)$. If an optional twelfth argument $t = t_f$ is added, the numerical solution $u_i(t_f)$ at each grid point is exhibited in tabular form. The global variable $PList$ can also be used to display a list of these solution values. If an optional twelfth argument *type = numeric* is added, a number of classical numerical solution methods for ODEs such as the *forward Euler, Heun, lower order Runge-Kutta, or predictor-corrector methods* can be invoked. The step sizes used in these methods can be controlled, as can the form of the output. (This is done by adding more optional arguments.) Solution plots can be constructed exactly as was done for $NumHeatLines$.

Example 11.9. A Problem with Robin Boundary Conditions. We consider an initial and mixed boundary value problem for the nonhomogeneous wave equation that has the exact solution $u(x,t) = x^2t^2$. The *method of lines* is applied to a problem with Robin boundary conditions, and the numerical solution at $t = 1$ is found.

If we enter $NumWaveLines(1, 2x^2 - 2t^2, t = 0, 0, 0, x = 0..1, 1, 0, 1, 3t^2, 4)$, the output exhibits the PDE $u_{tt}(x,t) - u_{xx}(x,t) = 2x^2 - 2t^2$, the initial conditions $u(x,0) = 0$, $u_t(x,0) = 0$, and the boundary conditions $-u_x(0,t) + u(0,t) = 0$, $u_x(1,t) + u(1,t) = 3t^2$. The interval $0 < x < 1$ is subdivided into four equal parts so that the x increment is $h = 1/4$. The final output of the procedure is *proc(rkf45_x)... end proc*. Then $SW(t)$ exhibits the numerical solution values at the time t. The numerical solution at $t = 1$ is given by $SW(1)$ as $[t = 1., u_0(t) = .41725 \times 10^{-8}$, $u_0'(t) = .36445 \times 10^{-8}$, $u_1(t) = .0625$, $u_1'(t) = .125, u_2(t) = .25$, $u_2'(t) = .5$, $u_3(t) = .5625$, $u_3'(t) = 1.125$, $u_4(t) = 1$, $u_4'(t) = 2.]$. We note that not only the solution but its first t-derivative are approximated at the grid points.

When a twelfth argument $t = 1$ is added to the procedure, the numerical solution at $t = 1$ is exhibited in tabular form as for the procedure $NumHeatLines$. The global variable $SysODE$ displays the coupled system of ODEs (11.3.14) appropriate for this problem together with the initial conditions. It has the form given in Example 11.6 except that first derivatives are replaced by second derivatives in t. Also, an additional set of initial values is exhibited. They are $\{u_0'(0) = 0, u_1'(0) = 0, u_2'(0) = 0, u_3'(0) = 0, u_4'(0) = 0\}$. On applying the procedure *dsolve* to this system, Maple obtains the exact solution $u_0(t) = 0$, $u_1(t) = t^2/16$, $u_2(t) = t^2/4$, $u_3(t) = 9t^2/16$, $u_4(t) = t^2$. We do not consider the application of $NumWaveLines$ to this problem with the use of specific discretization schemes for the term $u_{tt}(x,t)$, as was done in Example 11.6 for the heat equation.

This concludes our presentation of difference schemes for the solution of the wave equation. Since the wave equation can be expressed as a system of two first order PDEs, the methods presented below for the numerical solution of hyperbolic systems of PDEs can be applied to the wave equation as well.

Exercises 11.3

11.3.1. Use *mtaylor* to verify the result (11.3.4).

11.3.2. Use the procedure $NumWaveForw$ to reproduce the results of Example 11.7.

11.3.3. Reproduce the results given in Example 11.8 using $NumWaveBackw$.

11.3.4. Use $NumWaveLines$ to verify the results of Example 11.9.

11.3.5. Introduce optional arguments for $NumWaveLines$ to solve the problem given in Example 11.9 using (a) Euler's method; (b) Heun's method.

11.4 FINITE DIFFERENCE METHODS FOR TWO-DIMENSIONAL LAPLACE AND POISSON EQUATIONS

Poisson's equation in two dimensions is

$$u_{xx}(x,y) + u_{yy}(x,y) = F(x,y). \tag{11.4.1}$$

Let $h > 0$ and $k > 0$ be increments in x and y, respectively. We replace the partial derivatives by centered difference quotients. The resulting *difference scheme* will be used to solve boundary value problems for (11.4.1) in (bounded) rectangular regions in the (x,y)-plane. Although it is possible to deal with problems over more general regions by introducing a special treatment of the boundary conditions, we do not carry this out here. The *finite element method*, presented in Chapter 12, can handle fairly general boundaries. Boundary value problems for Laplace's and Poisson's equations in three dimensions are also treated below.

On approximating $u_{xx}(x,y)$ and $u_{yy}(x,y)$ in (11.4.1) by *centered second order difference quotients*, we obtain the difference equation

$$\frac{u(x+h,y) - 2u(x,y) + u(x-h,y)}{h^2}$$

$$+\frac{u(x,y+k) - 2u(x,y) + u(x,y-k)}{k^2} = F(x,y). \tag{11.4.2}$$

To determine the *truncation error* that results on replacing (11.4.1) by (11.4.2), we apply *mtaylor* to (11.4.2). This gives $u_{xx}(x,y) + u_{yy}(x,y) - F(x,y) \approx (1/12)u_{xxxx}h^2 - (1/12)u_{yyyy}k^2$ and shows that the truncation error is $O(h^2 + k^2)$. As $h, k \to 0$, the truncation error tends to zero, so that (11.4.2) is *consistent* with (11.4.1).

The points $(x_i, y_j) = (\alpha + ih, \beta + jk)$ with $i, j = 0, \pm 1, \pm 2, \ldots$, and α and β as arbitrary constants represent the *grid points* for the general difference scheme (11.4.2). Let $u_{i,j}$ represent the (numerical) approximation of the exact solution $u(x,y)$ of a given boundary value problem at the point (x_i, y_j) [i.e., $u_{i,j} \approx u(x_i, y_j)$], (11.4.2) can be given as

$$\frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2} + \frac{u_{i,j+1} - 2u_{i,j} + u_{i,j-1}}{k^2} = F_{i,j}, \tag{11.4.3}$$

with $F_{i,j} = F(x_i, y_j)$. Solving for $u_{i,j}$ yields

$$u_{i,j} = \frac{k^2(u_{i+1,j} + u_{i-1,j}) + h^2(u_{i,j+1} + u_{i,j-1}) - h^2k^2F_{i,j}}{2(h^2 + k^2)}. \tag{11.4.4}$$

If $h = k$, (11.4.4) becomes

$$u_{i,j} = \frac{1}{4}(u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1}) - \frac{h^2}{4}F_{i,j}. \tag{11.4.5}$$

With $F(x, y) = 0$, so that (11.4.1) becomes *Laplace's equation*, $u_{i,j}$ is given as the average of its four neighboring values in (11.4.5) and is a weighted average of its four neighboring values in (11.4.4). It is useful to rewrite (11.4.4) as

$$2\left(1 + \frac{h^2}{k^2}\right) u_{i,j} - (u_{i+1,j} + u_{i-1,j}) - \frac{h^2}{k^2}(u_{i,j+1} + u_{i,j-1}) = -h^2 F_{i,j}. \quad (11.4.6)$$

We consider boundary value problems of all three kinds for Laplace's or Poisson's equation in the rectangular region $a < x < b$, $c < y < d$. The x and y intervals are subdivided into n_x and n_y equal parts, respectively. Then the x and y increments are $h = (b - a)/n_x$ and $k = (d - c)/n_y$, respectively. Thus, the grid points are given as $(x_i, y_j) = (a + ih, c + jk)$ with $i = 0, 1, 2, \ldots, n_x$ and $j = 0, 1, 2, \ldots, n_y$. The sides of the rectangle are boundary lines, so that $i = 0$, $i = n_x$, $j = 0$, $j = n_y$ correspond to $x = a$, $x = b$, $y = c$, $y = d$, respectively. If *Dirichlet conditions* are assigned on the boundary, the grid for the problem comprises only the interior points of the rectangular region, so that $i = 1, 2, \ldots, n_x - 1$ and $j = 1, 2, \ldots, n_y - 1$. If *Neumann* or *Robin boundary conditions* are prescribed on one or more sides of the rectangle, the derivative terms $u_x(x, y)$ and/or $u_y(x, y)$ are replaced by centered differences. The introduction of ghost points and their elimination proceeds as in the discussion presented above for the heat and wave equations. We do not repeat it here. Then, with a *Neumann* or *Robin boundary condition problem*, the grid points are $i = 0, 1, 2, \ldots, n_x$ and $j = 0, 1, 2, \ldots, n_y$. With mixed boundary conditions, grid points that correspond to the sides of the rectangle where Dirichlet conditions are prescribed are eliminated from the solution grid set.

In each of the foregoing cases, the *difference scheme* (11.4.6) is used. It gives rise to a system of simultaneous linear equations for the $u_{i,j}$. The number of equations depends on the range of the indices i and j. For a Dirichlet problem, there are $N = (n_x - 1)(n_y - 1)$ equations for the $u_{i,j}$, $i = 1, 2, \ldots, n_x - 1$, $j = 1, 2, \ldots, n_y - 1$. For a Neumann or Robin problem there are $M = n_x n_y$ equations for the $u_{i,j}$, $i = 0, 1, 2, \ldots, n_x$, $j = 0, 1, 2, \ldots, n_y$. The number of equations lies between N and M when dealing with a mixed problem.

The difference scheme (11.4.6) can be represented in the matrix form $A\mathbf{u} = \mathbf{w}$, where A is the matrix of coefficients of the linear system, $\mathbf{u}$ is a column vector whose elements are all the unknown $u_{i,j}$ and $\mathbf{w}$ is a column vector whose elements are known boundary values and values of F at grid points. The general form of system of difference equations and the matrix equation $A\mathbf{u} = \mathbf{w}$ varies with the type of boundary conditions assigned and the sizes of the x and y increments. We use the Maple procedure $LaplaceMatrix$ to display the appropriate results in a number of specific cases and then in the general case. It is given as $LaplaceMatrix(F(x,y), x = a..b, lbcfactor, lbc, rbcfactor, rbc, y = c..d, lobcfactor, lobc, ubcfactor, ubc, n_x, n_y)$. Boundary value problems of all three kinds for Laplace's or Poisson's equation in the rectangle $a < x < b$, $c < y < d$ can be considered. The last two arguments n_x and n_y specify the number of x and y subintervals. The arguments *lbc, rbc, lobc, ubc* represent the boundary values prescribed on $x = a, x = b, y = c, y = d$, respectively. The arguments *lbcfactor, rbcfactor, lobcfactor, ubcfactor* determine the type of the boundary condition on each side of

the rectangle. For example, if $lbcfactor = dirichlet$ and $lbc = g(y)$, the Dirichlet condition $u(a, y) = g(y)$ is prescribed. If $lobcfactor = 0$ and $lobc = h(x)$, the Neumann condition $-u_y(x, c) = h(x)$ is assigned. If $ubcfactor = \lambda$ and $ubc = r(x)$, the Robin condition $u_y(x, d) + \lambda u(x, d) = r(x)$ is given.

The output of the procedure exhibits the Laplace or Poisson equation that is to be solved, together with the boundary conditions. In addition, the explicit matrix form $A\mathbf{u} = \mathbf{w}$ of these equations is displayed. The global variable $Syst$ displays the simultaneous linear system the $u_{i,j}$. (The procedure does not solve this system.) These equations can be solved either directly or by iteration. The global variable *MSyst* yields the matrix equation $A\mathbf{u} = \mathbf{w}$ and the global variables *CoeffMatrix, UK, NonHTerm* display the matrix A and the vectors $\mathbf{u}$ and $\mathbf{w}$, respectively.

Example 11.10. A Dirichlet Problem for Poisson's Equation. The procedure $LaplaceMatrix(F(x,y), x = 0..l, dirichlet, g_1(y), dirichlet, g_2(y), y = 0..L, dirichlet, g_3(x), dirichlet, g_4(x), 3, 3)$ solves the boundary value problem for Poisson's equation in the rectangle

$$u_{xx}(x, y) + u_{yy}(x, y) = F(x, y), \qquad 0 < x < l,\ 0 < y < L, \tag{11.4.7}$$

with the Dirichlet boundary condition

$$u(0, y) = g_1(y),\ u(l, y) = g_2(y),\ u(x, 0) = g_3(x),\ u(x, L) = g_4(x). \tag{11.4.8}$$

The x and y intervals are both subdivided into three equal parts and the step sizes are $h = l/3$ and $k = L/3$. The $u_{i,j}$ to be solved for and the $u(x, y)$ values they approximate are given as $u_{1,1} = u(l/3, L/3), u_{2,1} = u(2l/3, L/3), u_{1,2} = u(l/3, 2L/3), u_{2,2} = u(2l/3, 2L/3)$. The matrix system $A\mathbf{u} = \mathbf{w}$ is given as

$$\begin{bmatrix} 2 + 2h^2/k^2 & -1 & -h^2/k^2 & 0 \\ -1 & 2 + 2h^2/k^2 & 0 & -h^2/k^2 \\ -h^2/k^2 & 0 & 2 + 2h^2/k^2 & -1 \\ 0 & -h^2/k^2 & -1 & 2 + 2h^2/k^2 \end{bmatrix} \begin{bmatrix} u_{1,1} \\ u_{2,1} \\ u_{1,2} \\ u_{2,2} \end{bmatrix}$$

$$= \begin{bmatrix} g_1(k) + h^2 g_3(h)/k^2 - h^2 F(h, k) \\ -h^2 F(2h, k) + g_2(k) + h^2\, g_3(2h)/k^2 \\ h^2 g_4(h)/k^2 + g_1(2k) - h^2 F(h, 2k) \\ h^2 g_4(2h)/k^2 - h^2 F(2h, 2k) + g_2(2k) \end{bmatrix}. \tag{11.4.9}$$

If $h = k$, which is often the choice for the increments, the diagonal terms of the matrix A all equal 4 and the off-diagonal terms equal -1 or 0. The system of simultaneous linear equations for $u_{i,j}$ is easily determined from (11.4.9) or can be found by entering the global variable $Syst$.

We observe that the matrix A is essentially comprised of two submatrices

$$\hat{A} = \begin{bmatrix} 2 + 2h^2/k^2 & -1 \\ -1 & 2 + 2h^2/k^2 \end{bmatrix}, \quad \tilde{A} = \begin{bmatrix} -h^2/k^2 & 0 \\ 0 & -h^2/k^2 \end{bmatrix}, \tag{11.4.10}$$

in terms of which it assumes the block matrix form $A = \begin{bmatrix} \hat{A} & \tilde{A} \\ \tilde{A} & \hat{A} \end{bmatrix}$.

By inspection, it can be seen that in the general case, the matrix A assumes the block tridiagonal form

$$A = \begin{bmatrix} \hat{A} & \tilde{A} & & & \mathbf{0} \\ \tilde{A} & \hat{A} & \tilde{A} & & \\ & \ddots & \ddots & \ddots & \\ & & & \tilde{A} & \\ \mathbf{0} & & & \tilde{A} & \hat{A} \end{bmatrix}, \tag{11.4.11}$$

with

$$\hat{A} = \begin{bmatrix} 2+2h^2/k^2 & -1 & 0 & & \mathbf{0} \\ -1 & 2+2h^2/k^2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & & & -1 \\ \mathbf{0} & & & -1 & 2+2h^2/k^2 \end{bmatrix}, \tag{11.4.12}$$

$$\tilde{A} = \begin{bmatrix} -h^2/k^2 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & -h^2/k^2 \end{bmatrix}. \tag{11.4.13}$$

If $h = k$, the matrix $\tilde{A}$ becomes $-I$, where I is the identity matrix. In the general case, it is possible to multiply the matrix A by k^2/h^2 so that $\hat{A}$ is replaced by $-I$ and $\hat{A}$ is appropriately modified. The matrix A is not strictly diagonally dominant since there may be some rows whose elements sum to zero. Thus, we cannot conclude on that basis, as was done before, that A is nonsingular. An additional condition satisfied by the matrix guarantees that the matrix is nonsingular. Alternatively, it may be shown that zero is not an eigenvalue of A. We conclude that $A\mathbf{u} = \mathbf{w}$ has a unique solution.

As a concrete example we consider Laplace's equation,

$$u_{xx}(x,y) + u_{yy}(x,y) = 0, \qquad 0 < x < 1,\ 0 < y < 1, \tag{11.4.14}$$

with the Dirichlet boundary condition

$$u(0,y) = 1,\ u(1,y) = 0,\ u(x,0) = 0,\ u(x,1) = 0. \tag{11.4.15}$$

With four subdivisions of the x and y intervals, we use $LaplaceMatrix(0, x = 0..1, dirichlet, 1, dirichlet, 0, y = 0..1, dirichlet, 0, dirichlet, 0, 4, 4)$. The global variable $Syst$ displays the system of equations for the $u_{i,j}$, which is not exhibited here. The solution of the system found by *solve* is $u_{1,1} = 3/7, u_{1,2} = 59/112, u_{1,3} = 3/7, u_{2,1} = 3/16, u_{2,2} = 1/4, u_{2,3} = 3/16, u_{3,1} = 1/14, u_{3,2} = 11/112, u_{3,3} = 1/14$.

In terms of the random walk problems discussed in Section 1.3, the discrete problem associated with (11.4.14)–(11.4.15) determines the probabilities $u_{i,j}$ that a particle starting from an interior grid point $(i/4, j/4)$ reaches a grid point $(0, j/4)$ on the side $x = 0$ of the unit square, before it reaches a grid point on one of the other three sides where it is absorbed. As expected, the probabilities $u_{i,j}$ decrease as i increases from 1 to 3, since the distance from $x = 0$ is thereby increased.

Example 11.11. A Neumann Problem for Poisson's Equation. The procedure $LaplaceMatrix(F(x,y), x = 0..l, 0, g_1(y), 0, g_2(y), y = 0..L, 0, g_3(x), 0, g_4(x), 1, 1)$ solves the boundary value problem for Poisson's equation in the rectangle $u_{xx}(x,y) + u_{yy}(x,y) = F(x,y)$, $0 < x < l$, $0 < y < L$, with the Neumann boundary condition $u_x(0,y) = -g_1(y), u_x(l,y) = g_2(y), u_y(x,0) = -g_3(x), u_y(x,L) = g_4(x)$. To simplify the presentation the x and y intervals are not subdivided, so that the step sizes are $h = l$ and $k = L$. The $u_{i,j}$ to be solved for and the $u(x,y)$ values they approximate are given as $u_{0,0} = u(0,0), u_{1,0} = u(l,0), u_{0,1} = u(0,L), u_{1,1} = u(l,L)$. The four grid points coincide with the four vertices of the rectangle.

The matrix system $A\mathbf{u} = \mathbf{w}$ is given as

$$\begin{bmatrix} 2+2h^2/k^2 & -2 & -2h^2/k^2 & 0 \\ -2 & 2+2h^2/k^2 & 0 & -2h^2/k^2 \\ -2h^2/k^2 & 0 & 2+2h^2/k^2 & -2 \\ 0 & -2h^2/k^2 & -2 & 2+2h^2/k^2 \end{bmatrix} \begin{bmatrix} u_{0,0} \\ u_{1,0} \\ u_{0,1} \\ u_{1,1} \end{bmatrix}$$

$$= \begin{bmatrix} 2hg_1(0) + 2h^2 g_3(0)/k - h^2 F(0,0) \\ -h^2 F(h,0) + 2hg_2(0) + 2h^2\, g_3(h)/k \\ 2h^2 g_4(0)/k + 2hg_1(k) - h^2 F(0,k) \\ 2hg_2(k) - h^2 F(h,k) + 2h^2 g_4(h)/k \end{bmatrix}. \qquad (11.4.16)$$

If $h = k$, the diagonal terms of the matrix A all equal 4 and the off-diagonal terms equal -2 or 0. The system of simultaneous linear equations for $u_{i,j}$ is found from (11.4.16) or by entering the global variable $Syst$.

The sum of the four rows of the matrix A is zero (i.e., the zero vector), so that A is singular. As a result, (11.4.16) has no solution unless the sum of the four elements of the vector $\mathbf{w}$ add up to zero. (This is the *compatibility condition*.) Then the solution is not unique and is determined up to a constant multiple of the vector $[1,1,1,1]$. The compatibility condition yields $(1/2L)(g_1(0) + g_2(0) + g_1(L) + g_2(L)) + (1/2l)(g_3(0) + g_3(l) + g_4(0) + g_4(l)) = (1/4)(F(0,0) + F(l,0) + F(0,L) + F(l,L))$. Now it has been shown previously that the Neumann problem for Poisson's equation $\nabla^2 u = F$ in the region G with the boundary condition $\partial u/\partial n = g$ over the boundary ∂G has no solution unless the following *compatibility condition* is satisfied $\int_{\partial G} g\, ds = \iint_G F\, da$. If this condition is met, the solution of the problem is determined only up to an arbitrary constant. It is easily verified that on numerically approximating the integrals in the integral compatibility condition by the trapezoidal rule, using the subdivision of the rectangle given above, the discrete compatibility condition results.

As the number of grid points is increased, the resulting coefficient matrix A does not assume a form that is a straightforward generalization of (11.4.16). The main diagonal elements all equal $2 + 2h^2/k^2$. The matrix is *banded* in the sense that only a fixed number of diagonals adjacent to the main diagonal contain nonzero elements. Their number increases with number of grid points, but the matrix becomes increasingly *sparse* (i.e., most of its elements are zero) as the grid points increase in number. Additionally, each matrix is singular because its column vectors sum to the zero vector, so that a compatibility condition that guarantees a solution of the system must be found.

As an example we consider Poisson's equation with the Neumann boundary condition

$$u_{xx}(x,y) + u_{yy}(x,y) = 2y, \qquad 0 < x < 1,\ 0 < y < 1,$$

$$u_x(0,y) = 0,\ u_x(1,y) = 2y,\ u_y(x,0) = x^2,\ u_y(x,1) = x^2. \quad (11.4.17)$$

The exact solution of this problem is $u(x,y) = x^2y + c$, where c is an arbitrary constant. As a result, the compatibility condition is satisfied. With two subdivisions of the x and y intervals, we use the procedure $LaplaceMatrix(2y, x = 0..1, 0, 0, 0, 2y, y = 0..1, 0, -x^2, 0, x^2, 2, 2)$. The matrix system $A\mathbf{u} = \mathbf{w}$ is

$$\begin{bmatrix} 4 & -2 & 0 & -2 & 0 & 0 & 0 & 0 & 0 \\ -1 & 4 & -1 & 0 & -2 & 0 & 0 & 0 & 0 \\ 0 & -2 & 4 & 0 & 0 & -2 & 0 & 0 & 0 \\ -1 & 0 & 0 & 4 & -2 & 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & -1 & 4 & -1 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 & -2 & 4 & 0 & 0 & -1 \\ 0 & 0 & 0 & -2 & 0 & 0 & 4 & -2 & 0 \\ 0 & 0 & 0 & 0 & -2 & 0 & -1 & 4 & -1 \\ 0 & 0 & 0 & 0 & 0 & -2 & 0 & -2 & 4 \end{bmatrix} \begin{bmatrix} u_{0,0} \\ u_{1,0} \\ u_{2,0} \\ u_{0,1} \\ u_{1,1} \\ u_{2,1} \\ u_{0,2} \\ u_{1,2} \\ u_{2,2} \end{bmatrix} = \begin{bmatrix} 0 \\ -1/4 \\ -1 \\ -1/4 \\ -1/4 \\ 3/4 \\ -1/2 \\ -1/4 \\ 5/2 \end{bmatrix}, \quad (11.4.18)$$

with $u_{i,j}$ as an approximation to the exact solution $u(i/2, j/2)$.

The global variable $Syst$ displays the system of equations for the $u_{i,j}$. The (nonunique) solution of the system found by the Maple procedure *solve* is $u_{0,0} = -1/2 + c, u_{2,2} = 1/2 + c, u_{1,2} = -1/4 + c, u_{1,0} = -1/2 + c, u_{2,1} = c, u_{2,0} = -1/2 + c, u_{0,1} = -1/2 + c, u_{1,1} = -3/8 + c, u_{0,2} = -1/2 + c$. Here c is an arbitrary constant. If we put $c = 1/2$, the numerical solution values agree with the exact solution $u(x,y) = x^2y$ evaluated at the grid points.

Example 11.12. A Robin Problem for Poisson's Equation. The Maple procedure $LaplaceMatrix(F(x,y), x = 0..l, \lambda, g_1(y), \lambda, g_2(y), y = 0..L, \lambda, g_3(x), \lambda, g_4(x), 1, 1)$ solves the boundary value problem for Poisson's equation in the rectangle $u_{xx}(x,y) + u_{yy}(x,y) = F(x,y),\ 0 < x < l,\ 0 < y < L$, with the

Robin boundary conditions $-u_x(0,y) + \lambda u(0,y) = g_1(y)$, $u_x(l,y) + \lambda u(l,y) = g_2(y)$, $-u_y(x,0) + \lambda u(x,0) = g_3(x)$, $u_y(x,L) + \lambda u(x,L) = g_4(x)$. Again, the x and y intervals are not subdivided and the step sizes are $h = l$ and $k = L$. The $u_{i,j}$ to be solved for and the $u(x,y)$ values they approximate are given as $u_{0,0} = u(0,0)$, $u_{1,0} = u(l,0)$, $u_{0,1} = u(0,L)$, $u_{1,1} = u(l,L)$. The four grid points lie on the four vertices of the rectangle.

The matrix system $A\mathbf{u} = \mathbf{w}$ for the problem above is given as

$$A = \begin{bmatrix} 2+2C(h,k,\lambda) & -2 & -2h^2/k^2 & 0 \\ -2 & 2+2C(h,k,\lambda) & 0 & -2h^2/k^2 \\ -2h^2/k^2 & 0 & 2+2C(h,k,\lambda) & -2 \\ 0 & -2h^2/k^2 & -2 & 2+2C(h,k,\lambda) \end{bmatrix},$$

$$\mathbf{u} = \begin{bmatrix} u_{0,0} \\ u_{1,0} \\ u_{0,1} \\ u_{1,1} \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} 2hg_1(0) + 2h^2 g_3(0)/k - h^2 F(0,0) \\ -h^2 F(h,0) + 2hg_2(0) + 2h^2\, g_3(h)/k \\ 2h^2 g_4(0)/k + 2hg_1(k) - h^2 F(0,k) \\ 2hg_2(k) - h^2 F(h,k) + 2h^2 g_4(h)/k \end{bmatrix},$$

where $C(h,k,\lambda) = (h^2 + h^2 k\lambda + hk^2\lambda)/k^2$. If $h = k$, the diagonal terms of the matrix A all equal $4 + h\lambda$ and the off-diagonal terms equal -2 or 0. The system of simultaneous linear equations for $u_{i,j}$ is easily determined from the above or can be found by entering the global variable $Syst$. The matrix A is (strictly) diagonally dominant for $\lambda > 0$ and is nonsingular, so that the matrix system has a unique solution. Increasing the grid points yields a coefficient matrix A whose form is not a straightforward generalization of the matrix above, as was the case for the Neumann problem. We do not discuss it here but note that it is a banded matrix and is nonsingular.

We consider Poisson's equation $u_{xx}(x,y) + u_{yy}(x,y) = 2y$, $0 < x < 1$, $0 < y < 1$ with the Robin boundary condition $-u_x(0,y) + u(0,y) = 0, u_x(1,y) + u(1,y) = 3y, -u_y(x,0) + u(x,0) = -x^2, u_y(x,1) + u(x,1) = 2x^2$. The exact solution of this problem is $u(x,y) = x^2 y$. With two subdivisions of the x and y intervals, we use the procedure $LaplaceMatrix(2y, x = 0..1, 1, 0, 1, 3y, y = 0..1, 1, -x^2, 1, 2x^2, 2, 2)$. The output gives the matrix system $A\mathbf{u} = \mathbf{w}$ for the problem, which we do not display.

The global variable $Syst$ displays the system of equations for the $u_{i,j}$, which is not exhibited here. The solution of the system found by the Maple procedure *solve* is $u_{0,0} = 0, u_{2,2} = 1, u_{1,2} = 1/4, u_{1,0} = 0, u_{2,0} = 0, u_{0,1} = 0, u_{1,1} = 1/8, u_{0,2} = 0, u_{2,1} = 1/2$. The numerical solution values agree with the exact solution $u(x,y) = x^2 y$ evaluated at the grid points.

The procedure $LaplaceMatrix$ can also handle problems with mixed boundary conditions but we do not present any examples here. $LaplaceMatrix$ determines the linear system of difference equations that arise for each problem. The solution of this system can then be found by invoking Maple's *solve* procedure. This may not be practical if there are a large number of grid points and a corresponding large number of equations. Since $LaplaceMatrix$ also obtains the matrix form of the system for each problem, it is possible to make use of the $LinearSolve$ procedure from Maple's

LinearAlgebra package in each case. This enables the use of special techniques such as *LU Decomposition* to find solutions of the matrix equations. To deal with this matter, a number of special iteration techniques have been developed for the solution of the system of difference equations and their matrix equation representations. They have proven to be extremely useful if the system contains a large number of equations. We present a number of these iteration methods in the next subsection.

Jacobi, Gauss-Seidel, and Relaxation Methods for Two-Dimensional Laplace and Poisson Equations

We begin by considering a linear system of equations $A\mathbf{x} = \mathbf{b}$ with n equations and n unknowns given as

$$\sum_{j=1}^{n} a_{1,j}x_j = b_1, \sum_{j=1}^{n} a_{2,j}x_j = b_2, \ldots, \sum_{j=1}^{n} a_{n,j}x_j = b_n. \tag{11.4.19}$$

(We use single-subscript notation for the unknowns to simplify the presentation, but the results are valid for the all the difference schemes for Laplace's and Poisson's equations presented above.) The matrix A has the elements a_{ij}, $i, j = 1, 2, \ldots, n$ and we assume that the diagonal elements a_{ii}, $i = 1, 2, \ldots, n$ are all nonzero. Then (11.4.19) can then represented in the form

$$x_i = \frac{b_i - \sum_{j=1}^{i-1} a_{i,j}x_j - \sum_{j=i+1}^{n} a_{i,j}x_j}{a_{i,i}}, \qquad i = 1, 2, \ldots, n. \tag{11.4.20}$$

The system is now in a form where it can readily be solved by *iteration*. To do so, we begin with an (arbitrarily chosen) initial approximation to the x_i which we denote as $x_i(0)$, and substitute it in the right side of (11.4.20) to determine the next approximation $x_i(1)$. We iterate this process, and the approximation after the kth iteration is given as $x_i(k)$. Three iteration schemes—the Jacobi, the Gauss-Seidel, and the SOR iteration methods—are presented below. Each of these schemes differs in how the $x_i(k+1)$ are determined from the $x_i(k)$ in the iteration process.

The most obvious scheme is the *Jacobi iteration method*, given as

$$x_i(k+1) = \frac{b_i - \sum_{j=1}^{i-1} a_{i,j}x_j(k) - \sum_{j=i+1}^{n} a_{i,j}x_j(k)}{a_{i,i}}, \quad i = 1, 2, \ldots, n. \tag{11.4.21}$$

Although an improved value of x_1 is already known once approximate values of the x_i are substituted into the first equation of the system (11.4.20), this value is not used in finding x_2. This approach is applied in the determination of all the x_i in a single iteration, as seen from (11.4.21). Once an iteration is complete, the new values of the x_i are used to obtain the next approximations.

A scheme that uses the improved values of the x_i as soon as they are computed within an iteration is the *Gauss-Seidel iteration method.* It is

$$x_i(k+1) = \frac{b_i - \sum_{j=1}^{i-1} a_{i,j}x_j(k+1) - \sum_{j=i+1}^{n} a_{i,j}x_j(k)}{a_{i,i}}, \qquad i = 1, 2, \ldots, n. \tag{11.4.22}$$

We observe that if the Jacobi or Gauss-Seidel iteration scheme converges, so that $\lim_{k\to\infty} x_i(k) = x_i$, then x_i is a solution of (11.4.20).

In an effort to *accelerate* the rate of convergence of the Gauss-Seidel iteration scheme, we note that the right side of (11.4.22) can be expressed as

$$x_i(k+1) = x_i(k) + \left(\frac{b_i - \sum_{j=1}^{i-1} a_{i,j}x_j(k+1) - \sum_{j=i+1}^{n} a_{i,j}x_j(k)}{a_{i,i}} - x_i(k) \right). \tag{11.4.23}$$

The first term on the right of (11.4.23) is the previously determined value $x_i(k)$ and the second term is the correction to that value determined by the Gauss-Seidel scheme. We introduce a parameter ω, known as the *relaxation parameter*, to interpolate or extrapolate between these two values. This yields

$$x_i(k+1) = (1-\omega)x_i(k) + \frac{\omega\left(b_i - \sum_{j=1}^{i-1} a_{i,j}x_j(k+1) - \sum_{j=i+1}^{n} a_{i,j}x_j(k)\right)}{a_{i,i}} \tag{11.4.24}$$

for $i = 1, 2, \ldots, n$. This modified iteration scheme, for which it is found that we must have $0 < \omega < 2$ for convergence, is referred to as a *relaxation method.* (Then, with $\omega < 1$ and $\omega > 1$, it is an under- and overrelaxation method, respectively.) If $\omega = 1$, it reduces to the Gauss-Seidel method. It can be shown that we must take $\omega > 1$, in general, to accelerate the convergence rate. For this reason, (11.4.24) is called the *successive overrelaxation method* or *SOR method.* Optimal values of the relaxation parameter ω can be determined for specific schemes that minimize the number of iterations required to achieve a certain level of accuracy.

Each of the foregoing iteration methods can be formulated in *matrix form.* The matrix representation is useful in determining the stability and convergence properties of these schemes. Given the matrix equation $A\mathbf{x} = \mathbf{b}$, we decompose the (square) matrix A into a diagonal matrix D and a lower and upper triangular matrix L and U, respectively. The diagonal elements of D are those of A. The elements of L below the main diagonal are those of A, and all the elements of L on or above the main diagonal are zero. The elements of U above the main diagonal are those of A and all the elements of U on or below the main diagonal are zero. Hence, $A = L + D + U$. For example, if A is a 2×2 matrix we have

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix}, D = \begin{bmatrix} a_{1,1} & 0 \\ 0 & a_{2,2} \end{bmatrix}, L = \begin{bmatrix} 0 & 0 \\ a_{2,1} & 0 \end{bmatrix}, U = \begin{bmatrix} 0 & a_{1,2} \\ 0 & 0 \end{bmatrix}.$$

We assume that all the diagonal elements of the matrix A and D are nonzero, so that D has an inverse. Then, $A\mathbf{x} = (L + D + U)\mathbf{x} = D\mathbf{x} + (L + U)\mathbf{x} = \mathbf{b}$, yields $\mathbf{x} = -D^{-1}(L + U)\mathbf{x} + D^{-1}\mathbf{b}$. This corresponds to the system (11.4.20).

The matrix version of the *Jacobi method* is

$$\mathbf{x}(k+1) = -D^{-1}(L+U)\,\mathbf{x}(k) + D^{-1}\mathbf{b}, \tag{11.4.25}$$

where $\mathbf{x}(k)$ is the kth iteration. The Gauss-Seidel iteration scheme can be expressed as $\mathbf{x}(k+1) = -D^{-1}L\mathbf{x}(k+1) - D^{-1}U\mathbf{x}(k) + D^{-1}\mathbf{b}$. On multiplying across by D, solving for $\mathbf{x}(k+1)$, and then multiplying across by $(D + L)^{-1}$, we obtain the matrix version of the *Gauss-Seidel method*

$$\mathbf{x}(k+1) = -(D+L)^{-1}U\,\mathbf{x}(k) + (D+L)^{-1}\mathbf{b}. \tag{11.4.26}$$

Since $D + L$ is a lower triangular matrix with nonzero diagonal elements, it has a nonzero determinant and an inverse. Finally, the matrix version of the *SOR method* is found to be

$$\mathbf{x}(k+1) = (D+\omega L)^{-1}[(1-\omega)D - \omega U]\,\mathbf{x}(k) + \omega(D+\omega L)^{-1}\mathbf{b}. \tag{11.4.27}$$

Again, $D + \omega L$ is a nonsingular lower triangular matrix.

We have created a procedure $NumLaplace$ that uses the iteration methods presented above to solve the difference schemes presented in the preceding subsection. It is given as $NumLaplace(F(x,y), x = a..b, lbcfactor, lbc, rbcfactor, rbc, y = c..d, lobcfactor, lobc, ubcfactor, ubc, n_x, n_y, numits, inguess, err, par)$. As before, the first 13 arguments of the procedure prescribe the Laplace or Poisson equation, the rectangular region, the boundary conditions, and the number of x and y subdivisions. The fourteenth argument determines the maximum number of iterations that are to be carried out. The fifteenth argument specifies the initial guess for the iteration scheme at each grid point. The sixteenth argument err assigns a numerical value that the maximum of the residual errors (to be defined below) at each grid point cannot exceed at the conclusion of each iteration. If the value of err is greater than each of the residual errors, the iteration process is stopped. (This is a *closure* or *termination criterion* for the procedure and the value of err determines an *error tolerance*.) The output of the procedure displays the total number of iterations carried out. The final argument determines which iteration method is used. If it is $jacobi$ or $Jacobi$, the *Jacobi method* is used. If, instead, a relaxation parameter ω is entered, it must have a value between 0 and 2] otherwise, an error message is printed out. If $\omega = 1$, the *Gauss-Seidel method* is used. Otherwise, the *SOR method* is employed.

After $NumLaplace$ is invoked, a plot of the finite difference solution can be generated by using the procedure $PlotTab(Sol, x, y, lh, lv, z)$. The global variables lh and lv list the points in the subdvisions of the x and y intervals, respectively. The global variable Sol, given as $Sol(lh[i], lv[j])$, yields the finite difference solution for the ith and jth points in the respective subdivisions. (The remaining arguments

determine what labels are to be used in the plot.) The procedure constructs a set of points in (x, y, z)-space and uses $surfdata$ to plot the surface.

Each of the foregoing iteration schemes is applied to the difference equation (11.4.3). We define the residual $r_{i,j}$ at a grid point (i, j) to be

$$r_{i,j} = \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2} + \frac{u_{i,j+1} - 2u_{i,j} + u_{i,j-1}}{k^2} - F_{i,j}. \qquad (11.4.28)$$

If $u_{i,j}$ is an exact solution of the system of difference equations, the residual vanishes. We use the absolute value of the residual as a measure of how close an approximate solution $u_{i,j}(k)$, determined after the kth iteration, is to the exact solution $u_{i,j}$. Ideally, the difference between the approximate and exact solutions should be measured in some norm to determine the accuracy of the approximation, but the exact solution is unavailable. In addition, it may happen that an approximate solution has small residuals but is not close to the exact solution. Nevertheless, the procedure *NumLaplace* evaluates the maximum absolute value of all the residuals after the kth iteration, and if its sixteenth argument err exceeds this value, the iteration process is terminated.

To see the connection between the residual and the error more clearly, we return to the matrix form of the system $A\mathbf{x} = \mathbf{b}$. Let $\mathbf{X}$ be an approximate solution of $A\mathbf{x} = \mathbf{b}$, and define the (vector) residual to be $\mathbf{r} = A\mathbf{X} - \mathbf{b}$. The difference between the approximate and exact solution $\mathbf{X} - \mathbf{x}$ satisfies $A(\mathbf{X} - \mathbf{x}) = A\mathbf{X} - \mathbf{b} = \mathbf{r}$, $\mathbf{X} - \mathbf{x} = A^{-1}\mathbf{r}$, with A^{-1} as the inverse of A. Then it follows from a property of matrix and vector norms (some of these norms are presented in Section 11.6) that $\| \mathbf{X} - \mathbf{x} \| = \| A^{-1}\mathbf{r} \| \leq \| A^{-1} \| \| \mathbf{r} \|$. Thus, even if the residual vector $\mathbf{r}$ has a small norm, the absolute error $\| \mathbf{X} - \mathbf{x} \|$ can have a large value if $\| A^{-1} \|$ is large. Now, $\| \mathbf{b} \| = \| A\mathbf{x} \| \leq \| A \| \| \mathbf{x} \|$ and $1/\| \mathbf{x} \| \leq \| A \| / \| \mathbf{b} \|$. Thus, the relative error $\| \mathbf{X} - \mathbf{x} \| / \| \mathbf{x} \|$ is bounded by $\| \mathbf{X} - \mathbf{x} \| / \| \mathbf{x} \| \leq \| A \| \| A^{-1} \| \| \mathbf{r} \| / \| \mathbf{b} \|$. The *condition number* $C(A)$ of a matrix A is defined as $C(A) = \| A \| \| A^{-1} \|$. We see from this that even if the residual vector $\mathbf{r}$ has a small norm, a large condition number for the matrix can lead to a large relative error in the solution. If the condition number is close to 1, the magnitude of $\mathbf{r}$ gives a good indication of the size of the relative error in the solution. If the condition number is large, the problem is said to be *ill-conditioned.*

Example 11.13 A Dirichlet Problem for Laplace's Equation. Each of the foregoing iteration methods is used to solve the problem (11.4.14)–(11.4.15) of Example 11.10. $NumLaplace(0, x = 0..1, dirichlet, 1, dirichlet, 0, y = 0..1, dirichlet, 0, dirichlet, 0, 4, 4, 200, 0, 10^{-8}, Jacobi)$ solves the difference scheme using the *Jacobi method.* The maximum number of iterations permitted is 200 and the initial guess is 0. The maximum error allowed is 10^{-8}.

The output exhibits Laplace's equation and the boundary conditions, as well as the number of subdivisions and the maximum number of iterations. It declares that the Jacobi method was used and that 61 iterations were performed. The values of the

$u_{i,j}$ are displayed in tabular form as

$$\begin{bmatrix} y \setminus x & 0.0 & 0.25 & 0.50 & 0.75 & 1.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.25 & 1.0 & 0.429 & 0.187 & 0.0714 & 0.0 \\ 0.50 & 1.0 & 0.527 & 0.250 & 0.0982 & 0.0 \\ 0.75 & 1.0 & 0.429 & 0.187 & 0.0714 & 0.0 \\ 1.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \end{bmatrix}. \tag{11.4.29}$$

The first row and column of the table give the x and y values in the subdivisions, respectively. The remaining entries give the approximate solution values at these points.

The global variable $System$ exhibits the system of difference equations for the problem, and the global variable $GridPoints$ identifies the (i, j) values with corresponding (x_i, y_j) values. These values can then be entered into the global variable Sol in the form $Sol(x_i, y_j)$ to reproduce the solution values at each of the grid points. The solution values agree with the exact solution of the system of difference equations given in Example 11.10 if the exact solution values are converted to floating-point form.

The global variables lh and lv both equal $lh = lv = [0., .25, .50, .75, 1.0]$. The procedure $PlotTab(Sol, x, y, lh, lv, z)$ yields a plot of the finite difference solution. It is exhibited in Figure 11.5.

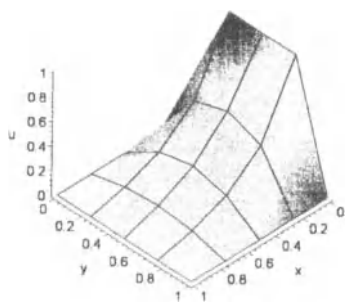

Figure 11.5 The finite difference solution.

The procedure $NumLaplace(0, x = 0..1, dirichlet, 1, dirichlet, 0, y = 0..1, dirichlet, 0, dirichlet, 0, 4, 4, 200, 0, 10^{-8}, 1)$ uses the *Gauss-Seidel method* to solve the foregoing problem, since the last argument is 1. (Otherwise, everything is the same as before.) The solution agrees with that obtained via the Jacobi method, but only 31 iterations are required to accomplish this. If 0.5 is the last argument in the procedure, the *SOR method* is used but 104 iterations are needed to get the required result. But if 1.2 is the last argument, only 16 iterations are required. It is a general principle that the Gauss-Seidel method requires approximately half the number of iterations of the Jacobi method, as is the case here. The SOR method can reduce the

number of iterations even further, but that generally requires the relaxation parameter to exceed 1. It is possible to consider a sequence of values of the relaxation parameter to determine an approximate optimal value of ω. For this problem, if we replace 1.2 by 1.21 in the SOR method, the number of iterations is reduced to 14.

Alternating-Direction Implicit Method for Two-Dimensional Laplace and Poisson Equations

The iteration schemes used in the Gauss-Seidel and SOR methods are carried out point by point over the grid appropriate to each problem. By simultaneously determining the new values in the iteration scheme over an entire row or column of grid points, an acceleration of the convergence rate of the scheme can be achieved. (This can also be done for the Jacobi method, but we do not treat it here.) As a result, the solution of a system of simultaneous linear equations must be found at each step (so that this is an *implicit method*). However, this does not present a significant problem since each system is of tridiagonal form and is easily solved. This approach yields *Gauss-Seidel* and *SOR row or column iteration methods.* A further improvement can be achieved if instead of determining only grid row or grid column values in each iteration, we first determine the grid row values, say, followed by a determination of the grid column values (using the values found from the grid row calculations), within each iteration. Since we alternate the direction from horizontal (rows) to vertical (columns) within each iteration step and the values along the rows and columns are determined implicitly, we obtain what is called the *alternating-direction implicit method* or *ADI method.* [We remark that $u_{i,j}$ corresponds to $u(x_i, y_j)$, so that if we consider the $u_{i,j}$ to be the elements of a matrix and we put i = constant, we obtains the (horizontal) ith row of the matrix. Yet, in the (x, y)-plane we have the values x_i = constant, which lie on a vertical line. The converse relationship exists for j = constant and y_j = constant. Our use of *horizontal* and *vertical* in the above, refers to matrix notation.]

The *row* and *column iteration methods* are carried out as follows. It is assumed that there are n rows and m columns in the (matrix) grid. On replacing (for notational clarity) the squared increments h^2 and k^2 by hx^2 and hy^2, respectively, in the difference equation approximation (11.4.3) to Poisson's equation, and solving for $u_{i,j}$, we obtain

$$u_{i,j} = \frac{u_{i-1,j} + u_{i+1,j} + (hx^2/hy^2)(u_{i,j-1} + u_{i,j+1}) - hx^2 F_{i,j}}{2 + 2hx^2/hy^2}. \qquad (11.4.30)$$

To carry out *row iterations*, we assume that a set of values $u_{i,j}$ is prescribed on the grid. We perform a sweep along the n rows in the grid and determine a new set of values $\hat{u}_{i,j}$ by solving the tridiagonal system (for each row)

$$\hat{u}_{i,j} = \frac{\hat{u}_{i-1,j} + u_{i+1,j} + (hx^2/hy^2)(\hat{u}_{i,j-1} + \hat{u}_{i,j+1}) - hx^2 F_{i,j}}{2 + 2hx^2/hy^2}, \quad i = 1, \ldots, n. \qquad (11.4.31)$$

A fixed value of i determines a row, and for each row the subscript j varies from 1 to m, the number of columns. Hence, there are m equations (on the ith row) and the unknowns in each equation are $\hat{u}_{i,j-1}$, $\hat{u}_{i,j}$, $\hat{u}_{i,j+1}$, assuming none of the corresponding grid points are boundary points. In the *Gauss-Seidel method*, which we consider here, the term $\hat{u}_{i-1,j}$ has already been determined when solving for all values on the $(i-1)$st row. To obtain the *SOR method* we introduce the *relaxation parameter* ω and replace $\hat{u}_{i,j}$ by $u_{i,j} = \omega\,\hat{u}_{i,j} + (1-\omega)\,u_{i,j}$, as the (new) grid point values on each row. Then, everything proceeds as before.

To carry out *column iterations*, we assume that a set of values $u_{i,j}$ is prescribed on the grid. We perform a sweep along the m columns in the grid and determine a new set of values $\tilde{u}_{i,j}$ by solving the tridiagonal system (for each column)

$$\tilde{u}_{i,j} = \frac{\tilde{u}_{i-1,j} + \tilde{u}_{i+1,j} + (hx^2/hy^2)(\tilde{u}_{i,j-1} + u_{i,j+1}) - hx^2 F_{i,j}}{2 + 2hx^2/hy^2},\ i = 1, \ldots, n. \tag{11.4.32}$$

A fixed value of j determines a column, and for each column the subscript i varies from 1 to n, the number of rows. Hence, there are n equations (on the jth column), and the unknowns in each equation are $\tilde{u}_{i-1,j}$, $\tilde{u}_{i,j}$, $\tilde{u}_{i+1,j}$, assuming that none of the corresponding grid points are boundary points. In the *Gauss-Seidel method*, the term $\tilde{u}_{i,j-1}$ has been determined when solving for all values on the $(j-1)$st row. To obtain the *SOR method*, we introduce the *relaxation parameter* ω and replace $\tilde{u}_{i,j}$ by $u_{i,j} = \omega\,\tilde{u}_{i,j} + (1-\omega)\,u_{i,j}$ as the grid point values on each column. Then everything proceeds as before.

If in each iteration we sweep through all the rows and then sweep through all the columns to determine new values for the $u_{i,j}$, we obtain the *ADI method*. In all methods, an initial guess $u_{i,j}(0)$ is given for each grid point (at which the numerical solution is to be found) to start the iteration process. The new values found after each row, column, or dual sweep are used as the starting values for the next iteration. While the ADI method generally requires fewer iterations than are needed for the row or column method, it should be noted that within each iteration the sweep over the rows and the columns effectively doubles the number of iterations carried out in the method. An improvement is expected to be achieved because row sweeps may yield better results than column sweeps, and vice versa, for a particular problem, and both are used in each iteration.

$NumLaplaceRowColADI$ carries out the foregoing iteration methods for Laplace's or Poisson's equation, in the form *NumLaplaceRowColADI*$(F(x,y), x = a..b, lbcfactor, lbc, rbcfactor, rbc, y = c..d, lobcfactor, lobc, ubcfactor, ubc, n_x, n_y, numits, inguess, err, par, method)$. The first 16 arguments are identical to those for the procedure $NumLaplace$ given above. The seventeenth argument is the value of the relaxation parameter ω and must lie between 0 and 2. (The option *jacobi* is not available for procedure.) The last argument *method* must take the value *row*, *col* or ADI, and determines which of the three iteration methods is to be used. Any problem that can be solved by $NumLaplace$ can be solved by $NumLaplaceRowColADI$.

After $NumLaplaceRowColADI$ is invoked, a plot of the finite difference solution can be generated by using the procedure $PlotTab(Sol, x, y, lh, lv, z)$. The global variables lh and lv list the points in the subdvisions of the x and y intervals, respectively. The global variable $SolRC$, given as $SolRC(lh[i], lv[j])$, yields the finite difference solution for the ith and jth points in the respective subdivisions. (The remaining arguments determine what labels are to be used in the plot.) The procedure constructs a set of points in (x, y, z)-space and uses $surfdata$ to plot the surface.

Example 11.14. A Dirichlet Problem for Laplace's Equation. We reconsider the problem of Example 11.13. The procedure $NumLaplaceRowColADI(0, x = 0..1, dirichlet, 1, dirichlet, 0, y = 0..1, dirichlet, 0, dirichlet, 0, 4, 4, 200, 0, 10^{-7}, \omega, p)$ applies the Row, $Column$, or ADI iteration method to that problem using the acceleration factor ω and the error tolerance 10^{-7}, according as p equals row, col, or ADI, respectively. The numerical output agrees with (11.4.29) and is not displayed here. The difference lies in the number of iterations required to achieve those values.

With $\omega = 1$, the Gauss-Seidel method is used. While 27 iterations are required with $NumLaplace$, only 16 iterations are needed with $NumLaplaceRowColADI$ using the row or col option. The ADI method needs only nine iterations. If we set $\omega = 1.2$, the SOR method is used. Then 13 iterations are required with $NumLaplace$, and 13 and 12 iterations are needed with $NumLaplaceRowColADI$ using the row and col option, respectively. The ADI method needs only seven iterations. We do not consider other choices of the relaxation parameter and other boundary conditions.

Exercises 11.4

11.4.1. Use the procedure $LaplaceMatrix$ to obtain the matrix system (11.4.9) associated with the Dirichlet problem for Laplace's equation (11.4.7)–(11.4.8).

11.4.2. Use Maple methods to obtain an exact solution of the system (11.4.9).

11.4.3. Obtain the solution of (11.4.14)–(11.4.15) given in Example 11.10.

11.4.4. Reproduce the results given in Example 11.11 for the Neumann problem (11.4.17).

11.4.5. Obtain the matrix system and the solution of the problem in Example 11.12.

11.4.6. Invoke Maple's $LinearAlgebra$ package and determine the form of the 3×3 Hilbert matrix. Determine the condition number of the Hilbert matrix in various norms by using the procedure $ConditionNumber$.

11.4.7. Use the procedure $NumLaplace$ as given in Example 11.13 to solve the problem (11.4.14)–(11.4.15) using (a) the Jacobi method; (b) the Gauss-Seidel method; (c) the SOR method with $\omega = 0.5$; (d) the SOR method with $\omega = 1.2$; (e) the SOR method with $\omega = 1.21$. In each case, note the number of iterations used to obtain the final result.

11.4.8. Reconsider Exercise 11.4.7, where the initial guess in the iteration procedures was 0, and determine if a significant decrease in the number of iterations results if the initial guess is taken to be 0.25 or some other positive number between 0 and 1.

11.4.9. Use the procedure $NumLaplaceRowColADI$ to verify the results of Example 11.14.

11.5 VON NEUMANN STABILITY OF DIFFERENCE METHODS FOR PDEs

The concept of *stability* for linear (time-dependent) PDEs with constant coefficients was introduced in Section 3.5. It was shown that problems for PDEs that are well posed can be unstable. The concept of stability or instability, as defined for difference schemes, has more serious consequences. An unstable difference scheme yields an invalid numerical approximation to the exact solution of the given problem. That is, we do not expect the numerical solution to converge to the exact solution, even if the difference equation is consistent with the PDE. This follows from the *Lax equivalence theorem* stated above. There are several ways to define and characterize stability for difference schemes. The most commonly used definition, for difference schemes that result from time-dependent problems, is based on the *von Neumann stability criterion*. This approach makes use of Fourier analysis, as was the case in the stability analysis of Section 3.5.

As we are dealing with a discrete problem, we make use of the *discrete Fourier transform* introduced in Section 5.8 and reformulate the results to fit our present needs. Given the interval $0 \leq x \leq l$, we divide it into N equal parts and define $x_n = nl/N$, $n = 0, 1, \ldots, N$. The increment in x is given as $h = l/N$. The increment in t is k, and we have $t_m = mk$. Then $u_{n,m}$ corresponds to $u(x_n, t_m)$ in our subscript notation. Also, $x_n \pm h = x_{n\pm1}$ and $t_m + k = t_{m+1}$.

In terms of the grid of x values, we define the *discrete Fourier transform* of $u(x_n, t_m)$ as

$$c_s(m) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} u(x_n, t_m) \exp\left(\frac{-2i\pi s x_n}{l}\right), \quad s = 0, 1, \ldots, N-1, \quad (11.5.1)$$

with the $c_s(m)$ as the *Fourier coefficients*. (This set of values determines the transform.) The *inverse discrete Fourier transform* is given as

$$u(x_n, t_m) = \frac{1}{\sqrt{N}} \sum_{s=0}^{N-1} c_s(m) \exp\left(\frac{2i\pi s x_n}{l}\right), \quad n = 0, 1, \ldots, N-1. \quad (11.5.2)$$

The function $u(x_n, t_m)$ is reproduced from the Fourier coefficients (or, equivalently, the transform values) $c_s(m)$. We note that $u(x_N, t_m) = u(x_0, t_m)$, so that $u(x_n, t_m)$ is periodic.

The discrete Fourier transforms of $u(x_n \pm h, t_m)$ and $u(x_n, t_m + k)$ are given as $c_s(m)\exp(\pm 2i\pi sh/l)$ and $c_s(m+1)$, respectively, with similar results for higher increments. As a result, if we consider the difference equation

$$u(x_n, t_m + k) = au(x_n + h, t_m) + bu(x_n, t_m) + cu(x_n - h, t_m), \qquad (11.5.3)$$

its discrete Fourier transform yields the recursion relation

$$c_s(m+1) = c_s(m)\,[a\exp(2i\pi sh/l) + b + c\exp(-2i\pi sh/l)]. \qquad (11.5.4)$$

The solution of the recursion relation is

$$c_s(m) = c_s(0)\,[a\exp(2i\pi sh/l) + b + c\exp(-2i\pi sh/l)]^m, \qquad (11.5.5)$$

where $c_s(0)$ is the transform of the discrete initial conditions for the problem. The solution of the difference equation is

$$u(x_n, t_m) = \frac{1}{\sqrt{N}} \sum_{s=0}^{N-1} c_s(0)\,[a\exp(2i\pi sh/l) + b + c\exp(-2i\pi sh/l)]^m e^{(2i\pi sx_n/l)} \qquad (11.5.6)$$

for $n = 0, 1, \ldots, N-1$.

For *stability* we require that the terms $[a\exp(2i\pi sh/l) + b + c\exp(-2i\pi sh/l)]^m$ in (11.5.6) must be bounded in absolute value as $m \to \infty$ for all relevant s. As a result, the solution $u(x_n, t_m)$ cannot grow (in absolute value) without bound as $t_m \to \infty$. This implies that

$$|a\exp(2i\pi sh/l) + b + c\exp(-2i\pi sh/l)| \le 1 \qquad (11.5.7)$$

for all relevant s, and this is the *von Neumann stability condition*. As the number of subdivisions $N \to \infty$, the increment $h = l/N$ tends to zero, and the expression $2\pi sh/l = 2\pi s/N$ effectively ranges over the entire interval $[0, 2\pi]$. Thus the von Neumann stability condition can be given as, with $\beta = 2\pi sh/l$,

$$|\alpha| \equiv |a\exp(i\beta) + b + c\exp(-i\beta)| \le 1, \qquad 0 \le \beta < 2\pi. \qquad (11.5.8)$$

If the PDE that is consistent with the foregoing difference scheme is *unstable* in the sense of Section 3.5, the von Neumann stability condition takes the form

$$|\alpha| \equiv |a\exp(i\beta) + b + c\exp(-i\beta)| \le 1 + O(k), \qquad 0 \le \beta < 2\pi, \qquad (11.5.9)$$

where k is the time increment. Furthermore, since $x_n = nl/N = nh$ we have $2\pi sx_n/l = 2\pi shn/l = \beta n$, so that $\exp(2i\pi sx_n/l) = \exp(i\beta n)$. Each of these terms in the sum (11.5.6) is multiplied by a term of the form α^m (apart from a constant multiple), as we have seen.

We have shown that $u(x_n, t_m)$ is expressed as a sum (over n) of constant multiples of $\alpha^m \exp(i\beta n)$, which may be referred to as *Fourier modes*. They are the analogs of the *normal modes* used in the analysis of stability for PDEs considered in Section 3.5. We must show that the absolute value of each of the Fourier modes cannot grow without bound with increasing m to achieve stability for the difference scheme. Therefore, on applying a *von Neumann stability analysis* to a difference scheme, we substitute $u(x_n, t_m) = \alpha^m \exp(i\beta n)$, into the difference equation. This yields an equation for $\alpha = \alpha(\beta)$, which is referred to as the *amplification factor*. The von Neumann stability criterion requires that we have

$$|\alpha| \leq 1 \qquad \text{or} \qquad |\alpha| \leq 1 + O(k) \tag{11.5.10}$$

for stability. (This must be shown for all β in the interval $[0, 2\pi]$.) Now, the amplification factor α is a function of the increments h and k that occur in the difference scheme. If (11.5.10) is satisfied without any restriction on h and k, the scheme is *unconditionally stable*. If (11.5.10) is valid only if h and k are related in some way, the scheme is *conditionally stable*. If the condition (11.5.10) can never be met, the scheme is *unstable*. We remark that the stability analysis is always applied to the homogeneous version of the difference equations, that is, with the nonhomogeneous term equated to zero.

The von Neumann stability condition is a *necessary condition* for the stability of a difference scheme, but it is not a *sufficient condition* for all problems. Thus, even if the condition is met, the scheme can be unstable in that it may fail to yield a satisfactory approximation to the exact solution of the given problem. One reason for this is that the von Neumann criterion does not take into account the boundary conditions assigned for the difference scheme. An alternative stability criterion based on the matrix representation of the difference scheme does take the boundary conditions into account and yields more reliable results. (It is presented below.) Nevertheless, the results of both methods are generally in agreement.

von Neumann Stability for the Heat Equation

In the following examples, we determine the *von Neumann stability condition* for a number of the difference schemes for the heat equation that were presented in Section 11.2. Cases of conditional and unconditional stability, as well as instability, will be encountered. We begin by considering a conditionally stable difference scheme.

Example 11.15. The Explicit Forward Difference Scheme. The explicit forward difference equation for the homogeneous heat equation (11.2.1) with $F(x, t) = 0$ is given as

$$u_{n,m+1} = (1 - 2r)u_{n,m} + ru_{n+1,m} + ru_{n-1,m}, \tag{11.5.11}$$

where $r = kc^2/h^2$ [see (11.2.5)]. We insert $u_{n,m} = \alpha^m \exp(i\beta n)$ into (11.5.11) and obtain $\alpha = 2r\cos(\beta) + 1 - 2r$ after some elementary simplification. For stability

we must have $-1 \leq \alpha \leq 1$ as β ranges over the interval $[0, 2\pi]$. Using elementary calculus to determine the maximum and minimum of $\alpha = \alpha(\beta)$, we determine the critical points to be $\beta = 0, \pi, 2\pi$. The minimum equals $1 - 4r$ and the maximum is 1 (we recall that $r > 0$). The condition $1 - 4r \geq -1$ yields $r \leq 1/2$. Thus, the scheme is *conditionally stable* with the stability condition $r = kc^2/h^2 \leq \frac{1}{2}$. This result is consistent with that obtained in Section 11.2, based on a random walk analogy.

Next, we present an unstable difference scheme.

Example 11.16. An Unstable Forward Difference Scheme. The homogeneous version of the explicit two-step forward difference scheme (11.2.29) for the heat equation is given as

$$u_{n,m+1} = u_{n,m-1} + 2r(u_{n+1,m} - 2u_{n,m} + u_{n-1,m}). \tag{11.5.12}$$

We insert $u_{n,m} = \alpha^m \exp(i\beta n)$ into (11.5.12) and obtain the quadratic equation $\alpha^2 + 4r(1-\cos(\beta))\alpha - 1 = 0$, whose roots are $\alpha = 2r(\cos(\beta) - 1) \pm \sqrt{1+(2r(\cos(\beta)-1))^2}$. If we pick the minus sign in the above and put $\beta = \pi$, we see immediately that $\alpha < -1$ for all positive r. Since we must have $|\alpha| \leq 1$ for all β, we conclude that the scheme is unconditionally *unstable*.

Next, we present a difference scheme for the heat equation with a radiation term

$$u_t(x,t) + bu(x,t) = c^2 u_{xx}(x,t). \tag{11.5.13}$$

If the constant $b > 0$, (11.5.13) characterizes heat conduction with radiation. If $b < 0$, the PDE is unstable.

Example 11.17. An Explicit Forward Difference Scheme. We consider the following explicit forward difference scheme for (11.5.13)

$$u_{n,m+1} = (1 - 2r)u_{n,m} + ru_{n+1,m} + ru_{n-1,m} - kbu_{n,m}, \tag{11.5.14}$$

with $r = kc^2/h^2$. Inserting $u_{n,m} = \alpha^m \exp(i\beta n)$ into (11.5.14) yields $\alpha = 2r\cos(\beta) - 2r + 1 - kb$. On proceeding as in Example 11.15, we find that the minimum of α equals $1 - 4r - kb$ and the maximum is $1 - kb$. On using the Maple procedure $solve(\{1 - 4r - kb \geq -1 - kb, 1 - kb \leq 1 - kb\}, \{r\})$, we find that $r \leq 1/2$, so that the stability condition $|\alpha| \leq 1 + O(k)$ of (11.5.10) yields $r \leq 1/2$ as in Example 11.13. Thus, the scheme is *conditionally stable*. It may be noted that this result does not depend on the sign of b in (11.5.14).

The following scheme for the heat equation is unconditionally stable.

Example 11.18. The Implicit Backward Difference Scheme. The implicit backward difference equation for (11.2.1) with $F(x,t) = 0$, is

$$u_{n,m} = (1+2r)u_{n,m+1} - ru_{n+1,m+1} - ru_{n-1,m+1}, \tag{11.5.15}$$

where $r = kc^2/h^2$ [see (11.2.18)]. We insert $u_{n,m} = \alpha^m \exp(i\beta n)$ into (11.5.15) and obtain $\alpha = 1/(1+2r(1-\cos(\beta)))$. Since $\cos(\beta) \leq 1$, we conclude immediately that $0 \leq \alpha \leq 1$ for all r and β. Thus, the scheme is *unconditionally stable.*

The foregoing approach applied to the θ *scheme* of (11.2.36) determines that it is unconditionally stable if $1/2 \leq \theta < 1$, while if $0 < \theta < 1/2$ it is conditionally stable under the restriction $r \leq 1/(2-4\theta)$. The choice $\theta = 1/2$ gives the *Crank-Nicolson scheme.* We conclude our discussion of the heat equation with a stability analysis of the *DuFort-Frankel difference scheme.*

Example 11.19. The DuFort-Frankel Difference Scheme. The DuFort-Frankel explicit forward difference equation for the homogeneous heat equation (11.2.1) with $F(x,t) = 0$ is given as

$$u_{n,m+1} = u_{n,m-1} + 2r(u_{n+1,m} - u_{n,m-1} - u_{n,m+1} + u_{n-1,m}), \tag{11.5.16}$$

where $r = kc^2/h^2$ [see (11.2.32)]. We insert $u_{n,m} = \alpha^m \exp(i\beta n)$ into (11.5.16) and obtain the quadratic equation $(1+2r)\alpha^2 - 4r\cos(\beta)\alpha + 2r - 1 = 0$. The two roots are $\alpha = \left[2r\cos(\beta) \pm \sqrt{1-4r^2\sin^2(\beta)}\right]/(1+2r)$. To determine the von Neumann stability condition, we consider two cases. From the quadratic equation we conclude that the product of the two roots equals $(2r-1)/(2r+1)$ and this is bounded by 1 in absolute value for $r > 0$. Consequently, if the roots are complex valued, in which case they are complex conjugates, we have $|\alpha| \leq 1$ for each root, regardless of the value of r. This case occurs if $4r^2\sin^2(\beta) > 1$. If $4r^2\sin^2(\beta) \leq 1$, the radical in the numerator of the above is bounded by 1 and we find, on using the triangle inequality for each root, that $|\alpha| \leq (2r+1)/(2r+1) = 1$. As the stability condition is satisfied without any restriction on r, the scheme is *unconditionally stable.*

---

von Neumann Stability for the Wave Equation

We determine the conditions for the stability of the *explicit forward* and *implicit backward difference schemes* for the homogeneous wave equation (11.3.1), [where we put $F(x,t) = 0$], presented in Section 11.3.

Example 11.20. The Explicit Forward Difference Scheme. The explicit forward difference equation for the homogeneous wave equation is given as

$$u_{n,m+1} = 2(1-r^2)u_{n,m} + r^2u_{n+1,m} + r^2u_{n-1,m} - u_{n,m-1}, \tag{11.5.17}$$

where $r = kc/h$ [see (11.3.5)]. On inserting $u_{n,m} = \alpha^m \exp(i\beta n)$ into (11.5.17), we obtain the quadratic equation

$$\alpha^2 + (4r^2 \sin^2(\beta/2) - 2)\,\alpha + 1 = 0, \tag{11.5.18}$$

after some simplification. The two roots can be written as $\alpha = 1 - 2r^2 \sin^2(\beta/2) \pm 2r \sin(\beta/2)\sqrt{r^2 \sin^2(\beta/2) - 1}$, since $\beta \in [0, 2\pi]$. To determine the von Neumann stability condition we consider two cases. If $r \leq 1$, the two roots are complex conjugates and (11.5.18) shows that the product of the two roots equals 1. Consequently, each root has unit modulus, and we have $|\alpha| = 1$ for each root. If $r > 1$, we put $\beta = \pi$ in the above and select the minus sign. This gives $\alpha = 1 - 2r^2 - 2r\sqrt{r^2 - 1}$. This expression has the value -1 at $r = 1$ and it has a negative derivative for $r > 1$, so that it is a decreasing function. As we have found a value of β for which $\alpha < -1$ for all $r > 1$, we conclude that the forward difference scheme is *conditionally stable*, with $r \leq 1$ as the stability condition. This agrees with the *CFL* stability condition obtained in Section 11.3.

Example 11.21. The Implicit Backward Difference Scheme. The implicit backward difference equation for the homogeneous wave equation is given as

$$(1 + r^2)u_{n,m+1} = 2u_{n,m} - (1 + r^2)u_{n,m-1} + \frac{r^2}{2}u_{n+1,m+1}$$

$$+\frac{r^2}{2}u_{n-1,m+1} + \frac{r^2}{2}u_{n+1,m-1} + \frac{r^2}{2}u_{n-1,m-1}, \tag{11.5.19}$$

where $r = kc/h$ [see (11.3.7)]. We insert $u_{n,m} = \alpha^m \exp(i\beta n)$ into (11.5.19) and obtain the quadratic equation $\alpha^2 - (2/(1 + 2r \sin^2(\beta/2))\alpha + 1 = 0$, after some simplification. The discriminant of this equation is $4/(1 + 2r \sin^2(\beta/2))^2 - 4$, and it is clearly negative for all $r > 0$. Thus, the two roots are complex conjugates and the quadratic equation shows that the product of the two roots equals 1. Consequently, each root has unit modulus, so that we have $|\alpha| = 1$ for each root. We conclude that the backward difference scheme is *unconditionally stable*.

Exercises 11.5

11.5.1. Verify the von Neumann stability result obtained in Example 11.15.

11.5.2. Verify that the difference scheme (11.5.12) is unconditionally von Neumann unstable.

11.5.3. Verify that the difference scheme (11.5.14) is conditionally von Neumann stable.

11.5.4. Show that the implicit difference scheme (11.5.15) is unconditionally von Neumann stable.

11.5.5. Determine that the von Neumann stability conditions for the θ-difference scheme (11.2.36), as θ ranges from 0 to 1, are as given in the text.

11.5.6. Verify that the DuFort-Frankel difference scheme (11.5.16) is unconditionally von Neumann stable.

11.5.7. Show that the ω-difference scheme (11.3.7) (with $F = 0$) is unconditionally von Neumann stable if $1/4 < \omega \leq 1$. If $0 \leq \omega < 1/4$, the von Neumann stability condition is $r^2 \leq 1/(1 - 4\omega)$.

11.6 STABILITY AND CONVERGENCE OF MATRIX DIFFERENCE METHODS FOR PDES

The one-step explicit and implicit θ difference schemes for the nonhomogeneous *heat equation* presented in Section 11.2 can be expressed in matrix form [see (11.2.41)] as

$$\mathbf{u}(k+1) = \hat{A}\mathbf{u}(k) + \mathbf{w}(k, k+1), \qquad k \geq 0, \tag{11.6.1}$$

with the matrix $\hat{A}$ independent of k and $\mathbf{w}(k, k+1)$ as a known vector that depends on the nonhomogeneous term $F(x,t)$ and the boundary conditions given for the problem. If the heat equation and the boundary conditions are homogeneous, then $\mathbf{w}(k, k+1) = \mathbf{0}$. The vector $\mathbf{u}(0)$ represents the initial condition for the difference scheme. The difference scheme is said to be *matrix stable* if a solution with small initial conditions, as measured in some vector norm [say, $||\mathbf{u}(0)|| \leq \epsilon \ll 1$], remains small and does not grow without bound as $k \to \infty$.

Thus, a condition for stability is determined from the iteration problem

$$\mathbf{u}(k+1) = \hat{A}\mathbf{u}(k), \quad \text{or equivalently,} \quad \mathbf{u}(k+1) = \hat{A}^{k+1}\mathbf{u}(0). \tag{11.6.2}$$

To examine the growth of the iterations, we first assume that the $n \times n$ matrix $\hat{A}$ has n linearly independent eigenvectors $\mathbf{r}_m$ that correspond to the eigenvalues λ_m, with $m = 1, 2, \ldots, n$. Then $\mathbf{u}(0)$ can be expressed uniquely as a linear combination of the eigenvectors as

$$\mathbf{u}(0) = \sum_{m=1}^{n} c_m \mathbf{r}_m, \tag{11.6.3}$$

so that $\mathbf{u}(k+1) = \hat{A}^{k+1}\mathbf{u}(0) = \sum_{m=1}^{n} c_m \lambda_m^{k+1} \mathbf{r}_m$. It is apparent (no matter which vector norm is used) that unless $|\lambda_m| \leq 1$, $m = 1, 2, \ldots, n$, $\mathbf{u}(k)$ will grow without bound as $n \to \infty$. Consequently, we conclude that the scheme is *matrix stable* if the *spectral radius* $\rho(\hat{A}) \leq 1$, with the spectral radius defined as

$$\rho(\hat{A}) = \max(|\lambda_m|), \qquad m = 1, 2, \ldots, n. \tag{11.6.4}$$

(Here $|\lambda_m|$ represents the modulus of the eigenvalue λ_m, which may be complex valued.) If the matrix $\hat{A}$ has fewer than n linearly independent eigenvectors, it can be shown, using the *Jordan canonical form* for the matrix, that the modulus of at least one of the elements of $\hat{A}^k$ grows without bound as $k \to \infty$ unless $\rho(\hat{A}) \leq 1$ is satisfied. For if the similarity transformation $Q^{-1}\hat{A}Q = J$ yields the Jordan matrix J, we have $\hat{A}^k = Q^{-1}J^kQ$. Then the matrix J^k is an upper triangular matrix whose diagonal elements are the kth powers of all the eigenvalues of $\hat{A}$.

Since the matrix $\hat{A}$ incorporates the boundary conditions for the difference scheme, the matrix stability condition yields more precise results than the von Neumann stability condition, where boundary conditions play no role. For the most part, however, both methods yield similar results, and the von Neumann method is easier to apply. In addition, the matrix stability condition is defined for one-step difference schemes, so that two-step difference schemes such as the DuFort-Frankel method for the heat equation and the various difference methods for the wave equation appear to be excluded. Although it is possible to use an artifice to convert two or more step difference schemes to a one-step form, the analysis of the stability properties of the resulting schemes is not as straightforward as in the above. We do not present this approach. The von Neumann stability analysis can easily be carried out for two-step difference schemes, as we have seen.

The matrix representations of the Jacobi, Gauss-Seidel and SOR iteration methods for Poisson's equation considered in Section 11.4 can be written in the general form $\mathbf{u}(k+1) = \hat{A}\mathbf{u}(k) + \mathbf{b}$, $k \geq 0$. If the iteration scheme converges, it must converge to a solution $\mathbf{u}$ of the matrix equation $\mathbf{u} = \hat{A}\mathbf{u} + \mathbf{b}$. Thus, the error $\mathbf{e}(k)$ after the kth iteration is $\mathbf{e}(k) = \mathbf{u} - \mathbf{u}(k)$. Then

$$\mathbf{e}(k+1) = \mathbf{u} - \mathbf{u}(k+1) = \hat{A}\mathbf{u} - \hat{A}\mathbf{u}(k) = \hat{A}\mathbf{e}(k) = \hat{A}^{k+1}\mathbf{e}(0), \qquad (11.6.5)$$

where $\mathbf{u}(0)$ is the initial guess in the iteration method. For the iteration scheme to converge, we must have $\lim_{k\to\infty} \mathbf{e}(n) = 0$. This implies that $\lim_{k\to\infty} \hat{A}^k = 0$, and as shown above, the convergence condition is met if the *spectral radius* $\rho(\hat{A})$ satisfies the inequality $\rho(\hat{A}) < 1$.

It is apparent that the smaller the spectral radius, the more rapid the convergence of the iteration scheme. This concept is made concrete in the following. We assume that $\mathbf{e}(0)$ can be expressed as in (11.6.3) in terms of the eigenvectors $\mathbf{r}_m$ of the matrix $\hat{A}$. As a result, we can write $\mathbf{e}(k) = \rho(\hat{A})^k \sum_{m=1}^{n} c_m(\lambda_m/\rho(\hat{A}))^k\mathbf{r}_m$. On introducing a vector norm for $\mathbf{e}(k)$, we conclude that for large k, $||\mathbf{e}(k)|| \approx c\rho(\hat{A})^k$, where c is independent of k. Consequently, $||\mathbf{e}(k+1)||/||\mathbf{e}(k)|| \approx \rho(\hat{A})$. To determine the number of additional iterations m that are needed to reduce the norm of the error (approximately) by 1/10, we put $||\mathbf{e}(k+m)||/||\mathbf{e}(k)|| \approx 1/10$ and note that $||\mathbf{e}(k+m)||/||\mathbf{e}(k)|| \approx \rho(\hat{A})^m$. This yields $\rho(\hat{A})^m \approx 1/10$. Taking the log (to the base 10) of both sides, we obtain $m \approx -1/\log(\rho(\hat{A}))$. If $\rho(\hat{A}) \leq 1$, its log is negative, so that the closer $\rho(\hat{A})$ is to 1, the larger the value of m and the slower the convergence of the iteration scheme. The value of $-\log(\rho(\hat{A}))$ is called the *convergence rate* for the iteration scheme.

The *matrix stability condition* for the heat equation requires that $\rho(\hat{A}) \leq 1$ and the *convergence condition* for the iteration methods for Laplace's and Poisson's equations requires that $\rho(\hat{A}) < 1$. For either case, it is not necessary to determine all the eigenvalues of the matrix $\hat{A}$ explicitly. Only the eigenvalue with largest modulus must be found or, more simply, we must show that its modulus is less than or equal to 1 or strictly less than 1. A simple method for determining these bounds on the largest (in modulus) eigenvalue makes use of *Gershgorin's disk (or circle) theorem.* It states that for the $n \times n$ matrix $\hat{A}$ with elements $\hat{a}_{ij}$, every eigenvalue lies within one of the disks in the complex λ-plane with center and radius, Center $= \hat{a}_{ii}$, Radius $= \sum_{j=1, j\neq i}^{n} |\hat{a}_{ij}|$, $i = 1, 2, \ldots, n$. All the eigenvalues lie in the union of these disks. The proof of this result is straightforward. Suppose that λ is an eigenvalue of $\hat{A}$. Then there exists a nonzero (column) eigenvector $\mathbf{x} = [x_1, x_2, \ldots, x_n]^T$ such that $\hat{A}\mathbf{x} = \lambda\mathbf{x}$. In component form we have $(\lambda - \hat{a}_{ii})\, x_i = -\sum_{j=1, j\neq i}^{n} \hat{a}_{ij} x_j$, $i = 1, 2, \ldots, n$. Pick a value of i such that $|x_j| \leq |x_i|$, $j = 1, 2, \ldots, n$, $j \neq i$. Then $|\lambda - \hat{a}_{ii}| \leq \sum_{j=1, j\neq i}^{n} |\hat{a}_{ij}||x_j|/|x_i| \leq \sum_{j=1, j\neq i}^{n} |\hat{a}_{ij}|$. This completes the proof.

A simple application of Gershgorin's theorem can be made to determine a matrix stability condition for the explicit forward difference scheme for the heat equation with Dirichlet conditions. The matrix B defined in (11.2.12) is relevant to that problem. Each diagonal element of B is $1 - 2r$ and the sum of the off-diagonal row elements equals r for the first and last row and $2r$ for the remaining rows. (Each off-diagonal element is nonnegative.) This yields two Gershgorin disks $|\lambda - (1 - 2r)| \leq r$ and $|\lambda - (1 - 2r)| \leq 2r$. Since one disk is contained within the other disk, we find that all the eigenvalues must lie within $|\lambda - (1 - 2r)| \leq 2r$, and this is equivalent to $1 - 4r \leq \lambda \leq 1$. As we must have $|\lambda| \leq 1$ for each eigenvalue to guarantee stability, we conclude that $1 - 4r \geq -1$ or $r \leq 1/2$ is the condition that ensures the matrix stability of the difference scheme. This restriction on r agrees with that obtained using the von Neumann stability criterion. We remark that the Gershgorin result does not exclude the possibility that $|\lambda| \leq 1$ if a larger disk than the foregoing is selected with a correspondingly larger value of r. The eigenvalues of the matrix B can be determined explicitly, and it then follows that the spectral radius condition $\rho(B) \leq 1$ requires that $r \leq 1/2$.

Matrix Stability for the Heat Equation

The θ difference scheme (11.2.36) (with $0 \leq \theta \leq 1$) represents all the one-step schemes for the heat equation introduced in Section 11.2. It was seen to reduce to the explicit forward, implicit backward and Crank-Nicolson schemes if $\theta = 0$, $\theta = 1$, and $\theta = 1/2$, respectively.

In the case of *Dirichlet boundary conditions*, the matrix representation of the θ scheme is given in (11.2.41) with the matrices A and B defined in (11.2.38) and (11.2.39). Using the special form of the matrices A and B, the eigenvalues of the matrix $A^{-1}B$ can be shown to be

$$\lambda_m = 1 - \frac{4r \sin^2(\pi m/(2(n+1)))}{1 + 4\theta r \sin^2(\pi m/(2(n+1)))}, \qquad m = 1, 2, \ldots, n. \tag{11.6.6}$$

We must determine a condition on r such that $-1 \leq \lambda_m \leq 1$ for all the eigenvalues. Now the term $\sin^2(\pi m/(2(n+1))) > 0$ for all $1 \leq m \leq n$. As a result, each $\lambda_m < 1$ for $r > 0$ and $\theta \geq 0$. To guarantee that all the $\lambda_m \geq -1$ we must have $(1-2\theta)\, 4r \sin^2(\pi m/(2(n+1))) \leq 2$. Clearly, if $1/2 \leq \theta \leq 1$, this condition is met for all r. If $0 \leq \theta < 1/2$, we have $1 - 2\theta > 0$, so that we obtain $r \leq 1/(2-4\theta)$. These conditions on r guarantee that the spectral radius $\rho(A^{-1}B) \leq 1$. As a result, we have the following matrix stability conditions for the θ scheme in the case of Dirichlet boundary conditions. If $1/2 \leq \theta \leq 1$, it is *unconditionally stable*. This includes the implicit backward and Crank-Nicolson schemes. If $0 \leq \theta < 1/2$ it is *conditionally stable* with the stability condition $r \leq 1/(2-4\theta)$. With $\theta = 0$ we obtain the explicit forward difference scheme and the conditional stability condition $r \leq 1/2$.

We do not exhibit general formulas for the eigenvalues of $A^{-1}B$ for the θ scheme with general boundary conditions. Such formulas can be found if the matrices A and B can be expressed as linear combinations of simple matrices whose eigenvalues can be determined easily, as is the case for Dirichlet conditions. Instead, we use the Maple procedure $HeatMatrix$, which constructs the matrix form of a specific θ scheme and exhibits the corresponding matrices A and B. Then the Maple procedure $HeatStability$ uses these matrices to determine the spectral radius $\rho(A^{-1}B)$ by finding the eigenvalues of $A^{-1}B$. Thereby, the stability or instability of specific schemes can be established.

The first procedure is given as $HeatMatrix(c^2, F(x,t), t = t_0, f(x), x = a..b, bcl, g(t), bcr, s(t), n, k, m, backward, \theta)$. The first 11 arguments are the same as those given for the procedures $NumHeatForw$ and $NumHeatBackw$ in Section 11.2, with the exception that only an initial time $t = t_0$ is entered. (We do not repeat the description of the arguments here.) The twelfth argument m prescribes the number of t steps in the scheme. If the thirteenth argument is *backward*, the θ scheme is employed with the value of θ prescribed in the fourteenth argument. The selection $\theta = 0$ in the fourteenth argument determines that the explicit forward difference scheme is used. The same result can be achieved by entering $forward$ as the 13th argument and deleting the fourteenth argument. The output of the procedure displays the related initial and boundary value problem for the heat equation together with the values of k, h, and $r = kc^2/h^2$, as is the case for $NumHeatBackw$, say. In addition, the matrix form of the difference scheme is displayed with r left unspecified. The final output of the procedure is the matrices A and B. (For the explicit forward scheme, A is always the identity matrix.)

Once A and B are determined by the procedure $HeatMatrix$, the procedure $HeatStability$ determines the spectral radius of the matrix $A^{-1}B$ by finding the eigenvalues of the matrix. It has the form $HeatStability(A, B)$ or $HeatStability(B)$. The latter form can be used if A is the identity matrix. If Maple can determine the relevant eigenvalues, and the output of the procedure, which is the value of the spectral radius, is less than or equal to 1, the schemes is stable. Otherwise, it is unstable.

Example 11.22. The θ Difference Scheme: Dirichlet Conditions. The Maple procedure $HeatMatrix(c^2, F(x,t), t = 0, f(x), x = a..b, dirichlet, g(t), dirichlet, h(t), 5, k, 1, backward, \theta)$ applies the θ difference scheme to the inhomogeneous heat equation (11.2.1) over the interval $[a, b]$ with the initial condition $u(x, 0) = f(x)$ and the Dirichlet boundary conditions $u(a, t) = g(t)$, $u(b, t) = h(t)$. The x interval is divided into five equal parts and the time increment equals k. For this problem we have $r = 25c^2k/(a - b)^2$.

The global variables LM and RM for the procedure determine the matrices A and B, and they agree with forms given in (11.2.38)–(11.2.39). (We are only interested in obtaining the matrices A and B in the matrix representation of the scheme.) On using Maple to determine the inverse of the matrix A and to obtain the eigenvalues of the matrix $A^{-1}B$ (invoking the procedure $Eigenvalues$ in the $LinearAlgebra$ package) we find that the eigenvalues coincide with those given in (11.6.6), where we must put $n = 4$.

We specialize the results by putting $c = 1$, $a = 0$, $b = 1$ and assigning specific values of θ in the procedure. If we put $\theta = 0$, the *explicit forward difference scheme* results. The matrices A and B are determined by putting $\theta = 0$ in (11.217)–(11.218) and A is the identity matrix. The output of $HeatStability(B)$ is $max(|3.618r - 1|, |1.382r - 1|, |0.38197r - 1|, |2.618r - 1|)$. This states that the spectral radius $\rho(B)$ of B is the maximum of the absolute values of its four (listed) eigenvalues, and this maximum depends on the choice of r. The use of the Maple procedure $solve(\rho(B) \leq 1)$ shows that forward difference scheme is *stable* if $r \leq 0.55279$. (The stability restriction $r \leq 1/2$ determined above applies as the number of x subdivisions increases without bound. If we increase the subdivisions for this problem, the upper bound for r decreases.)

Next, we put $\theta = 1/2$ in the procedure and obtain the *Crank-Nicolson difference scheme*. The matrices A and B are determined by putting $\theta = 1/2$ in the foregoing results. The output of $HeatStability(A, B)$ (which we do not display) states that the spectral radius $\rho(A^{-1}B)$ of $A^{-1}B$ is the maximum of the absolute values of its four eigenvalues, which again depends on the choice of r. The use of the Maple procedure $solve(\rho(A^{-1}B) \leq 1)$ determines that this inequality is satisfied for all $r > 0$. As a result the Crank-Nicolson scheme is *unconditionally stable*.

Finally, we put $\theta = 1/4$ and obtain an *implicit difference scheme*. The matrices A and B are determined by putting $\theta = 1/4$ in the matrices LM and RM. The output of $HeatStability(A, B)$ (which we do not display) states that the spectral radius $\rho(A^{-1}B)$ of $A^{-1}B$ is the maximum of the absolute values of its four eigenvalues, which again depends on the choice of r. The use of the Maple procedure $solve(\rho(A^{-1}B) \leq 1)$ determines that this inequality is satisfied if $r \leq 1.10557$. For this range of r, the $\theta = 1/4$ difference scheme is *stable*. Again, the upper bound on r exceeds the stability bound 1 given above in our general discussion.

Example 11.23. The θ Difference Scheme: Neumann and Robin Boundary Conditions. The procedure $HeatMatrix(c^2, F(x,t), t = 0, f(x), x = 0..l, \alpha, -g(t), \beta, h(t), 5, k, 1, backward, \theta)$ applies the θ difference scheme to the inhomogeneous heat equation (11.2.1) over the interval $[0, l]$ with the initial condition $u(x,0) = f(x)$ and the boundary conditions $-u_x(0,t)+\alpha\, u(0,t) = -g(t)$, $u_x(l,t)+\beta\, u(l,t) = h(t)$. On setting $\alpha = \beta = 0$, we obtain Neumann boundary conditions. The x interval is divided into five equal parts and the time increment equals k. For this problem $r = 25c^2k/l^2$. We are only interested in obtaining the matrices A and B in the matrix representation of the scheme. The global variables LM and RM determine these matrices to have the following form, which we denote by C:

$$\begin{bmatrix} 1+2r\hat{\theta}(1+\frac{\alpha l}{5}) & -2r\hat{\theta} & 0 & 0 & 0 & 0 \\ -r\hat{\theta} & 1+2r\hat{\theta} & -r\hat{\theta} & 0 & 0 & 0 \\ 0 & -r\hat{\theta} & 1+2r\hat{\theta} & -r\hat{\theta} & 0 & 0 \\ 0 & 0 & -r\hat{\theta} & 1+2r\hat{\theta} & -r\hat{\theta} & 0 \\ 0 & 0 & 0 & -r\hat{\theta} & 1+2r\hat{\theta} & -r\hat{\theta} \\ 0 & 0 & 0 & 0 & -2r\hat{\theta} & 1+2r\hat{\theta}(1+\frac{\beta l}{5}) \end{bmatrix}. \tag{11.6.7}$$

If we set $\hat{\theta} = \theta$, we have $C = A$, whereas if $\hat{\theta} = \theta - 1$, we have $C = B$. With $\alpha = 0$ and $\beta = 0$, the matrices assume the forms appropriate for the Neumann problem.

As in Example 11.22, we specialize the results by putting $c = 1$, $a = 0$, $b = 1$ and assigning specific values of θ, α, and β in the procedure. With $\theta = 0$, $\alpha = 0$, $\beta = 0$, the *explicit forward difference scheme* for the Neumann problem results. We insert these values in (11.6.7) and determine the matrices $A = I$ and B for this problem. $HeatStability(B)$ yields the spectral radius $\rho(A^{-1}B)$ as $max([1., |.38197r - 1|, |1.3820r - 1|, |4r - 1|, |3.6180r - 1|, |2.6180r - 1|)$. The use of $solve(\rho(A^{-1}B) \leq 1)$ determines that this inequality is satisfied for $r \leq 1/2$, and this is the condition for stability.

If we put $\theta = 1$, $\alpha = 0$, $\beta = 0$, the *implicit backward difference scheme* for the Neumann problem results. Then, on using the appropriate matrices A and B, we obtain as the output of $HeatStability(A, B)$, $\rho(A^{-1}B) = max(1., 1./|1 + 4r|, .5|.7639r + 2|/|r^2 + 3r + 1|, .5|7.236r + 2|/|5r^2 + 5r + 1|, .5|2.763r + 2|/|5r^2 + 5r + 1|, .5|5.236r + 2|/|r^2 + 3r + 1|)$. $solve(|\lambda_i| \leq 1)$ applied for each of the eigenvalues shows that $\rho(A^{-1}B) \leq 1$ for all $r > 0$, so the scheme is *unconditionally stable*. Setting $\theta = 1/2$ in the above yields the *Crank-Nicolson scheme*. The spectral radius is given as $max(1, 1 - 2r/|1 + 2r|, .5|2r^2| \pm 8.9443r - 8|/|r^2 + 6r + 4|, .5|10r^2 \pm 8.9443r - 8|/|5r^2 + 10r + 4|)$, so $\rho(A^{-1}B) \leq 1$ for all $r > 0$ and the scheme is *unconditionally stable*.

The choice $\theta = 0$, $\alpha = 1$, $\beta = 1$ yields the *explicit forward difference scheme* for a Robin boundary value problem. We insert these values into (11.6.7) to determine the matrix B. (The matrix A is the identity matrix.) The simplest method for determining the stability condition in this case is to determine the eigenvalues of the matrix B directly using the Maple procedure *eigenvals* from the *linalg* package. (The output from the procedure yields complex-valued eigenvalues with small imaginary parts,

which we equate to zero.) The list of absolute values of the eigenvalues is found to be $[|.068558r - 1|, |3.7934r - 1|, |1.5381r - 1|, |.52748r - 1|, |4.0943r - 1|, |2.7782r - 1|]$. We use *solve* to determine the maximum value of r for which the absolute value of each eigenvalue is less than or equal to 1. This gives $r \leq 0.48847$ as the stability condition for the difference scheme.

We do not consider backward difference schemes for problems with Robin boundary conditions.

Convergence of Matrix Iteration Methods for Laplace's and Poisson's Equations

The matrix forms of the *Jacobi, Gauss-Seidel,* and *SOR iteration methods* for *Poisson's equation* considered in Section 11.4 have the general form

$$\mathbf{u}(k+1) = \hat{A}\mathbf{u}(k) + \mathbf{b}, \qquad k \geq 0, \tag{11.6.8}$$

and the error was shown in (11.6.5) to be given as

$$\mathbf{e}(k+1) = \mathbf{u} - \mathbf{u}(k+1) = \hat{A}^{k+1}\mathbf{e}(0). \tag{11.6.9}$$

The iteration scheme converges if the spectral radius of $\hat{A}$ is less than 1. An alternative condition for the convergence of the iteration scheme makes use of vector and matrix norms.

The three basic *vector norms* are the L_1, the Euclidean L_2, and the maximum L_∞ norms. They are defined as

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad ||\mathbf{x}||_1 = \sum_{i=1}^{n} |x_i|, \quad ||\mathbf{x}||_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2}, \quad ||\mathbf{x}||_\infty = \max_{1 \leq i \leq n} |x_i|, \tag{11.6.10}$$

respectively. The corresponding *induced matrix norms* for an $n \times n$ matrix A with elements a_{ij}, can be shown to be given as

$$||A||_1 = \max_{1 \leq j \leq n} \sum_{i=1}^{n} |a_{ij}|, \; ||A||_2 = \sqrt{\rho(A^T A)}, \; ||A||_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^{n} |a_{ij}|, \tag{11.6.11}$$

where $\rho(A^T A)$ is the spectral radius of $A^T A$. (For example, the matrix norm $||A||_1$ corresponds to the vector norm $||\mathbf{x}||_1$.) For each norm we have the inequality $||A\mathbf{x}|| \leq ||A||||\mathbf{x}||$. We apply this inequality to (11.6.9) and obtain $||\mathbf{e}(k+1)|| \leq ||\hat{A}||^{k+1}||\mathbf{e}(0)||$. Thus, if $||\hat{A}|| < 1$ for any of the three norms, the iteration scheme converges. This is a useful result because it is generally easier to determine the L_1 or L_∞ norm of a

matrix than its spectral radius. Nevertheless, we try to find the spectral radius $\hat{A}$ of the iteration scheme because it gives the sharpest result and determines the convergence rate of the scheme.

The procedure $LaplaceMatrix$ determines the matrix form $A\mathbf{u} = \mathbf{w}$ for each of the finite difference schemes for *Poisson's equation* presented in Section 11.4. The output of the procedure is the matrix A. (The global variable *CoeffMatrix* displays this matrix.) Then $LaplaceConvergence(A, Method)$ constructs the matrix $\hat{A}$ in the equation $\mathbf{u}(k+1) = \hat{A}\mathbf{u}(k) + \mathbf{w}$ that occurs in the *Jacobi, Gauss-Seidel,* or *SOR iteration methods* for the given difference scheme. The second argument, $Method$, in the procedure must be *Jacobi, Gauss-Seidel,* or *SOR*. If it is *SOR*, the relaxation parameter ω must be added as a third argument. It must satisfy $0 < \omega < 2$. The procedure finds and displays the spectral radius of $\hat{A}$. If an arbitrary third argument is added for the *Jacobi* and *Gauss-Seidel* cases, and an arbitrary fourth argument is added in the *SOR* case, the L_1, L_2, and L_∞ norms of the matrix and the convergence rate of the iteration scheme are also displayed.

Finally, we have constructed a procedure $OptimalSOR(A, \omega_I, Increment)$ that determines an (approximate) optimal value of the parameter ω in the *SOR scheme*. This is accomplished by iterating the $LaplaceConvergence$ procedure. The first argument in the procedure is the matrix A, which is the output of the $LaplaceMatrix$ procedure. A (trial) value of the relaxation parameter ω, denoted by ω_I, is entered, together with the value of the $Increment$, which we refer to as s. Then a sequence of $LaplaceConvergence$ procedures is invoked, using the SOR scheme in which the relaxation parameter takes the set of values $\omega_I + is$, $i = 1, 2, \ldots$. This yields a sequence of spectral radii. The iteration ends when a newly determined spectral radius has a greater value than the one that precedes it in the sequence. When this happens, the spectral radius, the convergence rate, and the final (optimal) value of the relaxation parameter ω_F are displayed. To achieve the best results most rapidly, a larger value of s can be used on applying the procedure initially, say $s = 0.1$. Once the first approximate optimal value ω_F is found, the value $\omega_F - s$ with a new smaller value of the increment, say $s = 0.01$, is entered in the procedure. This process can then be iterated in an obvious way, to determine increasingly improved approximations to the optimal relaxation value.

Example 11.24. A Dirichlet Problem for Laplace's Equation. The problem (11.4.14)–(11.4.15) given in Example 11.10 was solved via iteration in Example 11.13. We apply $LaplaceMatrix(0, x = 0..1, dirichlet, 1, dirichlet, 0, y = 0..1, dirichlet, 0, dirichlet, 0, 4, 4)$ to determine the coefficient matrix A for the difference scheme for this problem. The output of the procedure or of *CoeffMatrix* gives the matrix A, which we do not display. Next, we apply the procedure $LaplaceConvergence$ to the matrix A. The use of $LaplaceConvergence$ $(A, Jacobi, r)$ yields $||\hat{A}||_1 = 1, ||\hat{A}||_2 = 0.70711, ||\hat{A}||_\infty = 1$, Spectral Radius $=$ 0.70711, Convergence Rate $= 0.15051$ for the matrix $\hat{A}$ that arises in the Jacobi iteration scheme. (The arbitrary third argument in the procedure, r, causes Maple to display the additional information exhibited here.) Then, $LaplaceConvergence$ $(A, Gauss - Seidel, r)$ gives, for the Gauss-Seidel method, $||\hat{A}||_1 = 0.86719$,

$||\hat{A}||_2 = 0.61070$, $||\hat{A}||_\infty = 0.81250$, Spectral Radius = 0.5, Convergence Rate = 0.30103. It is seen that the convergence rate for the Gauss-Seidel method is twice that for the Jacobi method. (This is the case in general.) Finally, $LaplaceConvergence\ (A, SOR, 1.2, r)$ gives, for the SOR method, $||\hat{A}||_1 = 0.8546$, $||\hat{A}||_2 = 0.57638$, $||\hat{A}||_\infty = 0.8600$, Spectral Radius = 0.2, Convergence Rate = 0.69897, with the relaxation parameter $\omega = 1.2$. The convergence rate for the SOR method is more than twice that of the Gauss-Seidel method. The foregoing results for the convergence rates are consistent with those obtained in Example 11.13, where the number of iterations required for each of the three schemes (for a given error tolerance) was obtained. We observe that for each case, the spectral radius is less than or equal to each of the three matrix norms.

To find an optimal value for the relaxation parameter in the SOR method we use the procedure $OptimalSOR$. We begin with $\omega_I = 1.2$. The output of $OptimalSOR\ (A, 1.2, 0.01)$ gives 1.2 as the value of the relaxation parameter. Thus, the optimal value is smaller than 1.2. With $\omega_I = 1.1$, $OptimalSOR(A, 1.1, .01)$ yields $\omega_F = 1.18$. Next, $OptimalSOR(A, 1.17, .001)$ gives $\omega_F = 1.172$. Finally, $OptimalSOR(A, 1.171, .0001)$ finds that Spectral Radius = 0.17160, Convergence Rate = 0.76548, and $\omega_F = 1.1716$. The convergence rate is now five times as great as that for the Jacobi method and more than two and one-half times great as that for the Gauss-Seidel method.

Exercises 11.6

11.6.1. Verify the numerical results of Example 11.22 using $HeatMatrix$ and $HeatStability$.

11.6.2. Verify the numerical results of Example 11.23 using $HeatMatrix$ and $HeatStability$.

11.6.3. Verify the numerical results of Example 11.24 using $LaplaceMatrix$, $LaplaceConvergence$, and $OptimalSOR$.

11.7 FINITE DIFFERENCE METHODS FOR FIRST ORDER HYPERBOLIC EQUATIONS AND SYSTEMS

Linear first order scalar PDEs and the exact and numerical solution of initial value and initial and boundary value problems for these PDEs have been considered a number of times in this book. In Chapter 1 it was shown that certain random walks are related to a class of finite difference schemes for the numerical solution of initial value problems for first order PDEs. In Chapter 2, numerical solution methods for first order scalar PDEs based on the method of characteristics were presented. First order hyperbolic systems of PDEs were studied in Chapter 3. Here we construct a number of difference

schemes for first order scalar PDEs, with constant coefficients, of the form

$$u_t(x,t) + au_x(x,t) = bu(x,t) + F(x,t), \tag{11.7.1}$$

and systems of PDEs in matrix form with constant coefficient matrices

$$\mathbf{u}_t(x,t) + A\mathbf{u}_x(x,t) = B\mathbf{u}(x,t) + \mathbf{F}(x,t). \tag{11.7.2}$$

The scalar equation (11.7.1) can be solved exactly, as was shown in Section 2.2, so that there appears to be no need for numerical solution methods. Nevertheless, it useful to develop such methods for the insight it yields into the construction of corresponding methods for equations with variable coefficients. The numerical method of characteristics can be used for these PDEs, but the results do not determine the solution over a regular grid in (x,t)-space without additional effort. In the case of systems with constant coefficients, it is rarely possible to obtain exact solutions.

A number of difference schemes will be constructed for (11.7.1)–(11.7.2) by replacing derivatives by difference quotients and introducing some modifications of the resulting expressions. The forward, backward, and centered differences defined in Section 11.1 will be used.

First Order Scalar PDEs

We assume in the following that k is the t increment and h is the x increment in the difference approximations. Then, on approximating $u_t(x,t)$ by a forward difference quotient and $u_x(x,t)$ by a backward difference quotient in (11.7.1), we obtain

$$u(x,t+k) = \left(1 - \frac{ka}{h} + kb\right)u(x,t) + \left(\frac{ka}{h}\right)u(x-h,t) + kF(x,t). \tag{11.7.3}$$

Next, we determine the *truncation error* that results on replacing (11.7.1) by the difference equation (11.7.3). Applying the Maple procedure *mtaylor* to (11.7.3) gives

$$u_t(x,t) + au_x(x,t) - bu(x,t) - F(x,t) \approx -\frac{1}{2}u_{tt}k - \frac{1}{6}u_{ttt}k^2 + \frac{a}{2}u_{xx}h - \frac{a}{6}u_{xxx}h^2. \tag{11.7.4}$$

Thus, the *truncation error* is of $O(k+h)$ and the difference equation is *consistent* with the first order PDE (11.7.1).

With $(x_i, t_j) = (ih, jk)$ we can write the difference equation (11.7.3) as

$$u_{i,j+1} = (1 - r + kb)u_{i,j} + ru_{i-1,j} + kF_{i,j}, \tag{11.7.5}$$

where $r = ka/h$. This is known as the *explicit forward-backward difference scheme* for the numerical solution of (11.7.1). The value of $u_{i,j+1}$ is known once $u_{i,j}$ and $u_{i-1,j}$ are determined. For an initial value problem for (11.7.1) with $u(x,0) = f(x)$, we have $u_{i,0} = f(x_i)$. If a boundary condition $u(0,t) = h(t)$ is prescribed at $x = 0$,

and the solution is to be found for $x > 0$, we have $u_{0,j} = h(t_j)$. In either case, the $u_{i,j}$, $j \geq 1$, can be determined step by step at the grid values for the problem.

If $a > 0$ in (11.7.1), the characteristic that passes through the point $(x_i, t_j) = (ih, jk)$ is the straight line $x - at = x_i - at_j = ih - ajk$. It intersects the x-axis at the point $x = x_i - at_j = ih - ajk$, which lies to the left of x_i. Thus, the *domain of dependence* of the solution of the initial value problem for the PDE (11.7.1) at $t = t_j$ with $x \geq x_i$ is given by the set of points in the (closed) region bounded on the left by the characteristic line, below by the interval $[x_i - at_j, \infty)$ on the x-axis, and above by the interval $[x_i, \infty)$ on the line $t = t_j$.

The corresponding *domain of dependence* of the solution $u_{i,j}$ of the *explicit forward-backward difference scheme* (11.7.3) at $t = t_j$ with $x \geq x_i$ is readily found to be set of grid points in the (closed) region bounded on the left by the line $x - ht/k = x_i - ht_j/k = ih - hj$, below by the interval $[x_i - ht_j/k, \infty)$ on the x-axis, and above by the interval $[x_i, \infty)$ on the line $t = t_j$. We see that if $-ht_j/k \leq -at_j$ or, equivalently, $r = ak/h \leq 1$, the numerical domain of dependence includes (or coincides with) the domain of dependence for the PDE. If $r > 1$, the reverse is true. The restriction $r \leq 1$ is the *CFL stability condition* or the *Courant stability condition* for this problem. It signifies that the numerical solution must make use of all the data that is required for the determination of the exact solution over the given interval; otherwise, the numerical solution will fail to converge to the exact solution. The domains of dependence for the PDE and numerical scheme in the unstable case are displayed in Figure 11.6.

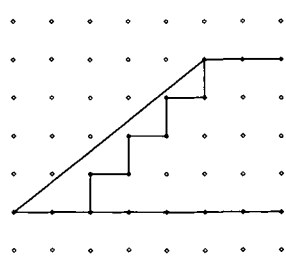

Figure 11.6 The unstable case: $r > 1$.

The *CFL stability condition* $r \leq 1$ agrees with the result obtained for the random walk characterized by (1.4.7), which corresponds to the foregoing difference scheme if we set $F(x,t) = 0$ and $b = 0$ in the problem. The von Neumann stability condition, to be obtained below, is also in agreement.

Next, we approximate both $u_t(x,t)$ and $u_x(x,t)$ by forward difference quotients in (11.7.1) and obtain

$$u(x,t+k) = \left(1 + \frac{ka}{h} + kb\right) u(x,t) - \left(\frac{ka}{h}\right) u(x+h,t) + kF(x,t). \quad (11.7.6)$$

The *truncation error* that results on replacing (11.7.1) by (11.7.6) is given as (on using the Maple procedure *mtaylor*) $u_t(x,t) + au_x(x,t) - bu(x,t) - F(x,t) \approx -(1/2)u_{tt}k - (1/6)u_{ttt}k^2 - (a/2)u_{xx}h - (a/6)u_{xxx}h^2$. Thus, the *truncation error*

is of $O(k + h)$ and the difference equation is *consistent* with the first order PDE (11.7.1).

Setting $(x_i, t_j) = (ih, jk)$, we can write the difference equation (11.7.6) as

$$u_{i,j+1} = (1 + r + kb)u_{i,j} - ru_{i+1,j} + kF_{i,j}, \tag{11.7.7}$$

where $r = ka/h$. This is the *explicit forward-forward difference scheme* for the numerical solution of (11.7.1). The value of $u_{i,j+1}$ is known once $u_{i,j}$ and $u_{i+1,j}$ are determined. For an initial value problem for (11.7.1) with $u(x, 0) = f(x)$, we have $u_{i,0} = f(x_i)$. If a boundary condition $u(x_N, t) = h(t)$ is prescribed at $x = x_N = Nh$, $N > 0$, and the solution is to be found for $x < x_N$, we have $u_{N,j} = h(t_j)$. In either case, the $u_{i,j}$, $j \geq 1$, can be determined step by step at the grid values for the problem.

If $a < 0$ in (11.7.1), the characteristic that passes through the point $(x_i, t_j) = (ih, jk)$ is a straight line $x - at = x_i - at_j = ih - ajk$ that intersects the x-axis at the point $x = x_i - at_j = ih - ajk$, which lies to the right of x_i. Thus, the *domain of dependence* of the solution of the initial value problem for the PDE (11.7.1) at $t = t_j$ with $x \leq x_i$ is given by the set of points in the (closed) region bounded on the right by the characteristic line, below by the interval $(-\infty, x_i - at_j]$ on the x-axis, and above by the interval $(-\infty, x_i]$ on the line $t = t_j$.

The corresponding *domain of dependence* of the solution $u_{i,j}$ of the *explicit forward-forward difference scheme* (11.7.3) at $t = t_j$ with $x \leq x_i$ is readily found to be set of grid points in the (closed) region bounded on the right by the line $x + ht/k = x_i + ht_j/k = ih + hj$, below by the interval $(-\infty, x_i + ht_j/k]$ on the x-axis, and above by the interval $(-\infty, x_i]$ on the line $t = t_j$. We see that if $ht_j/k \geq -at_j$ or, equivalently, $r = ak/h \geq -1$, the numerical domain of dependence includes (or coincides with) the domain of dependence for the PDE. If $r < -1$, the reverse is true. Thus, the *CFL stability condition* for this problem is $r \geq -1$. The domains of dependence for the PDE and numerical scheme in the unstable case are displayed in Figure 11.7.

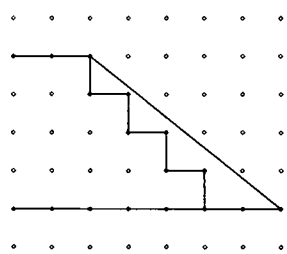

Figure 11.7 The unstable case: $r < -1$.

The *CFL stability condition* $r \geq -1$ agrees with the result obtained for the random walk characterized by (1.4.8), which corresponds to our difference scheme if we set $F(x, t) = 0$ and $b = 0$ in our problem. The von Neumann stability condition, to be obtained below, is also in agreement.

If we approximate $u_t(x,t)$ by a forward difference and $u_x(x,t)$ by a centered difference in (11.7.1), we obtain

$$u(x,t+k) = (1+kb)\,u(x,t) - \frac{ka}{2h}(u(x+h,t) - u(x-h,t)) + kF(x,t). \quad (11.7.8)$$

The *truncation error* that results on replacing (11.7.1) by the difference equation (11.7.8) is obtained on applying the Maple procedure *mtaylor* to (11.7.8) which gives

$$u_t(x,t) + au_x(x,t) - bu(x,t) - F(x,t) \approx -\frac{1}{2}u_{tt}k - \frac{1}{6}u_{ttt}k^2 - \frac{a}{6}u_{xxx}h^2. \quad (11.7.9)$$

Thus, the *truncation error* is of $O(k+h^2)$ and the difference equation is *consistent* with the first order PDE (11.7.1).

With $(x_i, t_j) = (ih, jk)$ we can write the difference equation (11.7.8) as

$$u_{i,j+1} = (1+kb)u_{i,j} - r/2(u_{i+1,j} - u_{i-1,j}) + kF_{i,j}, \quad (11.7.10)$$

where $r = ka/h$. This is the *explicit forward-centered difference scheme* for the numerical solution of (11.7.1). The value of $u_{i,j+1}$ is known once $u_{i,j}$, $u_{i-1,j}$, and $u_{i+1,j}$ are determined. For an initial value problem for (11.7.1) with $u(x,0) = f(x)$, we have $u_{i,0} = f(x_i)$.

It follows from the above that *CFL stability condition* for this difference scheme is $r \leq 1$. If we set $F = 0$ and $b = 0$ in the foregoing difference scheme, it has no random walk interpretation since one of the coefficients in the difference equation is negative whether a is positive or negative. A von Neumann stability analysis shows that the scheme is *unconditionally unstable*. As a result, we conclude that the *CFL stability condition* can at best serve as a necessary condition for stability but is not sufficient to guarantee stability.

The explicit forward-centered difference scheme can be modified to yield a conditionally stable difference scheme. In the forward difference approximation of $u_t(x,t)$ we replace $u(x,t)$ by $[u(x+h,t) + u(x-h,t)]/2$. Then

$$u(x,t+k) = kbu(x,t) + \left(\frac{1}{2} - \frac{ka}{2h}\right)u(x+h,t) + \left(\frac{1}{2} + \frac{ka}{2h}\right)u(x-h,t)) + kF(x,t). \quad (11.7.11)$$

The *truncation error* that results on replacing (11.7.1) by (11.7.11) is of $O(k+h^2)$, so that the difference equation is *consistent* with (11.7.1).

With $(x_i, t_j) = (ih, jk)$ we can write the difference equation (11.7.11) as

$$u_{i,j+1} = kbu_{i,j} + \frac{1}{2}(1-r)u_{i+1,j} + \frac{1}{2}(1+r)u_{i-1,j} + kF_{i,j}, \quad (11.7.12)$$

where $r = ka/h$. This is the *explicit Lax-Friedrichs difference scheme* for the numerical solution of (11.7.1). Proceeding as before, we find that the *CFL stability condition* is $|r| \leq 1$. If we set $F = 0$ and $b = 0$ in the scheme, it agrees with the

random walk characterized by (1.4.9). It was shown that we must have $|r| \leq 1$ for the validity of the random walk interpretation of (1.4.9). This result agrees with the von Neumann stability condition to be obtained below, so that the scheme is *conditionally stable.*

A difference scheme with a reduced truncation error can be obtained by proceeding as follows. We use the Taylor series to obtain the approximation $u(x, t + k) \approx u(x, t) + u_t(x, t)k + (1/2)u_{tt}(x, t)k^2$ and differentiate the PDE (11.7.1) to obtain $u_{tt}(x, t) = bu_t(x, t) + F_t(x, t) - au_{xt}(x, t)$. Substituting the foregoing equation and (11.7.1) into a Taylor approximation yields $u(x, t+k) \approx (1+kb+k^2b^2/2)u(x, t) - (ka + k^2ab)u_x(x, t) + (k^2a^2/2)u_{xx}(x, t) + (k + k^2b/2)F(x, t) + (k^2/2)F_t(x, t) - (k^2a/2)F_x(x, t)$. If we approximate $u_x(x, t)$ and $u_{xx}(x, t)$ in the above by centered differences, we obtain

$$u(x, t + k) = (1 - r^2 + kb + k^2b^2/2)u(x, t) + (-r/2 + r^2/2 - krb/2)u(x + h, t)$$

$$+ (r/2+r^2/2+krb/2)u(x-h, t)+(k+k^2b/2)F(x, t)+\frac{k^2}{2}F_t(x, t)-\frac{k^2a}{2}F_x(x, t), \tag{11.7.13}$$

with $r = ka/h$. The *truncation error* can be shown to be of $O(k^2 + h^2)$.

With $(x_i, t_j) = (ih, jk)$ we can write the difference equation (11.7.13) as

$$u_{i,j+1} = \left(1 - r^2 + kb + \frac{1}{2}k^2b^2\right) u_{i,j} + \frac{-r + r^2 + krb}{2} u_{i+1,j}$$

$$+ \frac{r + r^2 + krb}{2} u_{i-1,j} + \left(k + \frac{k^2b}{2}\right) F_{i,j} + \frac{k^2}{2}[(F_t)_{i,j} - a(F_x)_{i,j}], \tag{11.7.14}$$

where $r = ka/h$. This is the *explicit Lax-Wendroff difference scheme* for the numerical solution of (11.7.1). We again find that the *CFL stability condition* is $|r| \leq 1$. Next, we set $F = 0$ and $b = 0$ in the scheme and attempt to characterize the resulting difference equation as a random walk. This requires that each coefficient $1 - r^2$, $(r^2 - r)/2$, and $(r^2 + r)/2$ must be bounded below by 0 and above by 1, as it must represent a probability. From $1 - r^2 \geq 0$ we conclude that $|r| \leq 1$, which yields $r^2 \leq |r|$. Then, if $a > 0$, so that $r > 0$, we see that the coefficient $(r^2 - r)/2$ is negative. If $a < 0$, so that $r < 0$, the coefficient $(r^2 + r)/2$ is negative. As a result, the Lax-Wendroff scheme has no random walk representation if $F = 0$ and $b = 0$ for any nonzero r. Nevertheless, the von Neumann stability condition is found to be $|r| \leq 1$, in agreement with the *CFL* result, so that the scheme is *conditionally stable.*

We conclude from the above that the *CFL stability condition* is a *necessary condition* for the stability of an explicit difference scheme. If the scheme has a random walk representation when $F = 0$ and $b = 0$, subject to a restriction on the parameter r, this (apparently) determines a *sufficient condition* for the stability of the scheme. We expect the scheme to be stable in this case, because the coefficients yield weighted averages of the values of $u(x_i, t_j + 1) = u_{i,j+1}$ at the earlier time $t = t_j$, and this prevents the unbounded growth of the initial values. (If $b \neq 0$, some growth can be allowed.) Yet the possibility of constructing a related random walk representation for the scheme is not a *necessary condition* for the stability of a difference scheme.

There are *implicit difference schemes* that correspond to the explicit forward-backward and forward-forward difference schemes introduced above. The *implicit backward-backward difference scheme* is obtained by approximating both $u_t(x,t)$ and $u_x(x,t)$ by a backward difference quotient in (11.7.1). This yields $u_{i,j} = (1+r-kb)u_{i,j+1} - ru_{i-1,j+1} - kF_{i,j+1}$. The *truncation error* can be shown to be of $O(k+h)$. The *implicit backward-forward difference scheme* is obtained by approximating $u_t(x,t)$ by a backward difference quotient and $u_x(x,t)$ by a forward difference quotient in (11.7.1). This yields $u_{i,j} = (1-r-kb)u_{i,j+1} + ru_{i+1,j+1} - kF_{i,j+1}$. The *truncation error* can be shown to be of $O(k+h)$.

In both implicit methods, the solution values $u_{i,j+1}$ and $u_{i-1,j+1}$ or $u_{i+1,j+1}$ must be determined in terms of the known solution value $u_{i,j}$. In the case of an initial value problem, this requires that an infinite number of simultaneous equations be solved to move from one time step to the next, even if the solution is required at a finite number of grid points. However, if an initial and boundary value problem is to be considered, the foregoing implicit methods can be used without having to solve a system of simultaneous equations. For example, if $a > 0$ and $u(x,0) = f(x)$ and $u(0,t) = h(t)$ are prescribed (with the solution to be found for $x > 0$, $t > 0$), the *implicit backward-backward difference scheme* can be used to determine $u_{1,1}$ in terms of the known values $u_{0,1}$ and $u_{1,0}$. Next, $u_{2,1}$ can be determined from known values, and so on. This process can be continued to determine $u_{i,j}$ for all $i \geq 1$ and $j \geq 1$. A similar solution process can be carried out using the *implicit backward-forward difference scheme* if $a < 0$ and the solution is to be found for $x < 0$, $t > 0$. Both schemes are *unconditionally stable*.

In general, the numerical solution of initial and boundary value problems for first order PDEs using the Lax-Friedrichs, Lax-Wendroff, or implicit difference schemes requires the introduction of an additional boundary condition, either to reduce the number of calculations to a manageable size or to carry out the scheme at all. How this can be done is not discussed here. It is touched upon in Section 11.10, where built-in Maple finite difference methods are presented.

Each of the foregoing schemes is a *one-step scheme* in the time variable. If we approximate $u_t(x,t)$ and $u_x(x,t)$ by a centered difference quotient in (11.7.1), we obtain the *two-step difference scheme*

$$u_{i,j+1} = u_{i,j-1} + r(u_{i-1,j} - u_{i+1,j}) + 2kbu_{i,j} + 2kF_{i,j}, \qquad (11.7.15)$$

where $r = ka/h$. This *explicit centered-centered difference scheme* is also known as the *leapfrog scheme*. The *truncation error* can be shown to be of $O(k^2+h^2)$. A von Neumann stability analysis shows that it is stable if $|r| \leq 1$. If we apply this scheme to the initial value problem for the PDE, we need a method to determine the values $u_{i,1}$ since only $u_{i,0}$ is given in terms of the initial data. Any of the explicit one-step methods given above can be used to determine these values. Even the explicit forward-centered scheme can be used. Stability is not an issue, as we apply the method for only one step.

$NumHypSysExplicit([[a]],[[b]],[F(x,t)],t=0..\tau,[f(x)],x=0..l,\tau/k,l/h,Method)$ constructs numerical solutions of (11.7.1) with the initial condition

$u(x,0) = f(x)$ over the interval $[0,l]$ at the time $t = \tau$. The x and t increments are h and k, respectively. The last argument determines which of the above four one-step explicit methods is to be used in the procedure. It must be given as *Forward-Backward, Forward-Forward, Lax-Friedrichs,* or *Lax-Wendroff*. The output of the procedure exhibits the initial value problem to be solved and the values of h and k. Also, the value of $r = ka/h$ is displayed and we must have $|r| \leq 1$ for stability. The values of $u_{i,j}$ are given as an array at the x grid points evaluated at the time $t_j = jk = \tau$. The global variable $ScalSol$ displays the x grid points together with the approximate solution values as a list. $plot(ScalSol)$ plots the approximate solution.

Example 11.25. An Initial Value Problem. The procedure $NumHypSys\ Explicit([[4]],[[-1]],[8xt-8x+4+x^2t-4t], t = 0..1, [-x^2+\exp(-x/4)+8], x = 0..1, 5, 30, Method)$ uses any of the four one-step methods introduced above. The first order PDE to be solved is

$$u_t(x,t)+4\,u_x(x,t) = -u(x,t)+8xt-8x+4+x^2t-4t, \quad 0 < x < 1, \tag{11.7.16}$$

and the initial condition is

$$u(x,0) = -x^2 + \exp(-x/4) + 8. \tag{11.7.17}$$

The exact solution of the problem (11.7.16)–(11.7.17) is

$$u(x,t) = (t-1)x^2 + \exp(-x/4) + 8 - 4t. \tag{11.7.18}$$

The numerical solution is to be found at $t = 1$ with $h = 1/5$ and $k = 1/30$. This gives $r = ka/h = 4(5)/30 = 2/3$.

The following array displays the exact and numerical results at the time $t = 1$ at the listed values of x:

$$\begin{bmatrix} x & Exact & F-B & F-F & L-F & L-W \\ 0 & 5 & 4.806 & -7.887 & 4.708 & 5.002 \\ 0.2 & 4.951 & 4.756 & 15.06 & 4.658 & 4.953 \\ 0.4 & 4.905 & 4.708 & -2.176 & 4.609 & 4.907 \\ 0.6 & 4.861 & 4.663 & 11.84 & 4.563 & 4.862 \\ 0.8 & 4.819 & 4.620 & -3.608 & 4.519 & 4.820 \\ 1 & 4.779 & 4.579 & 16.90 & 4.478 & 4,780 \end{bmatrix}. \tag{11.7.19}$$

Each scheme is stable except for the *forward-forward scheme.* The instability of that scheme accounts for the wildly oscillatory values in the output of the procedure. As for the three other schemes, the *Lax-Wendroff scheme* yields the best agreement with the exact results. The results from the *forward-backward scheme* are better than those from the *Lax-Friedrichs scheme.*

Example 11.26. An Initial and Boundary Value Problem. We reconsider the problem of Example 11.25 and add a boundary condition $u(0,t) = 9 - 4t$ at $x = 0$, with the solution to be found for $x > 0$. The exact solution of the problem is (11.7.18), as given in Example 11.25. The procedure $NumHypSysExplicit$ can solve initial and boundary value problems using the *forward-forward* and *forward-backward* schemes. Here we must use the *forward-backward scheme*. All the arguments used in Example 11.25 are retained and we add the additional tenth argument $9 - 4t$, which represents the boundary condition at $x = 0$. We do not display the array that is the output of the procedure, but exhibit the output of the global variable $ScalSol$. It is $[[0., 5.0], [.20, 4.951], [.40, 4.902], [.60, 4.854], [.80, 4.806], [1.0, 4.759]]$, and it can be plotted using Maple's *plot* procedure.

First Order Hyperbolic Systems

Each of the four conditionally stable explicit one-step difference schemes introduced above for the scalar PDE (11.7.1) carries over to the hyperbolic system (11.7.2) with constant coefficient matrices. We recall that the system (11.7.2) is hyperbolic if all the eigenvalues of the $n \times n$ matrix A are real and there are n linearly independent eigenvectors. Although we are now working with vectors and matrices, we retain the notation introduced above in the scalar case. We represent the $n \times n$ identity matrix by I in our discussion.

The analysis of first order hyperbolic systems is closely connected with the study of first order scalar PDEs as was shown in Section 3.3. Each hyperbolic system of the form (3.3.42) [where y, C, and $\mathbf{d}$ are replaced by t, B, and $\mathbf{F}$, respectively, for our system (11.7.2)] can be represented in the *normal form* (3.3.45) or the *characteristic form* (3.3.46). We observe that if $B = 0$, the normal form (3.3.45) represents an uncoupled system of n first order scalar PDEs for the components of the vector $\mathbf{v}$, since A is a constant matrix. [The vector $\mathbf{v}$ in (3.3.45) is defined as $\mathbf{v} = R^{-1}\mathbf{u}$, where R is a matrix whose column vectors are the eigenvectors of A.] Each component $v_i(x,t)$, $i = 1, 2, \ldots, n$, of $\mathbf{v}(x,t)$ satisfies a first order PDE of the form (11.7.1), with $a = \lambda_i$, the ith eigenvalue of A, $b = 0$, and $F(x,t)$ defined appropriately. As a result, we can apply the foregoing difference schemes for each of the scalar equations satisfied by the components $v_i(x,t)$. [Even if $B \neq 0$ and the scalar equations for the $v_i(x,t)$ are coupled, this can still be done, since the principal parts of the equations are uncoupled.]

Each scalar equation can be dealt with separately, but we want to deal with all n equations simultaneously and apply the same difference scheme for each equation. Thus, the use of the *explicit forward-backward difference scheme* for each scalar equation would require that $\lambda_i > 0$ and that $r_i = k\lambda_i/h \leq 1$ for each eigenvalue if the scheme is to be stable for each equation. Similarly, the use of the *explicit forward-forward difference scheme* for each scalar equation would require that $\lambda_i < 0$ and that $r_i = k\lambda_i/h \geq -1$ for each eigenvalue if the scheme is to be stable for each equation. The use of the *Lax-Friedrichs* or *Lax-Wendroff difference schemes* for each scalar equation would require that $r_i = k|\lambda_i|/h \leq 1$ for each eigenvalue if the scheme is

to be stable for each equation. In keeping with these observations, we apply each of the four explicit difference schemes to the (unmodified) system (11.7.2) directly. In each case, except for the Lax-Wendroff scheme, the difference equations have the same general form as in the scalar case.

The *explicit forward-backward difference scheme* for the numerical solution of (11.7.2) is given in increment form as

$$\mathbf{u}(x,t+k) = \left(I - \frac{kA}{h} + kB\right)\mathbf{u}(x,t) + \frac{kA}{h}\mathbf{u}(x-h,t) + k\mathbf{F}(x,t), \quad (11.7.20)$$

and in subscript form as

$$\mathbf{u}_{i,j+1} = \left(I - \frac{kA}{h} + kB\right)\mathbf{u}_{i,j} + \frac{kA}{h}\mathbf{u}_{i-1,j} + k\mathbf{F}_{i,j}. \quad (11.7.21)$$

The *truncation error* is again of $O(k+h)$. The stability condition is $\lambda_i > 0$ and $r_i = k\lambda_i/h \leq 1$ for each eigenvalue of A.

The *explicit forward-forward difference scheme* for the numerical solution of (11.7.2) is given in increment form as

$$\mathbf{u}(x,t+k) = \left(I + \frac{kA}{h} + kB\right)\mathbf{u}(x,t) - \frac{kA}{h}\mathbf{u}(x+h,t) + k\mathbf{F}(x,t), \quad (11.7.22)$$

and in subscript form as

$$\mathbf{u}_{i,j+1} = \left(I + \frac{kA}{h} + kB\right)\mathbf{u}_{i,j} - \frac{kA}{h}\mathbf{u}_{i+1,j} + k\mathbf{F}_{i,j}. \quad (11.7.23)$$

The *truncation error* is again of $O(k+h)$. The stability condition is $\lambda_i < 0$ and $r_i = k\lambda_i/h \geq -1$ for each eigenvalue of A.

The *Lax-Friedrichs difference scheme* for the numerical solution of (11.7.2) is given in subscript form as

$$\mathbf{u}_{i,j+1} = kB\mathbf{u}_{i,j} + \left(\frac{I}{2} - \frac{kA}{2h}\right)\mathbf{u}_{i+1,j} + \left(\frac{I}{2} + \frac{kA}{2h}\right)\mathbf{u}_{i-1,j} + k\mathbf{F}_{i,j}. \quad (11.7.24)$$

The *truncation error* is of $O(k+h^2)$. The stability condition is $|r_i| = k|\lambda_i|/h \leq 1$ for each eigenvalue of A.

Finally, the *Lax-Wendroff difference scheme* for the numerical solution of (11.7.2) is given in subscript form as

$$\begin{aligned}\mathbf{u}_{i,j+1} = {} & \left(I - \frac{k^2A^2}{h^2} + kB + \frac{k^2B^2}{2}\right)\mathbf{u}_{i,j} \\ & + \left(-\frac{kA}{2h} + \frac{k^2A^2}{2h^2} - \frac{k^2(AB+BA)}{4h}\right)\mathbf{u}_{i+1,j} + \left(kI + \frac{k^2B}{2}\right)\mathbf{F}_{i,j} \quad (11.7.25) \\ & + \left(\frac{kA}{2h} + \frac{k^2A^2}{2h^2} + \frac{k^2(AB+BA)}{4h}\right)\mathbf{u}_{i-1,j} + \frac{k^2}{2}\left(\frac{\partial \mathbf{F}}{\partial t}\right)_{i,j} - \frac{k^2A}{2}\left(\frac{\partial \mathbf{F}}{\partial x}\right)_{i,j}.\end{aligned}$$

The *truncation error* is of $O(k^2 + h^2)$. The stability condition is $|r_i| = k|\lambda_i|/h \leq 1$ for each eigenvalue of A.

We see that the *forward-backward* and *forward-forward difference schemes* can only be used if all the eigenvalues of A have the same sign. The *Lax-Friedrichs* and *Lax-Wendroff difference schemes* permit A to have positive, negative or even zero eigenvalues. It is a straightforward matter to extend the *leap-frog difference scheme* to systems. Similarly, *implicit forward-backward* and *forward-forward difference schemes* can be constructed. These schemes can be used for initial and boundary value problems. The number of boundary conditions imposed at a boundary line depends on the signs of the eigenvalues of the matrix A. Since the eigenvalues are all of one sign, by assumption, if the forward-backward and forward-forward difference schemes are to be stable, it is an easy matter to assign boundary conditions for these schemes, just as for the scalar case. We do not consider these matters here.

The Maple procedure $NumHypSysExplicit$ can also be used to construct numerical solutions of the hyperbolic systems of first order PDEs (11.7.2) over the interval $[0, l]$ at the time $t = \tau$, with the initial condition $\mathbf{u}(x, 0) = \mathbf{f}(x)$. The x increments equal h and the t increments are given as k. The first two arguments of the procedure are a list of lists that prescribe the coefficient matrices A and B for the system. The third and fifth arguments are lists that specify the inhomogeneous term $\mathbf{F}(x, t)$ and the initial value $\mathbf{f}(x)$, respectively, for the problem. As was the case before, the last argument of the procedure determines which of the four one-step explicit methods introduced above is to be used in the procedure. It must be given as *Forward-Backward, Forward-Forward, Lax-Friedrichs,* or *Lax-Wendroff.*

The output of the procedure exhibits the initial value problem to be solved and the values of the increments h and k. Additionally, the eigenvalues of A and the stability condition for the discrete problem are displayed. The components of the numerical solution $\mathbf{u}_{i,j}$ are given as an array that exhibits their values at the x grid points for the time $t_j = jk = \tau$.

The foregoing procedure is applied to a specific initial value problem in the Example 11.27.

Example 11.27. An Initial Value Problem for a Hyperbolic System. The procedure $NumHypSysExplicit([[3, -1], [-1, 2]], [[-2, 4], [1, -3]], [6x + 2x^2 - 4t^2, 2t - 2x - x^2 + 3t^2], t = 0..1, [x^2, 0], x = 0..1, 8, 30, Method)$ constructs a numerical solution of the following initial value problem for the two-component vector $\mathbf{u}(x, t)$, where the $Method$ must be prescribed:

$$\mathbf{u}_t(x,t) + \begin{bmatrix} 3 & -1 \\ -1 & 2 \end{bmatrix} \mathbf{u}_x(x,t) = \begin{bmatrix} -2 & 4 \\ 1 & -3 \end{bmatrix} \mathbf{u}(x,t) + \begin{bmatrix} 6x + 2x^2 - 4t^2 \\ 2t - 2x - x^2 + 3t^2 \end{bmatrix},$$

$$\mathbf{u}(x,t) = \begin{bmatrix} u_1(x,t) \\ u_2(x,t) \end{bmatrix}, \qquad \mathbf{u}(x,0) = \begin{bmatrix} x^2 \\ 0 \end{bmatrix}. \tag{11.7.26}$$

As indicated by the arguments of the procedure, the x interval $[0, 1]$ is divided into eight equal parts, so that $h = 0.125$. The value of the time step is $k = 1/30$ and the numerical solution is to be found at $t = 1$. The exact solution of the initial value

problem is $\mathbf{u}(x,t) = [x^2, t^2]^T$. We enter $Forward-Backward$ for the method to be used in the procedure. Then the eigenvalues of A are found to be $\lambda_1 = 1.381966011$ and $\lambda_2 = 3.618033989$ and the stability parameter has the value 0.9648091, so that the method is stable.

As can be seen from the output of the procedure given below, and comparing the numerical and exact solution values, the agreement between the numerical and exact results is not that good:

$$\begin{bmatrix} x & u_1(x,1) & u_2(x,1) \\ 0.0 & 0.1285 & 0.9841 \\ 0.12 & 0.1442 & 0.9841 \\ 0.25 & 0.1911 & 0.9841 \\ 0.38 & 0.2693 & 0.9841 \\ 0.50 & 0.3788 & 0.9841 \\ 0.62 & 0.5195 & 0.9841 \\ 0.75 & 0.6914 & 0.9841 \\ 0.88 & 0.8946 & 0.9841 \\ 1.0 & 1.129 & 0.9841 \end{bmatrix}. \tag{11.7.27}$$

For example, at $x = 0$ the exact solution value of $u_1(x,t)$ equals 0, whereas the procedure yields 0.1285. The exact value of $u_2(x,1)$ equals 1, whereas the procedure gives 0.9841. A comparison of all solution values shows the maximum error to be of $O(h)$, with $h = 0.125$.

Since the eigenvalues are both positive, the *forward-forward scheme* is unstable and the output of the procedure has no relation to the exact solution values. Finally, we use the *Lax-Wendroff scheme,* and this yields

$$\begin{bmatrix} x & u_1(x,1) & u_2(x,1) \\ 0.0 & -0.0005062 & 1.0 \\ 0.12 & 0.01519 & 1.0 \\ 0.25 & 0.06214 & 1.0 \\ 0.38 & 0.1403 & 1.0 \\ 0.50 & 0.2498 & 1.0 \\ 0.62 & 0.3905 & 1.0 \\ 0.75 & 0.5624 & 1.0 \\ 0.88 & 0.7656 & 1.0 \\ 1.0 & 1.0 & 1.0 \end{bmatrix}. \tag{11.7.28}$$

The numerical results are now extremely accurate.

von Neumann Stability for First Order PDES and Hyperbolic Systems of PDEs

The *von Neumann stability condition* for scalar difference schemes was introduced in Section 11.5. If $u_{n,m}$ is a solution of a linear, homogeneous difference equation, we put $u_{n,m} = \alpha^m \exp(i\beta n)$. The von Neumann stability condition requires that for all β in the interval $[0, 2\pi]$, the *amplification factor* α must satisfy $|\alpha| \leq 1$ or $|\alpha| \leq 1 + O(k)$, where k is the time increment.

For vector difference schemes for $\mathbf{u}_{n,m}$, we put $\mathbf{u}_{n,m} = \boldsymbol{\alpha}^m \exp(i\beta n)$, where $\boldsymbol{\alpha}^m = [\alpha_1^m, \alpha_1^m, \ldots, \alpha_N^m]^T$ (the vector $\mathbf{u}_{n,m}$ has N components) into the difference equation. This yields an equation $\boldsymbol{\alpha}^{m+1} = M\,\boldsymbol{\alpha}^m$, where the *amplification matrix* M depends on β, h, and k. Stability requires that M^m does not grow without bound as $m \to \infty$, as will be the case if the *spectral radius* $\rho(M)$ satisfies $\rho(M) \leq 1$. This is the *von Neumann stability condition*. It requires that $\max(|\lambda_i| \leq 1)$ for all the eigenvalues λ_i.

Example 11.28. The Forward-Backward and Backward-Backward Difference Schemes: The Scalar Case. The *explicit forward-backward difference scheme* with $F = 0$ is

$$u_{n,m+1} = (1 - r + kb)u_{n,m} + ru_{n-1,m}, \tag{11.7.29}$$

where $r = ka/h$. On inserting $u_{n,m} = \alpha^m \exp(i\beta n)$ into (11.7.29), we obtain

$$\alpha = (1 - r(1 - \cos(\beta)) + kb) - (r\sin(\beta))i, \tag{11.7.30}$$

on separating the exponential into real and imaginary parts. For stability the modulus of α must satisfy $|\alpha| \leq 1$. This yields, for the squared modulus, $|\alpha|^2 = (1 - r(1 - \cos(\beta)) + kb)^2 + (r\sin(\beta))^2 = 1 - 2r(\cos(\beta) - 1)(r - 1) + O(k) \leq 1$. The inequality must be satisfied for all β. Now, if $\beta = \pi/2$, the inequality becomes $1 + 2r(r - 1) + O(k) \leq 1$. Then, if $r < 0$ or $r > 1$, the foregoing inequality is not valid. If $0 \leq r \leq 1$, the inequality is valid [if we rewrite it as $|\alpha|^2 \leq 1 + O(k)$] since $\cos(\beta) - 1 \leq 0$ for all β. [We note that if $b = 0$, the $O(k)$ term is absent.] Thus, the *von Neumann stability condition* is that $a \geq 0$ and $r = ka/h \leq 1$.

The *implicit backward-backward difference scheme* with $F = 0$ is

$$u_{n,m} = (1 + r - kb)u_{n,m+1} - ru_{n-1,m+1}, \tag{11.7.31}$$

where $r = ka/h$. On inserting $u_{n,m} = \alpha^m \exp(i\beta n)$ into (11.7.31), we obtain, after some simplification, $1/|\alpha|^2 = 1 + 2r(1 - \cos(\beta))(r + 1) + O(k)$. The condition $1 + 2r(1 - \cos(\beta))(r + 1) \geq 1$ results in $|\alpha| \leq 1 + O(k)$. If $a \geq 0$, the foregoing condition is satisfied for all $r \geq 0$, so that the implicit scheme is *unconditionally stable*.

Example 11.29. The Explicit Forward-Centered Difference Scheme: The Scalar Case. The *explicit forward-centered difference scheme* with $F = 0$ is given as

$$u_{n,m+1} = (1 + kb)u_{n,m} - (r/2)(u_{n+1,m} - u_{n-1,m}), \tag{11.7.32}$$

where $r = ka/h$. On inserting $u_{n,m} = \alpha^m \exp(i\beta n)$ into (11.7.32), we obtain $\alpha = (1+kb)+(r\sin(\beta))i$, so that we must have $|\alpha|^2 = 1+(r\sin(\beta))^2+O(k) \leq 1$, for stability. If we put $\beta = \pi/2$, the stability condition cannot be satisfied for any r, so the scheme is *unconditionally unstable*.

Example 11.30. The Explicit Forward-Backward Difference Scheme: The Vector Case. The *explicit forward-backward difference scheme* with $\mathbf{F} = \mathbf{0}$ and $B = 0$ is given as

$$\mathbf{u}_{n,m+1} = \left(I - \frac{kA}{h}\right)\mathbf{u}_{n,m} + \left(\frac{kA}{h}\right)\mathbf{u}_{n-1,m}. \tag{11.7.33}$$

(We assume that $B = 0$.) We insert $\mathbf{u}_{n,m} = \boldsymbol{\alpha}^m \exp(i\beta n)$ into (11.7.33) and obtain $\boldsymbol{\alpha}^{m+1} = (I - (1 - \cos(\beta))\,(kA/h) - i\sin(\beta)\,(kA/h))\,\boldsymbol{\alpha}^m$, with I as the identity matrix. The coefficient of $\boldsymbol{\alpha}^m$ is the *amplification matrix* M for the difference scheme. We must determine a condition that guarantees that all the eigenvalues of M are bounded by 1 in absolute value. Now the amplification matrix M is a linear function $M = L(A)$ of the matrix A. It is easy to show that if λ_j is an eigenvalue of A, then $L(\lambda_j)$ is an eigenvalue of $L(A)$. Thus, $L(\lambda_j) = 1 - (1 - \cos(\beta))r_j - i\sin(\beta)r_j$, where $r_j = k\lambda_j/h$. If we put $r = r_j$ and $b = 0$ in (11.7.30), the amplification factor α has the same form as $L(\lambda_j)$. Therefore, it follows that the *von Neumann stability condition* is that all the eigenvalues $\lambda_j \geq 0$ and that $k\lambda_j/h \leq 1$.

Example 11.31. The Lax-Wendroff Difference Scheme: The Scalar Case. The *explicit Lax-Wendroff difference scheme* with $F = 0$ and $b = 0$ is given as

$$u_{n,m+1} = \left(1 - r^2\right)u_{n,m} + \frac{1}{2}\left(-r + r^2\right)u_{n+1,m} + \frac{1}{2}\left(r + r^2\right)u_{n-1,m}, \tag{11.7.34}$$

where $r = ka/h$. We insert $u_{n,m} = \alpha^m \exp(i\beta n)$ into (11.7.34) and obtain $\alpha = 1-(1-\cos(\beta))r^2 - ir\sin(\beta)$. The condition given above on the amplification factor yields $|\alpha|^2 = 1 + r^2(r^2-1)(\cos(\beta)-1)^2 \leq 1$, Clearly, we must have $|r| \leq 1$ if this is to be satisfied. Consequently, the Lax-Wendroff difference scheme is *conditionally stable* with $|r| \leq 1$ as the von Neumann stability condition.

Exercises 11.7

11.7.1. Verify the truncation errors (11.7.4) and (11.7.9).

11.7.2. Find the truncation errors for the Lax-Friedrichs and Lax-Wendroff difference schemes.

11.7.3. Determine the truncation error for the leapfrog scheme (11.7.15).

11.7.4. Use the procedure *NumHypSysExplicit* to obtain the results given in the array (11.7.19) in Example 11.25.

11.7.5. Use the forward-backward difference scheme in *NumHypSysExplicit* to solve the initial and boundary value problem given in Example 11.26, with the arguments given in that example. Plot the exact and numerical solutions.

11.7.6. Solve the initial value problem for the hyperbolic system given in Example 11.27 using the procedure *NumHypSysExplicit* and invoking the forward-backward difference scheme.

11.7.7. Solve the initial value problem for the hyperbolic system given in Example 11.27 using the procedure *NumHypSysExplicit* and invoking the forward-forward difference scheme. Note the exact solution as given in the example and observe that the output of the procedure indicates instability.

11.7.8. Solve the initial value problem for the hyperbolic system given in Example 11.27 using the procedure *NumHypSysExplicit* and invoking the Lax-Friedrichs difference scheme.

11.7.9. Solve the initial value problem for the hyperbolic system given in Example 11.27 using the procedure *NumHypSysExplicit* and invoking the Lax-Wendroff difference scheme. Obtain the array (11.7.28).

11.7.10. Verify the von Neumann stability results obtained in Example 11.28.

11.7.11. Verify the von Neumann stability results obtained in Example 11.29.

11.7.12. Verify the von Neumann stability results obtained in Example 11.30.

11.7.13. Verify the von Neumann stability results obtained in Example 11.31.

11.8 FINITE DIFFERENCE METHODS FOR PDEs WITH VARIABLE COEFFICIENTS

We have presented a number of finite difference methods for the solution of (linear) heat and wave equations and Laplace and Poisson equations. Now we extend some of these methods to deal with linear and semilinear parabolic, hyperbolic, and elliptic equations with variable coefficients.

The *parabolic PDEs* that we treat have the general form

$$u_t(x,t) - (\alpha(x,t)u_x(x,t))_x + \beta(x,t)u_x(x,t) + \gamma(x,t)u(x,t) = F(x,t,u(x,t)). \tag{11.8.1}$$

Apart from the addition of the terms $\beta(x,t)u_x(x,t)$, (11.8.1) has the general form of the parabolic PDEs considered in Chapter 4. The PDE is semilinear because of the term $F(x,t,u(x,t))$, but this is not the most general semilinear form that the PDE can take. The *hyperbolic PDEs* have the general form

$$u_{tt}(x,t) - (\alpha(x,t)u_x(x,t))_x + \beta(x,t)u_x(x,t) + \gamma(x,t)u_t(x,t) + \delta(x,t)u(x,t) = F(x,t,u(x,t)). \tag{11.8.2}$$

Again, apart from the addition of the terms $\beta(x,t)u_x(x,t)$ and $\gamma(x,t)u_t(x,t)$, (11.8.2) has the general form of the hyperbolic PDEs considered in Chapter 4. [We note that the coefficients in (11.8.1)–(11.8.2) are not required to be independent of t as was the case in Chapter 4.] The semilinear *elliptic PDEs* are of the form

$$\alpha(x,y)u_{xx}(x,y) + \beta(x,y)u_{yy}(x,y) + \gamma(x,y)u_x(x,y) + \delta(x,y)u_y(x,y) + \omega(x,y)u(x,y) = F(x,y,u(x,y)). \tag{11.8.3}$$

Initial and boundary value problems for (11.8.1) and (11.8.2) will be solved by the *method of lines*, while *iteration methods* will be used to solve the boundary value problems for (11.8.3). As we have seen, characteristics play a significant role in the solution of problems for hyperbolic equations and systems. Consequently, we present a number of methods for the solution of problems for quasilinear second order hyperbolic PDEs and hyperbolic systems of first order PDEs that use the *method of characteristics*.

Method of Lines for Linear and Semilinear Parabolic Equations

The presentation of numerical solution methods for initial and boundary value problems for the parabolic equation (11.8.1) will be restricted to the semi-discrete *method of lines*. Fully discrete difference methods analogous to those given above for the heat equation can also be constructed for (11.8.1) but are not presented here. We have created a Maple procedure *NumParabolicLinesSL* for the solution of the system of first order ODE that arises in the method of lines. This system can be solved exactly, if possible, or approximately, by the use of finite difference methods for ODEs.

To apply the method of lines to the parabolic equation (11.8.1) we replace derivatives in x by difference quotients and leave the derivative in t intact. Centered difference quotients are used throughout except for the terms that involve $\alpha(x,t)$, where we use the difference approximation

$$\alpha(x-h,t)\frac{u(x-h,t)-u(x,t)}{2h^2}+\alpha(x,t)\frac{u(x-h,t)-2u(x,t)+u(x+h,t)}{2h^2}$$

$$+\alpha(x+h,t)\frac{u(x+h,t)-u(x,t)}{2h^2}=\frac{\partial}{\partial x}\left(\alpha(x,t)\frac{\partial u(x,t)}{\partial x}\right)+O\left(h^2\right). \quad (11.8.4)$$

With h as the increment in x and $x_i = a + ih$ (with an arbitrary a), this yields the system of ODEs

$$\frac{du_i(t)}{dt}=\frac{\alpha_{i-1}(t)(u_{i-1}(t)-u_i(t))}{2h^2}+\frac{\alpha_i(t)(u_{i+1}(t)-2u_i(t)+u_{i-1}(t))}{2h^2}$$

$$+\frac{\alpha_{i+1}(t)(u_{i+1}(t)-u_i(t))}{2h^2}-\frac{\beta_i(t)(u_{i+1}(t)-u_{i-1}(t))}{2h}-\gamma_i(t)u_i(t)+F_i(t), \quad (11.8.5)$$

where $u_i(t)$ represents the approximation to the exact solution $u(x,t)$ of the given problem at the point x_i at the time t, and $\alpha_i(t) = \alpha(a + ih, t)$, $\beta_i(t) = \beta(a + ih, t)$, $\gamma_i(t) = \gamma(a + ih, t)$, $F_i(t) = F(a + ih, t)$. Given an initial and boundary value problem for the parabolic equation (11.8.1), the initial and boundary conditions for the related system of ODEs (11.8.5) are determined as in the method of lines for the heat equation, and is not be repeated here.

Initial and boundary value problems are solved by the method of lines using *NumParabolicLinesSL*$(\alpha(x,t), \beta(x,t), \gamma(x,t), F(x,t,u(x,t)), t = t_0, f(x), x = a..b, bcl, g(t), bcr, s(t), n)$. The first four arguments in the procedure determine the coefficients and the inhomogeneous term in the parabolic PDE (11.8.1). The remaining arguments are as described above for the procedure *NumHeatLines*. Boundary conditions of all three types can be treated. The output is given as a Maple procedure *proc(rkf45_x)... end proc*. The global variable SP can be used to exhibit the numerical solution values $u_i(t_f)$ at all the x grid points, by invoking $SP(t_f)$. If an optional thirteenth argument $t = t_f$ is added, the numerical solution $u_i(t_f)$ at each grid point is exhibited in tabular form. The global variable $PList$ can also be used to display a list of these solution values. If an optional thirteenth argument *type = numeric* is added, a number of classical numerical solution methods for ODEs such as the *forward Euler, Heun, lower order Runge-Kutta*, or *predictor-corrector methods* can be invoked. The step sizes used in these methods can be controlled, as can the form of the output. (This is done by adding more optional arguments.)

Example 11.32. A Semilinear Parabolic Equation. We consider two initial and boundary value problems for the semilinear parabolic equation

$$u_t(x,t) - (xu_x(x,t))_x = F(x,t,u(x,t)), \quad 0 < x < 1,\ t > 0, \quad (11.8.6)$$

where $F(x,t,u(x,t)) = u^2(x,t) - e^x(2 + t + xt + x + e^x)/(t+1)^2$. The function $u(x,t) = e^x/(t+1)$ is an exact solution of (11.8.6). The initial condition for both problems is $u(x,0) = e^x$. In one problem, we introduce Dirichlet boundary conditions at $x = 0$ and $x = 1$, and in the other problem Neumann boundary conditions

are used. The interval $0 < x < 1$ is subdivided into five equal parts, and the solution at the grid points is to be found at $t = 1$.

We invoke the procedure $NumParabolicLinesSL(x, 0, 0, F(x,t,u(x,t), t = 0, e^x, x = 0..1, bcl, g(t), bcr, h(t), 10, t = 1)$. For the Dirichlet problem we put $bcl = bcr = dirichlet$, $r(t) = 1/(t+1)$ and $s(t) = e/(t+1)$. For the Neumann problem we put $bcl = bcr = 0$, $r(t) = -1/(t+1)$ and $s(t) = e/(t+1)$. The exact and numerical solution values (as given by the procedures) are displayed in the following array:

$$\begin{bmatrix} x & Exact & Dirichlet & Neumann \\ 0 & .5000 & .5000 & .5000 \\ 0.2 & .6105 & .6156 & .5992 \\ 0.4 & .7460 & .7508 & .7284 \\ 0.6 & .9110 & .9148 & .8888 \\ 0.8 & 1.113 & 1.115 & 1.087 \\ 1 & 1.359 & 1.359 & 1.331 \end{bmatrix}. \tag{11.8.7}$$

Increasing the number of grid points will improve the numerical results.

In Section 4.7 we applied *nonlinear stability theory* to an initial and boundary value problem for the semilinear parabolic equation (4.7.1). In the following example, we use the *method of lines* to analyze this problem.

Example 11.33. Nonlinear Stability. We determine a numerical solution of the semilinear parabolic PDE

$$u_t(x,t) - u_{xx}(x,t) = \hat{\lambda} u(x,t)(1 - u^2(x,t)), \qquad 0 < x < \pi,\ t > 0, \tag{11.8.8}$$

with the initial and boundary conditions

$$u(x,0) = \epsilon \sin(x), \qquad u(0,t) = 0,\ u(\pi,t) = 0. \tag{11.8.9}$$

Here ϵ is a small parameter which we fix as $\epsilon = 0.01$. It was shown in Section 4.7, using a perturbation approach, that the solution of (11.8.8)–(11.8.9) decays to zero as $t \to \infty$ if $\hat{\lambda} < \hat{\lambda}_c = 1$, the critical value of the parameter. If $\hat{\lambda} \approx 1$ but $\hat{\lambda} > 1$, the solution does not grow unboundedly, as predicted by linear stability theory, but tends to a steady state.

We begin by putting $\hat{\lambda} = 0.9$ in (11.8.8). The interval is subdivided into 10 equal parts and the procedure $NumParabolicLinesSL(1, 0, 0, 0.9u(x,t)(1 - u(x,t)^2), t = 0, 0.01\sin(x), x = 0..\pi, dirichlet, 0, dirichlet, 0, 10)$ is invoked. [The x grid points are $0,\ \pi/10,\ 2\pi/10, \ldots,\ \pi$.] The output is a procedure that yields the numerical solution at the grid points at prescribed values of t. It shows that the solution tends to zero with increasing t. For example, at $t = 100$ we have $u_1 = .328 \times 10^{-6}, u_2 = .587 \times 10^{-6}, u_3 = .859 \times 10^{-6}, u_4 = .950 \times 10^{-6}, u_5 = .106 \times 10^{-6}$. The values

of $u_6, \ldots, u_9$ equal those of $u_1, \ldots, u_4$ in reverse order. (Here u_1 represents the numerical solution at $x = \pi/10$.)

Next we set $\hat{\lambda} = 1.1$ in (11.8.8) and apply $NumParabolicLinesSL$ with the arguments given above, except that in the fourth argument we replace 0.9 by 1.1. It is found that for $t \geq 80$, the solution stabilizes around the values $u_1 = .1136, u_2 = .2151, u_3 = .2942, u_4 = .3442, u_5 = .3613$, with the values of $u_6, \ldots, u_9$ equal those of $u_1, \ldots, u_4$ in reverse order.

The steady-state result (as t grows large) given in (4.7.24) has the form $u(x,t) \approx (2/\sqrt{3})\sqrt{(\hat{\lambda}-1)/\hat{\lambda}}\,\sin(x)$ if we put $h(x) = \sin(x)$ in (4.7.24). We find that for large t, $u(\pi/10, t) \approx .1076, u(2\pi/10, t) \approx .2046, u(3\pi/10, t) \approx .2816, u(4\pi/10, t) \approx .331, u(5\pi/10, t) \approx .3481$. The values of $u(x,t)$ at $x = 6\pi/10, \ldots, 9\pi/10$ equal the values at $x = \pi/10, \ldots, 4\pi/10$ in reverse order. There is fairly good agreement between the perturbation and numerical results. The perturbation result indicates that the choice of $h(x)$ in the initial value affects only the sign of the steady-state approximation. Although we have set $h(x) = \sin(x)$ in the procedure, it can be shown that the numerical steady state result is independent of the choice of $h(x)$, provided that $h(x)$ is not orthogonal to $\sin(x)$ over the interval $(0, \pi)$.

Method of Lines for Linear and Semilinear Hyperbolic Equations

Numerical solution methods for initial and boundary value problems for the hyperbolic equation (11.8.2) are restricted to the semidiscrete *method of lines*. The procedure $NumHyperbolicLinesSL$ solves the system of first order ODEs that arises in the method of lines. This system can be solved exactly, or approximately by the use of finite difference methods for ODEs, as seen above.

To apply the method of lines to (11.8.2), we replace derivatives in x by difference quotients and leave the t derivatives intact. Centered difference quotients are used throughout except for the terms that involve $\alpha(x,t)$, where we use the difference approximation (11.8.4). With h as the x-increment and $x_i = a + ih$ (with arbitrary a), this yields the system of ODEs

$$u_i''(t)+\gamma_i(t)u_i'(t) = \frac{\alpha_{i-1}(t)(u_{i-1}(t) - u_i(t))}{2h^2}+\frac{\alpha_i(t)(u_{i+1}(t) - 2u_i(t) + u_{i-1}(t))}{2h^2}$$

$$+\frac{\alpha_{i+1}(t)(u_{i+1}(t) - u_i(t))}{2h^2} - \frac{\beta_i(t)(u_{i+1}(t) - u_{i-1}(t))}{2h} - \delta_i(t)u_i(t) + F_i(t), \tag{11.8.10}$$

where $u_i(t)$ represents the approximation to the exact solution $u(x,t)$ of the given problem at the point x_i at the time t, and $\alpha_i(t) = \alpha(a + ih, t)$, $\beta_i(t) = \beta(a + ih, t)$, $\gamma_i(t) = \gamma(a + ih, t)$, $\delta_i(t) = \delta(a + ih, t)$, $F_i(t) = F(a + ih, t)$. Given an initial and boundary value problem for (11.8.2), the determination of initial and boundary conditions for (11.8.10) follows that given above for the method of lines for the wave equation and is not repeated here.

The Maple procedure that solves initial and boundary value problems by the method of lines is $NumHyperbolicLinesSL(\alpha(x,t),\beta(x,t),\gamma(x,t),\delta(x,t),F(x,t,u(x,t)),t=t_0,f(x),g(x),x=a..b,bcl,g(t),bcr,s(t),n)$. The first five arguments in the procedure determine the coefficients and the inhomogeneous term in the hyperbolic PDE (11.8.2). The remaining arguments are as in $NumWaveLines$. Boundary conditions of all three types can be treated. The output is given as a Maple procedure *proc(rkf45_x)... end proc*. Then the global variable SH can be used to exhibit the numerical solution values $u_i(t_f)$, as well as approximate first order derivatives in t, at all the x grid points, by invoking $SH(t_f)$. If an optional fifteenth argument $t=t_f$ is added, the numerical solution $u_i(t_f)$ at each grid point is exhibited in tabular form. The global variable $PList$ can also be used to display a list of these solution values. If an optional fifteenth argument *type = numeric* is added, a number of classical numerical solution methods for ODEs such as the *forward Euler, Heun, lower order Runge-Kutta*, or *predictor-corrector methods*, can be invoked. The step sizes used in these methods can be controlled, as can the form of the output. (This is done by adding more optional arguments.)

We consider only one example and show that the solution of an initial and boundary value problem for a damped wave equation reduces to that of a related diffusion equation for large time. The relationship between these solutions has been discussed a number of times in this book.

Example 11.34. A Damped Wave Equation and a Related Diffusion Equation. We consider an initial and boundary value problem for the *damped wave equation*

$$u_{tt}(x,t)-u_{xx}(x,t)+u_t(x,t)=0, \tag{11.8.11}$$

and for the related *diffusion equation*

$$u_t(x,t)-u_{xx}(x,t)=0, \tag{11.8.12}$$

given over the interval $0<x<1$. The Dirichlet boundary conditions for both problems are $u(0,t)=1$, $u(1,t)=0$. The initial conditions are $u(x,0)=x^2$, $u_t(x,0)=\sin(\pi x)$, with the initial derivative prescribed only for (11.8.11).

The unit interval is subdivided into five equal parts and the numerical solutions of both problems are obtained at $t=15$. We use $NumHyperbolicLinesSL(1,0,1,0,0,t=0,x^2,\sin(\pi x),x=0..1,dirichlet,1,dirichlet,0,5,t=15)$ for (11.8.11) and $NumParabolicLinesSL(1,0,0,0,t=0,x^2,x=0..1,dirichlet,1,dirichlet,0,5,t=15)$ for (11.8.12) to determine the numerical solutions at $t=15$. We obtain

$$\begin{bmatrix} x & Wave & Diffusion \\ 0 & 1. & 1. \\ 0.2 & .7998 & .8000 \\ 0.4 & .6000 & .6000 \\ 0.6 & .4004 & .4000 \\ 0.8 & .2004 & .2000 \\ 1 & 0. & 0. \end{bmatrix}. \tag{11.8.13}$$

The results agree nicely, as expected. As t increases, the difference between corresponding numerical solution values tends to zero.

Second Order Quasilinear Hyperbolic Equations: Method of Characteristics for Initial Value Problems

The quasilinear second order PDE

$$\alpha\, u_{xx}(x,t) + 2\beta\, u_{xt}(x,t) + \gamma\, u_{tt}(x,t) = \delta, \tag{11.8.14}$$

where α, β, γ and δ are all functions of $(x, t, u(x,t), u_x(x,t), u_t(x,t))$, is of hyperbolic type at the points (x,t) where $\beta^2 - \alpha\gamma > 0$. [Clearly, the classification of (11.8.14) depends not only on the point (x,t) but also on the solution $u(x,t)$.] In Section 3.1 it was shown that if (11.8.14) is linear, its characteristics can be used to bring it into a simple canonical form. This simplification is not possible for the quasilinear case. Instead, we express the PDE in a characteristic form.

Suppose that $\gamma \neq 0$ in (11.8.14) and assume, without loss of generality, that γ is positive. Let the characteristic curves of (11.8.14) be expressed as $x = x(t)$. Then it follows from (3.1.8), appropriately modified, that the two families of characteristics $x = x(t)$ satisfy the equations

$$\frac{dx(t)}{dt} = \frac{\beta \pm \sqrt{\beta^2 - \alpha\gamma}}{\gamma}, \tag{11.8.15}$$

where α, β, and γ are all functions of $(x(t), t, u(x(t),t), u_x(x(t),t), u_t(x(t),t))$. Now, the chain rule for differentiation for a function $z = z(x(t), t)$ gives $dz/dt = \partial z/\partial t + (dx(t)/dt)\partial z/\partial x$. We set $v(t) = u_x(x(t),t)$ and $w(t) = u_t(x(t),t)$. It is then easily verified that

$$\frac{\beta - \sqrt{\beta^2 - \alpha\gamma}}{\gamma}\frac{dv(t)}{dt} + \frac{dw(t)}{dt} = \frac{\delta}{\gamma} \quad \text{on} \quad \frac{dx(t)}{dt} = \frac{\beta + \sqrt{\beta^2 - \alpha\gamma}}{\gamma}, \tag{11.8.16}$$

$$\frac{\beta + \sqrt{\beta^2 - \alpha\gamma}}{\gamma}\frac{dv(t)}{dt} + \frac{dw(t)}{dt} = \frac{\delta}{\gamma} \quad \text{on} \quad \frac{dx(t)}{dt} = \frac{\beta - \sqrt{\beta^2 - \alpha\gamma}}{\gamma}. \tag{11.8.17}$$

[Since $\gamma > 0$ by assumption, we find that the value of $dx(t)/dt$ in (11.8.16) exceeds that in (11.8.17) at common points.] These equations correspond to the characteristic normal form (3.3.76) obtained in Section 3.3.

Suppose that two points $Q = (x_Q, t_Q)$ and $R = (x_R, t_R)$ are specified at which the values of u, u_x, and u_t are known. These values determine two characteristic directions at each point Q and R from (11.8.15). Four characteristic curves are thereby determined. Of these, we choose those characteristics that issue from Q and R and intersect at some later value of t (i.e., the forward characteristics). It follows from the above, that this will (generally) be the case if we choose the plus sign in (11.8.15) at Q and the minus sign in (11.8.15) at R. We denote the intersection point

by $P = (x_P, t_P)$ and seek to determine P and the values of u, u_x, and u_t at that point from (11.8.16)–(11.8.17). (u_Q, u_R, v_Q, v_R, w_Q, w_R are known.) No exact solution is generally available for the system, so we construct a numerical method for its approximate solution. Derivatives are replaced by difference quotients, and we obtain

$$\frac{x_P - x_Q}{t_P - t_Q} = \left.\frac{\beta + \sqrt{\beta^2 - \alpha\gamma}}{\gamma}\right|_Q, \quad \left.\frac{\beta - \sqrt{\beta^2 - \alpha\gamma}}{\gamma}\right|_Q \frac{v_P - v_Q}{t_P - t_Q} + \frac{w_P - w_Q}{t_P - t_Q} = \left.\frac{\delta}{\gamma}\right|_Q \tag{11.8.18}$$

$$\frac{x_P - x_R}{t_P - t_R} = \left.\frac{\beta - \sqrt{\beta^2 - \alpha\gamma}}{\gamma}\right|_R, \quad \left.\frac{\beta + \sqrt{\beta^2 - \alpha\gamma}}{\gamma}\right|_R \frac{v_P - v_R}{t_P - t_R} + \frac{w_P - w_R}{t_P - t_R} = \left.\frac{\delta}{\gamma}\right|_R. \tag{11.8.19}$$

(The $|_Q$ and $|_R$ indicate that all the terms are evaluated at the points Q or R.) The four equations (11.8.18)–(11.8.19) must be solved simultaneously for x_P, t_P, v_P, and w_P.

Once a solution is obtained, the approximate intersection point P of the characteristics is determined and approximate values of $v_P = u_x$ and $w_p = u_t$ are found. To obtain an approximate value for u at the point P (i.e., u_P), we can use either of the following approximations

$$u_P \approx u_Q + \frac{1}{2}(v_Q + v_P)(x_P - x_Q) + \frac{1}{2}(w_Q + w_P)(t_P - t_Q), \tag{11.8.20}$$

$$u_P \approx u_R + \frac{1}{2}(v_R + v_P)(x_P - x_R) + \frac{1}{2}(w_R + w_P)(t_P - t_R). \tag{11.8.21}$$

The leading terms in the Taylor series of u_P yield $v_Q(x_P - x_Q)$ and $w_Q(t_P - t_Q)$ in (11.8.20) and a corresponding expression in (11.8.21). The (improved) approximation to u_P given in (11.8.20)–(11.8.21) is obtained by averaging the values of v and w at the points Q, R, P, as shown. A similar averaging approach can be used to develop an iteration scheme that improves the accuracy of the approximate values of x_P, t_P, v_P, and w_P determined from (11.8.18)–(11.8.19).

NumHypScalChar automates the foregoing *numerical method of characteristics*. It is presented in two forms. If a noncharacteristic initial value problem is considered with the initial curve given as the line $t = t_Q$ and the initial values as $u(x, t_Q) = f(x)$ and $u_t(x, t_Q) = g(x)$, two values of x must be prescribed, $x = x_Q$ and $x = x_R$. On using $NumHypScalChar(\alpha, \beta, \gamma, \delta, f(x), g(x), [x, t, u, u_x, u_t], [x_Q, t_Q], [x_R, t_Q], n)$, the (hyperbolic) PDE (11.8.14) with the foregoing initial values, supplemented by $u_x(x, t_Q) = f'(x)$, is considered. These values are used to determine u_Q, u_R, v_Q, w_Q, v_R, w_R at the points Q and R specified in the procedure. The output of the procedure exhibits the coordinates of the point P where the characteristics that issue from R and Q intersect. The values of u_P, v_P, w_P are also displayed. The last argument, n, in the procedure determines the number of iterations carried out if the errors from one step to the next exceed a built-in error tolerance.

If we wish to determine the solution at a point P using solution values given at the arbitrary points R and Q, as discussed above, we invoke $NumHypScalChar(\alpha, \beta, \gamma, \delta, [u_Q, u_R], [[v_Q, w_Q], [v_R, w_R]], [x, t, u, u_x, u_t], [x_Q, t_Q], [x_R, t_R], n)$. The output has the form given by the foregoing procedure. For each form of the procedure, the output at two different points can be used to find the solution at a new point. Thereby a set of solution values can be found at a grid of points in the (x,t)-plane. In general, this grid is quite irregular, but if the principal part of the PDE has constant coefficients, so that the characteristics are straight lines, it is possible to determine solution values on a regular grid.

Example 11.35. The Numerical Method of Characteristics for a Quasilinear Hyperbolic PDE. The procedure *NumHypScalChar*$(-u_x, 0, u^2, -\exp(2x)\sin^2(1+t) - \exp(3x)\sin^3(1+t), \exp(x)\sin(1), \exp(x)\cos(1), [x, t, u, u_x, u_t], [1.5, 0], [1.6, 0], 8)$ finds the intersection point of the two characteristics for the hyperbolic PDE $-u_x(x,t)u_{xx}(x,t) + u^2(x,t)u_{tt}(x,t) = -\exp(2x)\sin^2(1+t) - \exp(3x)\sin^3(1+t)$, with the initial data $u(x,0) = e^x \sin(1)$, $u_t(x,0) = e^x \cos(1)$ that issue from the points $Q = (1.5, 0)$ and $R = (1.6, 0)$. Maple's output is

$$\begin{bmatrix} x & t & u(x,t) & u_x(x,t) & u_t(x,t) \\ 1.550607 & 0.101064 & 4.203614 & 4.202042 & 2.132819 \end{bmatrix}. \tag{11.8.22}$$

The point of intersection P is given as $(x_P, t_P) = (1.550607, 0.101064)$. The approximate values of u and its first derivatives at that point are also displayed.

If we use the points $Q = (1.6, 0)$ and $R = (1.7, 0)$ in the procedure, the output is

$$\begin{bmatrix} x & t & u(x,t) & u_x(x,t) & u_t(x,t) \\ 1.65061 & 0.106319 & 4.65799 & 4.65603 & 2.33245 \end{bmatrix}. \tag{11.8.23}$$

The exact solution of the initial value problem is $u(x,t) = e^x \sin(1+t)$. Its values at the intersection points as determined from the procedure are $u(1.550607, .101064) = 4.20371$ and $u(1.65061, .106319) = 4.65817$. They are in excellent agreement with the values found from the procedure.

The values obtained from the two procedures can be used to determine a new characteristic intersection point P and approximate values for the solution and its derivatives there. This is done by invoking $NumHypScalChar(-u_x, 0, u^2, -\exp(2x)\sin^2(1+t) - \exp(3x)\sin^3(1+t)$, [4.2036,4.65799],[[4.20204,2.133], [4.65603, 2.3325]], $[x, t, u, u_x, u_t]$, $[1.55061, .10106]$, $[1.65061, .10632]$, $8)$. The intersection points found by the two foregoing procedures now play the role of Q and R. The values of u, u_x and u_t are entered in the fifth and sixth arguments of the procedure. The output is

$$\begin{bmatrix} x & t & u(x,t) & u_x(x,t) & u_t(x,t) \\ 1.60247 & 0.210176 & 4.64543 & 4.64219 & 1.749125 \end{bmatrix}. \tag{11.8.24}$$

The value of the exact solution at the intersection points determined from the procedure, $u(1.60247, .210176) = 4.64591$, agrees nicely with the value found from the procedure.

A set of approximate solution values has been found at an irregular set of grid points. It is clearly possible to determine additional grid points and solution values at these points and thereby obtain an approximation to the solution for $t > 0$ within a range of x and t values.

Method of Characteristics for Hyperbolic Systems of Two Quasilinear Equations In Two Unknowns

We consider the quasilinear hyperbolic system of the form (3.3.73)

$$A\mathbf{u}_x(x,t) + B\mathbf{u}_t(x,t) = \mathbf{c}, \tag{11.8.25}$$

where

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \; B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}, \; \mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}, \; \mathbf{u}(x,t) = \begin{bmatrix} u_1(x,t) \\ u_2(x,t) \end{bmatrix}.$$

The elements of $A, B, \mathbf{c}$ can be functions of $(x, t, u_1(x,t), u_2(x,t))$. With λ_1 and λ_2 as the (real) eigenvalues of the eigenvalue problem $A^T\mathbf{r} = \lambda B^T\mathbf{r}$ [they are the roots of $\det(A - \lambda B) = 0$] we obtain the *characteristic normal form*

$$\sigma_{i1}\frac{du_1(t)}{dt} + \sigma_{i2}\frac{du_2(t)}{dt} = \delta_i \quad \text{on} \quad \frac{dx(t)}{dt} = \lambda_i, \qquad i = 1,2. \tag{11.8.26}$$

Here $u_i(t) = u_i(x(t), t)$ and $dx(t)/dt = \lambda_i$ determines the characteristic curves for $i = 1, 2$.

The system of equations (11.8.26) is in the form of the system (11.8.16)–(11.8.17). We proceed as before. Consider the two points $Q = (x_Q, t_Q)$ and $R = (x_R, t_R)$ at which the values of u_1 and u_2 are known. We choose the characteristics that issue from Q and R and intersect at some later value of t (i.e., the forward characteristics). We denote the intersection point by $P = (x_P, t_P)$ and seek to determine P and the values of u_1 and u_2 at that point, using (11.8.26). Again, we construct a numerical method for its approximate solution by replacing derivatives by difference quotients. This yields the four (linear) equations for x_P, t_P, v_P, w_P,

$$\frac{x_P - x_Q}{t_P - t_Q} = \lambda_1\big|_Q, \qquad \sigma_{11}\big|_Q\frac{v_P - v_Q}{t_P - t_Q} + \sigma_{12}\big|_Q\frac{w_P - w_Q}{t_P - t_Q} = \delta_1\big|_Q, \tag{11.8.27}$$

$$\frac{x_P - x_R}{t_P - t_R} = \lambda_2\big|_R, \qquad \sigma_{21}\big|_R\frac{v_P - v_R}{t_P - t_R} + \sigma_{22}\big|_R\frac{w_P - w_R}{t_P - t_R} = \delta_2\big|_R, \tag{11.8.28}$$

where v_Q, w_Q, v_R, w_R, and v_P, w_P represent u_1, u_2 evaluated at Q, R, and P, respectively.

NumHypSystChar automates the foregoing numerical method of characteristics. It is again presented in two forms. If a noncharacteristic initial value problem is considered with the initial curve given as the line $t = t_Q$ and the initial values as

$u_1(x,t_Q) = f(x)$ and $u_2(x,t_Q) = g(x)$, two values of x must be prescribed, $x = x_Q$ and $x = x_R$. On invoking the procedure $NumHypSystChar(A, B, \mathbf{c}, [f(x), g(x)], [u_1, u_2], [x, t], [x_Q, t_Q], [x_R, t_Q], n)$ the (hyperbolic) system (11.8.25) with the foregoing initial values is considered. These values are used to determine v_Q, w_Q, v_R, w_R at the points Q and R specified in the procedure. The output of the procedure exhibits the coordinates of the point P where the characteristics that issue from R and Q intersect. Also, the values of v_P, w_P are displayed. The last argument n in the procedure determines the number of iterations carried out in the procedure if the errors from one step to the next exceed a built-in error tolerance.

If we wish to determine the solution at a point P using solution values given at the arbitrary points R and Q, the procedure takes the form *NumHypSystChar*$(A, B, \mathbf{c}, [[v_Q, w_Q], [v_R, w_R]], [u_1, u_2], [x, t], [x_Q, t_Q], [x_R, t_R], n)$. The output has the form given by the foregoing procedure.

Example 11.36. The Numerical Method of Characteristics for a Quasi-linear Hyperbolic System. *NumHypSystChar*$([[u_1^3, 0], [5, u_2]], [[1, 0], [0, 1]], [[3u_1 + 2x^7 \exp(4t) - 2x^2 \exp(t)], [-u_2 + 10x \exp(t) - 3x^2t^2 + 3x^5 + 2t + t^2 - x^3]], [x^2, -x^3], [u_1, u_2], [x, t], [1, 0], [1.2, 0], 25)$ solves the initial value problem for the system (11.8.25) [i.e., $A\mathbf{u}_x + B\mathbf{u}_t = \mathbf{c}$] with

$$A = \begin{bmatrix} u_1^3 & 0 \\ 5 & u_2 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} 3u_1 + 2x^7e^{4t} - 2x^2e^t \\ -u_2 + 10xe^t - 3x^2t^2 + 3x^5 + 2t + t^2 - x^3 \end{bmatrix},$$

$B = I$ (the identity matrix), $[u_1(x,0), u_2(x,0)] = [x^2, -x^3]$, $Q = (1,0)$, and $R = (1.2, 0)$. The exact solution of this problem is $u_1(x,t) = x^2e^t$, $u_2(x,t) = t^2 - x^3$. Maple's output,

$$\begin{bmatrix} x & t & u_1(x,t) & u_2(x,t) \\ 1.1027 & 0.064050 & 1.3023 & -1.3087 \end{bmatrix}, \tag{11.8.29}$$

displays the point P where the forward characteristics intersect and the values of u_1 and u_2 at that point. The values of the exact solution at the point P found by the procedure are $u_1(1.1027, 0.06405) = 1.296376981$ and $u_2(1.1027, 0.06405) = -1.336722674$.

The value of t_P is greater than zero (the value of t_Q) for this problem. If it turns out for a specific choice of Q and R that $t_P < 0$, we put $Q = R$ and $R = Q$ and reapply the procedure. We do not consider an example where the values u_1 and u_2 are known only at Q and R.

Characteristic Difference Methods for Linear Hyperbolic Systems

In the two preceding subsections, the method of characteristics was used to advance the numerical solution values in time, using given solution values at an earlier time. The point(s) at which the solution was found were not known a priori and had to be determined. Here we obtain an (approximate) numerical solution at a prescribed

point $P = (x_P, t_P)$ in the (x,t)-plane, in terms of initial data for $\mathbf{u}(x,t)$ given at the earlier time $t = t_Q$. We consider the linear hyperbolic system

$$A(x,t)\mathbf{u}_x(x,t) + B(x,t)\mathbf{u}_t(x,t) = C(x,t)\mathbf{u}(x,t) + \mathbf{d}(x,t), \qquad (11.8.30)$$

where A, B, C are $n \times n$ matrices and $\mathbf{u}$, $\mathbf{d}$ are n-vectors.

The method proceeds by first reducing (11.8.30) to a *characteristic normal form* specialized to the linear case, given as

$$\sum_{j=1}^{n} \sigma_{ij} \frac{du_j}{dt} = \sum_{j=1}^{n} \gamma_{ij} u_j + \delta_i \qquad \text{on} \qquad \frac{dx}{dt} = \lambda_i, \qquad i = 1, ..., n. \quad (11.8.31)$$

[Although a further simplification, the *characteristic form*, is available for linear systems, we work with (11.8.31) since the numerical solution method we use is not much simpler when applied to the characteristic form of the system.] For the linear system (11.8.30), the n equations for the characteristics $dx(t)/dt = \lambda_i(x(t),t)$, $i = 1, ..., n$, can be solved independently of the equations for the $u_i(x,t)$. For each characteristic $x = x_i(t)$, we determine the point $Q_i = (x_{Q_i}, t_Q)$ where the characteristic curve that passes through the point $P = (x_P, t_P)$ intersects the line $t = t_Q$. Once the n intersection points Q_i have been determined, the system of equations for the $u_i(x,t)$ is replaced by the linear (algebraic) system

$$\sum_{j=1}^{n} \left[\sigma_{ij}\big|_P + \sigma_{ij}\big|_{Q_i}\right] \frac{u_j - u_j\big|_{Q_i}}{t_P - t_Q} = \sum_{j=1}^{n} \left[\gamma_{ij}\big|_P u_j + \gamma_{ij}\big|_{Q_i} u_j\big|_{Q_i}\right] + \left[\delta_i\big|_P + \delta_i\big|_{Q_i}\right], \qquad (11.8.32)$$

with $i = 1, ..., n$, where u_j, $j = 1, ..., n$, represents the numerical solution values at the point P that are to be determined. [We have replaced derivatives in (11.8.31) by (forward) difference quotients and the coefficients and the nonhomogeneous term have been averaged over their values at P and the Q_i.]

The use of (11.8.32) to determine the u_j at the time $t = t_P$ requires that exact or approximate values of $\mathbf{u}(x,t)$ are known at the points x_{Q_i} at the time $t = t_Q$. If $t = t_Q$ represents the initial time for the initial value problem, it is (generally) assumed that the values $\mathbf{u}(x, t_Q)$ are prescribed. However, if $\mathbf{u}(x, t_Q)$ is only given for a set of values of x that differ from the x_{Q_i}, the Maple procedure $CurveFitting[PolynomialInterpolation]$ or $CurveFitting[Spline]$ can be used to create an (approximate) vector function $\mathbf{u}(x, t_Q)$. [The points $x = x_{Q_i}$ where the characteristics that issue from the point P intersect the line $t = t_Q$ can be used to determine the size of the x-interval for which values of $\mathbf{u}(x, t_Q)$ are needed.] Additionally, if it desired to determine the solution values over a regular grid in the (x,t)-plane, the values of $\mathbf{u}(x,t)$ must be determined for $t > t_P$. Since solution values are available only at grid points for $t = t_P$, interpolation is unavoidable for this process to succeed.

We have constructed the Maple procedure $NumHypSystCharBack$ that finds the value of the solution vector $\mathbf{u}(x,t)$ at a prescribed point $P = (x_P, t_P)$ in terms

of data given at a previous time $t = t_Q$, using the foregoing method. It is assumed that $\mathbf{u}(x, t_Q)$ is known. If $\mathbf{u}(x, t_Q)$ is given only on a set of points, interpolation must be used to create a vector function for the initial condition. The procedure solves the characteristic system numerically, using Maple's default ODE solver. The points $x = x_{Q_i}$ where the characteristics that issue from the point P intersect the line $t = t_Q$ are found and they can be displayed by invoking the global variable $PSet$. Interpolation can then be used to create a required vector function, as indicated above. Any number of equations can be treated by the procedure. The procedure takes the form $NumHypSystCharBack(A, B, C, \mathbf{d}, t = t_P, x = x_P, t = t_Q, \mathbf{u}(x, t_Q))$. This corresponds to the problem formulated above for the system (11.8.30). The output of the procedure displays the value of $\mathbf{u}(x_P, t_P)$ and the points $x = x_{Q_i}$. If a ninth argument *print* is added, additional information relating to the problem is exhibited. The procedure attempts to find an exact solution of the initial value problem if the ninth argument is *exact*. Also, other numeric ODE solvers can be invoked.

Example 11.37. The Numerical Method of Characteristics for a Linear Hyperbolic System. The procedure *NumHypSystCharBack*($A, B, C, \mathbf{d}$, $t = t_P, x = x_P, t = 0, [0, 0]$) applies the foregoing method of characteristics to solve the initial value problem for (11.8.30) with

$$A = \begin{bmatrix} 0 & -1 \\ -x & 0 \end{bmatrix}, \; B = I, \; C = \begin{bmatrix} x & 0 \\ 0 & 0 \end{bmatrix}, \; \mathbf{d} = \begin{bmatrix} -2x^2t^2 \\ 2x^2 - 2xt^2 \end{bmatrix}, \quad (11.8.33)$$

$t_Q = 0$, and the initial condition $\mathbf{u}(x, 0) = \mathbf{0}$. The exact solution of this problem is $\mathbf{u}(x, t) = [u_1(x, t), u_2(x, t)] = [2xt^2, 2x^2t]$.

We apply the procedure three times with $t_P = 0.2$ and $x_P = 0.3,\ 0.5,\ 0.7$. This yields

$$\begin{bmatrix} x & t & u_1 & u_2 & t_Q & x_{Q1} & x_{Q2} \\ 0.3 & 0.2 & 0.024826 & 0.035980 & 0.0 & 0.20046 & 0.41954 \end{bmatrix},$$

$$\begin{bmatrix} x & t & u_1 & u_2 & t_Q & x_{Q1} & x_{Q2} \\ 0.5 & 0.2 & 0.040845 & 0.099980 & 0.0 & 0.36858 & 0.65142 \end{bmatrix}, \quad (11.8.34)$$

$$\begin{bmatrix} x & t & u_1 & u_2 & t_Q & x_{Q1} & x_{Q2} \\ 0.7 & 0.2 & 0.056858 & 0.19598 & 0.0 & 0.54267 & 0.87733 \end{bmatrix}.$$

The exact solution values at these points are $[u_1 = 0.024, u_2 = 0.036], [u_1 = 0.040, u_2 = 0.100], [u_1 = 0.056, u_2 = 0.196]$, respectively.

To determine the numerical solution at the point $P = (0.5, 0.4)$ by using the numerical results obtained at $t = 0.2$, we must use interpolation to construct a vector function whose values at $t = 0.2$ and $x = 0.3,\ 0.5,\ 0.7$ agree with those given by the procedure. On applying the $NumHypSystCharBack$ procedure with the last four arguments given as $t = 0.4,\ x = 0.5,\ t = 0.2,\ [0, 0]$, we determine that the characteristics intersect the line $t = 0.2$ within the x-interval $[0.3, 0.7]$. We construct

the components u_1 and u_2 of this function by using Maple's *Spline* procedure to generate the piecewise linear interpolation functions

$$u_1 = \begin{cases} 0.0005 + 0.081\,x, & x < 0.5, \\ 0.001 + 0.080\,x, & x \geq 0.5, \end{cases} \quad u_2 = \begin{cases} -0.06 + 0.32\,x, & x < 0.5, \\ -0.14 + 0.48\,x, & x \geq 0.5. \end{cases} \tag{11.8.35}$$

On replacing $[0, 0]$ in the foregoing by $[u_1, u_2]$ as given in (11.8.35), the output of the procedure is

$$\begin{bmatrix} x & t & u_1 & u_2 & t_Q & x_{Q1} & x_{Q2} \\ 0.5 & 0.4 & 0.1631 & 0.2032 & 0.2 & 0.3686 & 0.6514 \end{bmatrix}. \tag{11.8.36}$$

There is good agreement with the exact solution value.

Characteristic Difference Methods for Quasilinear Hyperbolic Systems

We consider the quasilinear hyperbolic system of the form (3.3.73)

$$A\mathbf{u}_x(x,t) + B\mathbf{u}_t(x,t) = \mathbf{c}, \tag{11.8.37}$$

where A and B are $n \times n$ matrices and $\mathbf{u}$ and $\mathbf{c}$ are n-component vectors. The elements of A, B, $\mathbf{c}$ can be functions of $(x, t, u_1(x,t), u_2(x,t), \ldots, u_n(x,t))$. As for the linear system considered above, we determine an (approximate) numerical solution at a prescribed point $P = (x_P, t_P)$ in the (x,t)-plane, in terms of initial data for $\mathbf{u}(x,t)$ given at the earlier time $t = t_Q$.

We first reduce (11.8.37) to a *characteristic normal form* given as

$$\sum_{j=1}^{n} \sigma_{ij} \frac{du_j}{dt} = \delta_i \qquad \text{on} \qquad \frac{dx}{dt} = \lambda_i, \qquad i = 1, \ldots, n. \tag{11.8.38}$$

In contrast to the linear system (11.8.30) considered above, the n equations for the characteristics $dx(t)/dt = \lambda_i(x(t), t)$, $i = 1, \ldots, n$, cannot be solved independently of the equations for the $u_i(x,t)$. Nevertheless, for a particular solution $\mathbf{u}(x,t)$, each characteristic $x = x_i(t)$ determines a point $Q_i = (x_{Q_i}, t_Q)$ where the characteristic curve that passes through the point $P = (x_P, t_P)$ intersects the line $t = t_Q$. To determine these points and the solution at the point P numerically, we follow the approach used for the linear case that gave rise to (11.8.32). The system of equations (11.8.38) for the $u_i(x,t)$ and characteristics $x = x_i(t)$ is replaced by the system

$$\sum_{j=1}^{n} \left(\sigma_{ij}\big|_P + \sigma_{ij}\big|_{Q_i} \right) \frac{u_j - u_j\big|_{Q_i}}{t_P - t_Q} = \delta_i\big|_P + \delta_i\big|_{Q_i}, \quad \frac{x_P - x_{Q_i}}{t_P - t_Q} = \frac{\lambda_i\big|_P + \lambda_i\big|_{Q_i}}{2}, \tag{11.8.39}$$

with $i = 1, \ldots, n$, where the u_j, $j = 1, \ldots, n$, represent the numerical solution values at the point P, and the x_{Q_i}, $i = 1, \ldots, n$ are to be determined. As for the linear problem,

$\mathbf{u}(x, t_Q)$ must be given or determined by interpolation. To extend the solution beyond $t = t_P$, interpolation must be used as before.

We have constructed the Maple procedure $NumQuasiHypSystCharBack$ that finds the value of the solution vector $\mathbf{u}(x, t)$ at a prescribed point $P = (x_P, t_P)$ in terms of data given at a previous time $t = t_Q$, using the foregoing method. It is assumed that $\mathbf{u}(x, t_Q)$ is known. If $\mathbf{u}(x, t_Q)$ is given only on a set of points, interpolation must be used to create a vector function for the initial condition. The procedure takes the form *NumQuasiHypSystCharBack*$(A, B, \mathbf{c}, \mathbf{u}, t = t_P, x = x_P, t = t_Q, \mathbf{u}(x, t_Q))$. (The term $\mathbf{u}$ represents a list of the dependent variables $u_1, u_2, \ldots, u_n$.) This corresponds to the problem formulated above for the system (11.8.37). The output of the procedure displays the value of $\mathbf{u}(x_P, t_P)$ and the points $x = x_{Q_i}$. If an eighth argument *print* is added, additional information relating to the problem is exhibited.

Example 11.38. The Numerical Method of Characteristics for a Quasilinear Hyperbolic System. *NumQuasiHypSystCharBack*$(A, B, \mathbf{c}, [u_1, u_2], t = .2, x = .5, t = 0, [0, 0])$ applies the foregoing method of characteristics to solve the initial value problem for (11.8.37) with the matrices A and B defined as in (11.8.33), $\mathbf{c} = [xu_1^2 - 4x^2t^4,\ u_2^2 - 4x^4t^2 + 2x^2 - 2xt^2]$, $t_P = 0.2$, $x_P = 0.5$ and $\mathbf{u}(x, 0) = \mathbf{0}$. The exact solution of the (full) initial value problem is $\mathbf{u}(x, t) = [2xt^2, 2x^2t]$ and is identical with that of the problem in Example 11.37. The output of the procedure is

$$\left[\begin{array}{cccccccc} x & t & u_1 & u_2 & t_Q & x_{Q1} & x_{Q2} \\ 0.5 & 0.2 & 0.03687 & 0.1002 & 0.0 & 0.3686 & 0.6514 \end{array}\right]. \tag{11.8.40}$$

It agrees fairly well with the result given in (11.8.34).

Difference Methods for the Solution of BVPs for Semilinear Elliptic Equations with Variable Coefficients

We consider the two-dimensional semilinear elliptic PDE

$$\alpha(x, y)\, u_{xx}(x, y) + \beta(x, y)\, u_{yy}(x, y) + \gamma(x, y)\, u_x(x, y)$$

$$+ \delta(x, y)\, u_y(x, y) + \omega(x, y)\, u(x, y) = F(x, y, u(x, y)). \tag{11.8.41}$$

[It is assumed that $\alpha(x, y) \neq 0$ and $\beta(x, y) \neq 0$ in the region of interest and that they both have the same sign, so that the PDE is of elliptic type. Although $\gamma(x, y)u(x, y)$ can be included in $F(x, y, u(x, y))$, we use (11.8.41) for convenience.] Let $h > 0$ be an increment in x and $k > 0$ be an increment in y. We replace the partial derivatives by centered difference quotients. The *difference scheme* generated thereby will be used to solve boundary value problems for (11.8.41) in (bounded) rectangular regions in the (x, y)-plane. It is

$$\alpha(x,y)\frac{u(x+h,y)-2u(x,y)+u(x-h,y)}{h^2}+\gamma(x,y)\frac{u(x+h,y)-u(x-h,y)}{2h}$$

$$+\beta(x,y)\frac{u(x,y+k)-2u(x,y)+u(x,y-k)}{k^2}+\delta(x,y)\frac{u(x,y+k)-u(x,y-k)}{2k}$$

$$+\omega(x,y)u(x,y)=F(x,y,u(x,y)). \tag{11.8.42}$$

The *truncation error* is $O(h^2+k^2)$.

As before, the points $(x_i,y_j)=(a+ih,b+jk)$ with $i=0,\pm1,\pm2,\ldots$, and $j=0,\pm1,\pm2,\ldots$, and a and b as arbitrary constants, represent the *grid points* for the general difference scheme (11.8.42). Let $u_{i,j}$ represent the (numerical) approximation of the exact solution $u(x,y)$ of a boundary value problem at the point (x_i,y_j) [i.e., $u_{i,j}\approx u(x_i,y_j)$]. Then (11.8.42) can be written as

$$u_{i,j}=\frac{k^2\alpha_{ij}\,(u_{i+1,j}+u_{i-1,j})}{\sigma}+\frac{h^2\beta_{ij}\,(u_{i,j+1}+u_{i,j-1})}{\sigma}$$

$$+\frac{k^2h\gamma_{i,j}(u_{i+1,j}-u_{i-1,j})}{2\sigma}+\frac{h^2k\delta_{i,j}(u_{i,j+1}-u_{i,j-1})}{2\sigma}-\frac{h^2k^2F_{i,j}}{\sigma}, \tag{11.8.43}$$

where $\sigma=2k^2\alpha_{ij}+2h^2\beta_{ij}-h^2k^2\omega_{ij}$, $\alpha_{ij}=\alpha(x_i,y_j)$—with the other coefficients defined similarly—and $F_{i,j}=F(x_i,y_j,u_{i,j})$.

Boundary value problems of all three kinds are considered for the semilinear equation (11.8.41) in the rectangular region $a<x<b$, $c<y<d$. (We must keep in mind that not every boundary value problem has a solution and the consideration of a nonlinear PDE raises additional problems. We will assume that the problems we consider have unique solutions.) The x and y intervals are subdivided into n_x and n_y equal parts, respectively. The x and y increments are $h=(b-a)/n_x$ and $k=(d-c)/n_y$, respectively. Thus, the grid points are given as $(x_i,y_j)=(a+ih,b+jk)$ with $i=0,1,2,\ldots,n_x$ and $j=0,1,2,\ldots,n_y$. The sides of the rectangle are the boundary lines so that $i=0$, $i=n_x$, $j=0$, $j=n_y$ correspond to $x=a$, $x=b$, $y=c$, $y=d$. respectively. If *Dirichlet conditions* are assigned on the boundary, the grid for the problem comprises only the interior points of the rectangular region, so that $i=1,2,\ldots,n_x-1$ and $j=1,2,\ldots,n_y-1$. If *Neumann* or *Robin conditions* are prescribed on one or more sides of the rectangle, the derivative terms $u_x(x,y)$ and/or $u_y(x,y)$ are replaced by centered differences. The introduction of ghost points and their elimination proceeds as in the discussion presented above for Poisson's equation and is not repeated here. Then, if we are dealing a *Neumann* or *Robin problem*, the grid points are $i=0,1,2,\ldots,n_x$ and $j=0,1,2,\ldots,n_y$. In the case of mixed boundary conditions, grid points that correspond to the sides of the rectangle where Dirichlet conditions are prescribed are eliminated from the solution grid set.

For each of the foregoing boundary value problems, the *difference scheme* (11.8.43) is used. It gives rise to a system of simultaneous, possibly nonlinear equations for the $u_{i,j}$. The number of equations depends on the range of the indices i and j as

was indicated above. Although it may be possible to obtain an exact solution of the system, as was the case for Laplace's and Poisson's equations, we move directly to the iteration method for its solution. To that end, the equations in the difference scheme have been written in the form (11.8.43), for which it is straightforward to carry out an iteration process. The *Jacobi, Gauss-Seidel,* and *SOR methods* can each be applied to (11.8.43), and the general ideas presented above apply here as well. The details as they apply for our problem are not given here. Instead, we present a Maple procedure $NumEllipticSL$ that automates the iteration process. However, it is restricted to the use of the Gauss-Seidel and SOR iteration methods.

The procedure has the form $NumEllipticSL(\alpha, \beta, \gamma, \delta, \omega, F(x,y,u(x,y)), u, x = a..b, lbcfactor, lbc, rbcfactor, rbc, y = c..d, lobcfactor, lobc, ubcfactor, ubc, n_x, n_y, numits, inguess, err, par)$. The first seven arguments prescribe the coefficients and the right-hand side of (11.8.41), and the remaining arguments have the same meaning as those in the procedure $NumLaplace$ presented above. We consider two examples.

Example 11.39. A Dirichlet Problem for a Linear Elliptic Equation. We consider the Dirichlet problem in the unit square $0 < x < 1$, $0 < y < 1$ for the following linear elliptic PDE:

$$\left(2+\cos\left(y^2\right)\right)u_{xx}(x,y)+\left(1+x^2+y^2\right)u_{yy}(x,y)+u_x(x,y)=x^3ye^{xy}$$
$$+\left(2+\cos\left(y^2\right)\right)\left(6xe^{xy}+6x^2ye^{xy}+x^3y^2e^{xy}\right)+\left(1+x^2+y^2\right)x^5e^{xy}+3x^2e^{xy} \tag{11.8.44}$$

with the boundary conditions $u(x,0)=x^3$, $u(x,1)=x^3e^x$, $u(0,y)=0$, $u(1,y)=e^y$. The exact solution of this problem is $u(x,y)=x^3e^{xy}$.

$NumEllipticSL(2+\cos(y^2), 1+x^2+y^2, 1, 0, 0, F(x,y), u, x = 0..1, dirichlet, 0, dirichlet, e^y, y = 0..1, dirichlet, x^3, dirichlet, x^3e^x, 10, 10, 1000, 0, .0001, 1)$ is used, where $F(x,y)$ is the right side of (11.8.44). The x and y intervals are both subdivided into 10 equal parts. The maximum number of iterations allowed is 1000. The initial guess is $u(x,y)=0$. The error tolerance is 0.0001 and the last argument in the procedure, the number 1, signifies that the *Gauss-Seidel method* is to be used. The output of the procedure is

$y \setminus x$	0	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.0
0	0	.0010	.008	.027	.064	.125	.216	.343	.512	.729	1.0
0.1	0	.0012	.009	.028	.067	.131	.229	.369	.557	.800	1.11
0.2	0	.0014	.009	.030	.070	.140	.245	.394	.602	.872	1.22
0.3	0	.0016	.009	.031	.074	.146	.259	.424	.653	.955	1.35
0.4	0	.0018	.010	.032	.077	.155	.278	.457	.704	1.05	1.49
0.5	0	.0019	.010	.033	.081	.164	.293	.489	.766	1.14	1.65
0.6	0	.0020	.011	.035	.084	.172	.311	.522	.827	1.25	1.82
0.7	0	.0020	.011	.036	.087	.180	.333	.563	.901	1.37	2.01
0.8	0	.0019	.011	.036	.091	.189	.352	.601	.969	1.49	2.23
0.9	0	.0016	.011	.037	.093	.197	.373	.644	1.05	1.64	2.46
1.	0	.0011	.010	.036	.095	.206	.393	.689	1.14	1.79	2.72

The first row in the array lists the x grid values, and the first column lists the y grid values. The number of iterations carried out to get this result equals 150. The numerical solution is a good approximation to the exact solution.

Example 11.40. A Linear and Nonlinear Stability Problem. We consider the stability problem of Exercise 4.7.6. The semilinear PDE is

$$u_{xx}(x,y) + u_{yy}(x,y) = -\lambda u(x,y)\left(1 - u^2(x,y)\right), \tag{11.8.45}$$

with $0 < x < \pi$, $y > 0$ (with $\hat{\lambda}$ replaced by λ), and the linearized PDE is

$$u_{xx}(x,y) + u_{yy}(x,y) = -\lambda u(x,y), \tag{11.8.46}$$

with the boundary conditions $u(x,0) = \epsilon h(x)$, $u(0,y) = 0$, $u(\pi,y) = 0$, $\lim_{y\to\infty} u(x,y) = 0$. Although the problem considers the unbounded interval $y > 0$ and requires that $u(x,y) \to 0$ as $y \to \infty$, we must assign a finite range for y in order to apply the procedure *NumEllipticSL* to this problem.

The critical value of the parameter is $\lambda_c = 1$ and we put $h(x) = \sin(x)$ and $\epsilon = 0.01$. We use the procedures $NumEllipticSL(1,1,0,0,0,-\lambda u(x,y),u,x = 0..\pi,dirichlet,0,dirichlet,0,y = 0..100,dirichlet,.01\sin(x),dirichlet,.1,4,20,250,0,.0020,1.6)$ and $NumEllipticSL(1,1,0,0,0,-\lambda u(x,y)(1-u^2(x,y)),u,x = 0..\pi,dirichlet,0,dirichlet,0,y = 0..100,dirichlet,.01\sin(x),dirichlet,.1,4,20,250,0,.0020,1.6)$ for the linear PDE and the semilinear PDE, respectively. In each case a finite y interval $[0,100]$ is used and a small but nonzero value $u(x,100) = 0.1$ is prescribed. We have $u(x,0) = 0.01\sin(x)$ and subdivide the x interval into four parts and the y interval into 20 parts. The *SOR method* with the relaxation parameter $\omega = 1.6$ is used.

First we put $\lambda = 0.9$, which is below the critical value and apply both procedures. The full output of the procedures is not exhibited here. We use the global variable SE to obtain the numerical results at $y = 60$ with $x = \pi/4,\ \pi/2,\ 3\pi/4$. They are $[6.013 \times 10^{-12}, 1.857 \times 10^{-11}, 2.650 \times 10^{-11}]$, for both the linear and semilinear problem. The solution decays, as expected, for increasing y. [However, as y approaches 100 the solution nears the assigned boundary value $u(x,100) = 0.1$ of the procedures.]

Next we apply both procedures with $\lambda = 1.1$, a value that exceeds the critical value. We again use the global variable SE to obtain the numerical results at $y = 60$ with $x = \pi/4,\ \pi/2,\ 3\pi/4$. For the linear case they are $[0.2927 \times 10^{13}, 0.4463 \times 10^{13}, 0.3403 \times 10^{13}]$, which is consistent with the instability of the linear problem. For the semilinear problem we obtain $[.306, .424, .306]$. They are the solution values given by the procedure for $y = 15, 20, \ldots, 90$. The solution does not grow without bound but appears to stabilize around the function $u(x,y) \approx 0.4244\sin(x)$. Replacing the boundary value at $y = 0$ by $u(x,0) = 0.01x(\pi - x)$ yields the same output as before, for y values away from $y = 0$ and $y = 100$. If the number of x intervals is increased, the numerical solution again exhibits stability, but the solution values differ from those given above at common points.

Exercises 11.8

11.8.1. Consider the initial and boundary value problems for the semilinear parabolic PDE given in Example 11.32. Use the exact solution and the procedure $NumPara-bolicLinesSL$ to obtain the values given in the array (11.8.7).

11.8.2. Reproduce the results in Example 11.33 via $NumParabolicLinesSL$.

11.8.3. Carry out Exercise 11.8.2 with the function $h(x) = \sin(\pi x)$ in Example 11.33 replaced by $h(x) = x(\pi - x)$.

11.8.4. Use $NumHyperbolicLinesSL$ and $NumParabolicLinesSL$ to obtain the results given in Example 11.34.

11.8.5. Reproduce the results of Example 11.35 by invoking the procedure $Num-HypScalChar$.

11.8.6. Use the procedure $NumHypSystChar$ to obtain the results of Example 11.36.

11.8.7. Apply the procedure $NumHypSystCharBack$ to obtain the results of Example 11.37.

11.8.8. Use the procedure $NumQuasiHypSystChar$ to obtain the results of Example 11.38.

11.8.9. Use the procedure $NumEllipticSL$ to reproduce the results of Example 11.39.

11.8.10. Use the procedure $NumEllipticSL$ to verify the stability results of Example 11.40.

11.9 FINITE DIFFERENCE METHODS FOR HIGHER-DIMENSIONAL PDEs

In this section we extend a number of the numerical methods introduced above to problems in two or three spatial dimensions for the heat and wave equations and to three-dimensional problems for Laplace's and Poisson's equations. Forward and backward difference schemes in two dimensions and the method of lines in two and three dimensions are introduced for the numerical solution of the heat and wave equations. Iteration schemes are presented for the numerical solution of Laplace's and Poisson's equations in three dimensions. Maple procedures have been constructed to carry out each of these numerical schemes. Systems of hyperbolic equations and PDEs with variable coefficients are not considered.

Explicit Forward Difference Method for the Two-Dimensional Heat Equation

The two-dimensional nonhomogeneous heat or diffusion equation is

$$u_t(x,y,t) - c^2(u_{xx}(x,y,t) + u_{yy}(x,y,t)) = F(x,y,t). \tag{11.9.1}$$

Let $h_x > 0$ and $h_y > 0$ be increments in x and y, respectively, and $k > 0$ be an increment in t. We replace $u_t(x,y,t)$ by a *forward first order difference quotient* in t and $u_{xx}(x,y,t)$ and $u_{yy}(x,y,t)$ by *centered second order difference quotients* in x and y, respectively. With $r_x = kc^2/h_x^2$ and $r_y = kc^2/h_y^2$, we obtain the *explicit forward difference scheme*

$$u(x,y,t+k) = (1 - 2r_x - 2r_y)\,u(x,t) + r_x\,(u(x+h_x,y,t) + u(x-h_x,y,t))$$

$$+r_y\,(u(x,y+h_y,t) + u(x,y-h_y,t)) + kF(x,y,t), \tag{11.9.2}$$

after some simplification. The *truncation error* is readily found to be $O(k+h_x^2+h_y^2)$, so that (11.9.2) is *consistent* with the heat equation (11.9.1).

The points $(x_p, y_j, t_m) = (a + ph_x, b + jh_y, t_0 + mk)$, with $p = 0, \pm 1, \pm 2, \ldots$, $j = 0, \pm 1, \pm 2, \ldots$, $m = 0, 1, 2, \ldots$, and a, b, and t_0 as arbitrary constants, represent the *grid points* for the difference scheme (11.9.2). Let $u_{p,j,m}$ represent the (numerical) approximation of the exact solution $u(x,y,t)$ of a given boundary value problem at the point (x_p, y_j, t_m). Then (11.9.2) can be given as

$$u_{p,j,m+1} = (1 - 2r_x - 2r_y)u_{p,j,m} + r_x(u_{p+1,j,m} + u_{p-1,j,m})$$

$$+ r_y(u_{p,j+1,m} + u_{p,j-1,m}) + kF_{p,j,m}, \tag{11.9.3}$$

with $F_{p,j,m} = F(x_p, y_j, t_m)$.

The difference equation (11.9.3) with $F = 0$ characterizes a two-dimensional random walk, with r_x as the probability of a step to the right or left (i.e., in the x direction) and r_y as the probability of a step up or down (i.e., in the y direction). For the coefficients in the resulting difference equation to represent probabilities, we must restrict the values of the nonnegative r_x and r_y so that $r_x + r_y \leq 1/2$.

The von Neumann stability condition for the difference scheme is obtained by inserting

$$u_{p,j,m} = \alpha^m \exp(i(\beta p + \sigma j)) \tag{11.9.4}$$

into (11.9.3) with $F = 0$. This yields $\alpha = 2r_x \cos(\beta) + 2r_y \cos(\sigma) + 1 - 2r_x - 2r_y$ after some elementary simplification. For stability, the *amplification factor* α must satisfy $-1 \leq \alpha \leq 1$ as β and σ range over the interval $[0, 2\pi]$. Using elementary calculus to determine the maximum and minimum of $\alpha = \alpha(\beta, \sigma)$, we easily conclude that the minimum equals $1 - 4r_x - 4r_y$ and the maximum is 1. This yields the stability

condition $r_x+r_y \leq 1/2$. Thus, the explicit forward difference scheme is *conditionally stable* with the stability condition

$$r_x + r_y = kc^2/h_x^2 + kc^2/h_y^2 \leq 1/2. \tag{11.9.5}$$

We consider initial conditions at $t = t_0$ and boundary conditions of all three kinds for the heat equation (11.9.1) in the rectangular region $a < x < b$, $c < y < d$. The x and y intervals are subdivided into n_x and n_y equal parts, respectively. The x and y increments are $h_x = (b-a)/n_x$ and $h_y = (d-c)/n_y$, respectively, and the t increment is k. Thus, the x and y grid points are given as above, but with $p = 0, 1, 2, \ldots, n_x$ and $j = 0, 1, 2, \ldots, n_y$. The sides of the rectangle are boundary lines (for each value of t) so that $p = 0$, $p = n_x$, $j = 0$, $j = n_y$ correspond to $x = a$, $x = b$, $y = c$, $y = d$, respectively. The size of the grid depends on the type of boundary conditions assigned for the problem, as discussed in Section 11.4.

We consider an initial and boundary value problem for the heat equation (11.9.1) in the region $a < x < b$, $c < y < d$, $t > t_0$ with the initial condition $u(x, y, t_0) = f(x, y)$. Boundary conditions of all three kinds can be prescribed on the four sides of the rectangular region. If Dirichlet conditions are assigned, for example, we have $u(a, y, t) = g(y, t), u(b, y, t) = h(y, t), u(x, c, t) = r(x, t), u(x, d, t) = s(x, t)$.

We have constructed the procedure $NumHeatForw2d(c^2, F(x, y, t), t = t_0..t_f, f(x, y), x = a..b, lbcfactor, g(y, t), rbcfactor, h(y, t), y = c..d, lobcfactor, r(x, t), ubcfactor, s(x, t), n_x, n_y, k)$ for the numerical solution of the foregoing initial and boundary value problems. The second argument specifies the initial time $t = t_0$ and the time $t = t_f$ when the solution is to be found. The last argument specifies the time increment k. All other arguments are defined as in the $LaplaceMatrix$ procedure of Section 11.4.

We consider only one example.

Example 11.41. A Neumann Problem for the Heat Equation. We consider an initial and boundary value problem for the homogeneous heat equation (11.9.1) with $c = 1$ and $F(x, y, t) = 0$. The initial condition is $u(x, y, 0) = xy^2$ and we have $0 < x < 1$, $0 < y < 1$ with homogeneous Neumann boundary conditions prescribed on the boundary. The x and y intervals are both subdivided into 10 equal parts, and the solution is to be found at $t = 5$. It is expected that the solution will have reached a steady state at that time. We invoke the procedure $NumHeatForw2d(1, 0, t = 0..5., xy^2, x = 0..1, 0, 0, 0, 0, y = 0..1, 0, 0, 0, 0, 10, 10, k)$. If we put $k = 0.001$, then $r_x+r_y = 0.2$ and the scheme is stable. The solution values at the grid points (with $t = 5$) are either 0.167 or 0.168. For example, the global variable $SolHF$ determines that the numerical solution at $(x, y, t) = (0, 0.1, 5)$ equals $SolHF(0., 0.10) = 0.167$. This represents the average of the initial temperature integrated over the unit square, which is given as $\int_0^1 \int_0^1 xy^2\, dx\, dy = 0.1666666667$.

If we put $k = 0.01$, then $r_x + r_y = 2$ and the scheme is unstable. The global variable $SolHF$ has the value $SolHF(0., 0.10) = 0.159 \times 10^{416}$. The solution values at all the other grid points are of the same order of magnitude.

Implicit Backward Difference Methods for the Two-Dimensional Heat Equation

We replace $u_t(x, y, t)$ by a *backward first order difference quotient* in t and $u_{xx}(x, y, t)$ and $u_{yy}(x, y, t)$ by *centered second order difference quotients* in x and y, respectively, in the heat equation (11.9.1). On retaining the notation introduced above, we obtain the *implicit backward difference scheme*

$$u_{p,j,m-1} = (1 + 2r_x + 2r_y)u_{p,j,m} - r_x(u_{p+1,j,m} + u_{p-1,j,m}) - r_y(u_{p,j+1,m} + u_{p,j-1,m}) - kF_{p,j,m}. \tag{11.9.6}$$

The *truncation error* for the difference scheme is $O(k + h_x^2 + h_y^2)$, so that (11.9.6) is *consistent* with the heat equation (11.9.1).

The von Neumann stability condition for the difference scheme is obtained by inserting (11.9.4) into (11.9.6), where we put $F = 0$. This yields

$$\alpha = \frac{1}{(2 - 2\cos(\beta))r_x + (2 - 2\cos(\sigma))r_x + 1} \tag{11.9.7}$$

after some simplification. Clearly, $|\alpha| \leq 1$ for all β and σ, so that the implicit backward difference scheme is *unconditionally stable*.

The *implicit Crank-Nicolson difference scheme* for (11.9.1) is constructed analogously to that for the one-dimensional heat equation. We obtain

$$(1 + r_x + r_y)u_{p,j,m} - \frac{r_x}{2}(u_{p+1,j,m} + u_{p-1,j,m}) - \frac{r_y}{2}(u_{p,j+1,m} + u_{p,j-1,m}) = (1 - r_x - r_y)u_{p,j,m-1} + \frac{r_x}{2}(u_{p+1,j,m-1} + u_{p-1,j,m-1}) + \frac{r_y}{2}(u_{p,j+1,m-1} + u_{p,j-1,m-1}) + \frac{k}{2}F_{p,j,m} + \frac{k}{2}F_{p,j,m-1}. \tag{11.9.8}$$

The *truncation error* for the difference scheme is $O(k^2 + h_x^2 + h_y^2)$, so that (11.9.8) is *consistent* with the heat equation (11.9.1). The difference scheme is *unconditionally stable*, but we do not demonstrate this.

The procedure $NumHeatBackw2d(c^2, F(x, y, t), t = t_0..t_f, f(x, y), x = a..b, lbcfactor, g(y, t), rbcfactor, h(y, t), y = c..d, lobcfactor, r(x, t), ubcfactor, s(x, t), n_x, n_y, k, par)$ constructs numerical solutions of the initial and boundary value problems discussed above. All arguments are as in the $NumHeatForw2d$ except for the added last argument *par*, which must take the value 0.5 or 1. If the value is 0.5 the *Crank-Nicolson scheme* is used, and if the value is 1, the *implicit backward difference scheme* is invoked.

Both implicit schemes require that a system of simultaneous linear equations be solved in moving from one time step to the next. They do not take on as simple a form as in the one-dimensional case. Alternating direction methods, which are presented below, introduce systems of simultaneous equations that are easier to solve.

We apply the procedure $NumHeatBackw2d$ to the *Neumann problem* considered in Example 11.41, using the same arguments. Both of the foregoing implicit methods are used. For each choice of k prescribed above, the steady-state values 0.167 or 0.168 are obtained at each grid point.

Peaceman-Rachford and Douglas-Rachford ADI Difference Methods for the Two-Dimensional Heat Equation

As indicated above, the use of the implicit backward and Crank-Nicolson difference schemes in two dimensions requires the solution of simultaneous linear systems that do not have the simple tridiagonal forms encountered in the one-dimensional versions of these schemes. To retain the simplicity of the one-dimensional schemes, Peaceman and Rachford developed a modification of the Crank-Nicolson scheme, and Douglas and Rachford developed a modification of the implicit backward scheme. Both methods replace the two-dimensional schemes by two (effectively) one-dimensional schemes, one in the x direction and the other in the y direction. They are both implicit methods. As such, they are referred to as *alternating direction implicit* (ADI) methods. Both methods are consistent with the nonhomogeneous heat equation and have the same truncation error as that of the scheme they modify, and they are both *unconditionally stable*. Each of these methods can be derived by factoring the difference operators that arise in the Crank-Nicolson and backward difference methods. We do not present their derivation but state the difference schemes directly. Their consistency and stability properties can then be verified directly.

The *Peaceman-Rachford alternating direction implicit difference scheme* is given as

$$(1+r_x)\,w_{p,j,m} - \frac{r_x}{2}\,(w_{p+1,j,m} + w_{p-1,j,m}) - \frac{k}{4}\,(F_{p,j,m-1} + F_{p,j,m})$$

$$= (1-r_y)\,u_{p,j,m-1} + \frac{r_y}{2}\,(u_{p,j+1,m-1} + u_{p,j-1,m-1})\,, \tag{11.9.9}$$

$$(1+r_y)\,u_{p,j,m} - \frac{r_y}{2}\,(u_{p,j+1,m} + u_{p,j-1,m}) - \frac{k}{4}\,(F_{p,j,m-1} + F_{p,j,m})$$

$$= (1-r_x)\,w_{p,j,m} + \frac{r_x}{2}\,(w_{p+1,j,m} + w_{p-1,j,m})\,. \tag{11.9.10}$$

The *Douglas-Rachford alternating direction implicit difference scheme* is given as

$$(1+r_x)\,w_{p,j,m} - r_x\,(w_{p+1,j,m} + w_{p-1,j,m}) + \frac{k}{2}F_{p,j,m}$$

$$= (1-r_y)\,u_{p,j,m-1} + r_y\,(u_{p,j+1,m-1} + u_{p,j-1,m-1})\,, \tag{11.9.11}$$

$$(1+r_y)\,u_{p,j,m} - r_y\,(u_{p,j+1,m} + u_{p,j-1,m}) - \frac{k}{2}F_{p,j,m}$$

$$= w_{p,j,m} - r_y\,(u_{p,j+1,m-1} - 2u_{p,j,m-1} + u_{p,j-1,m-1})\,. \tag{11.9.12}$$

The (subscript) notation used for the terms $u_{p,j,m}$ and $w_{p,j,m}$ is as before. At each time step $t = t_m$, $m \geq 1$, it is assumed that $u_{p,j,m-1}$ is known. Then (11.9.9) and (11.9.11) yield simultaneous linear tridiagonal systems for $w_{p,j,m}$. Once the $w_{p,j,m}$ are determined, (11.9.10) and (11.9.12) yield simultaneous linear tridiagonal systems for $u_{p,j,m}$. Each of the initial and boundary value problems considered above can be solved by either of these methods.

The procedures $NumHeatPRADI2d$ and $NumHeatDRADI2d$ obtain numerical solutions of initial and boundary value problems (of the form that were solved above by explicit forward and implicit backward difference methods) by the use of Peaceman-Rachford and Douglas-Rachford methods. The arguments for each procedure are the same as in the procedure $NumHeatForw2d$.

Example 11.42. A Mixed Problem for the Heat Equation. We consider the initial and boundary value problem for the nonhomogeneous heat equation $u_t(x,y,t) - u_{xx}(x,y,t) - u_{yy}(x,y,t) = \cos(\pi x)\left(-y^2 + \pi^2 y^2 - 2\right) e^{-t}$, with $0 < x, y < 1$, $t > 0$, the initial condition $u(x,y,0) = \cos(\pi x)\, y^2$, and the mixed boundary condition $u_x(0,y,t) = 0$, $u(1,y,t) = -y^2 e^{-t}$, $u_y(x,0,t) - u(x,0,t) = 0$, $u(x,1,t) = \cos(\pi x) e^{-t}$. The exact solution of this problem is $u(x,y,t) = \cos(\pi x)\, y^2 e^{-t}$.

We subdivide the x and y intervals into four equal parts, with the t increment equal to 0.01. The numerical solution is to be found at $t = 1$. We begin with the *Peaceman-Rachford method* and use $NumHeatPRADI2d(1, \cos(\pi x)(-y^2 + \pi^2 y^2 - 2)\, e^{-t}, t = 0..1., \cos(\pi x)\, y^2, x = 0..1, 0, 0, dirichlet, -y^2\, e^{-t}, y = 0..1, 1, 0, dirichlet, \cos(\pi x)\, e^{-t}), 4, 4, .01)$. The Maple output is

$t = 1$	x	0.0	0.25	0.50	0.75	1.0
y						
0.0		0.00267	0.00225	0.00131	0.000421	0.0
0.25		0.0267	0.0193	0.00157	−0.0160	−0.023
0.50		0.0969	0.0689	0.00157	−0.0653	−0.092
0.75		0.212	0.150	0.00107	−0.147	−0.207
1.0		0.368	0.260	$-7.55\ 10^{-11}$	−0.260	−0.368

(11.9.13)

Next we use the *Douglas-Rachford method* and enter $NumHeatDRADI2d$ with the same arguments as above. The Maple output is

$t = 1$	x	0.0	0.25	0.50	0.75	1.0
y						
0.0		0.00113	0.00123	0.00157	0.00230	0.0
0.25		0.0254	0.0184	0.00173	−0.0146	−0.0230
0.50		0.0968	0.0688	0.00140	−0.0657	−0.0920
0.75		0.213	0.151	0.000666	−0.149	−0.207
1.0		0.368	0.260	$-7.55\ 10^{-11}$	−0.260	−0.368

(11.9.14)

It may be noted that the exact solution vanishes when $x = 0.5$, whereas the numerical solution values are small but not zero. The agreement between the numerical and exact solution values can be improved by increasing the number of subdivisions.

Method of Lines for the Two-Dimensional Heat Equation

We continue our discussion of numerical methods for the solution of initial and boundary value problems for the nonhomogeneous heat equation (11.9.1). Let $h_x > 0$ and $h_y > 0$ be increments in x and y, respectively. We replace $u_{xx}(x, y, t)$ and $u_{yy}(x, y, t)$ by *centered second order difference quotients* in x and y, respectively, but retain t as a continuous variable. Our discussion is restricted to a rectangular region in the (x, y)-plane given as $0 \le x \le l$, $0 \le y \le L$. The initial condition is given at $t = 0$, and boundary conditions of all three kinds can be given at $x = 0$, $x = l$, $y = 0$, $y = L$. The x and y intervals are divided into n and m equal parts, respectively. Since t is not discretized, the grid comprises the (x, y) values at which the solution is to be found, and is given as $x_i = il/n$, $y_j = jL/m$ with $i = 0, 1, 2, \ldots, n$, $j = 0, 1, 2, \ldots, m$. Thus $h_x = l/n$ and $h_y = L/m$. The approximate solution is to be found at a given time $t > 0$. The semidiscrete difference scheme can then be written as

$$u'_{i,j}(t) = \frac{c^2}{h_x^2}\left(u_{i+1,j}(t) - 2u_{i,j}(t) + u_{i-1,j}(t)\right)$$

$$+ \frac{c^2}{h_y^2}\left(u_{i,j+1}(t) - 2u_{i,j}(t) + u_{i,j-1}(t)\right) + F_{i,j}(t), \qquad (11.9.15)$$

where $u_{i,j}(t)$ represents the approximation to the exact solution $u(x, y, t)$ of the given problem at the point $(x_i, y_j) = (il/n, jL/m)$ at the time t, and where $F_{i,j}(t) = F(il/n, jL/m, t)$. As before, the set of grid points at which the solution is to be found depends on the type of boundary conditions that are prescribed. Numerical solution values at boundary points are determined as above. The use of the coupled system of ODEs to solve a given initial and boundary value problem for the heat equation numerically is the two-dimensional version of the *method of lines*.

$NumHeatLines2d(c^2, F(x,y,t), t = t_0, f(x,y), x = a..b, lbcfactor, g(y,t), rbcfactor, h(y,t), y = c..d, lobcfactor, r(x,t), ubcfactor, s(x,t), n_x, n_y)$ applies the method of lines to determine numerical solutions of the foregoing initial and boundary value problems. The arguments of the procedure are defined as in $NumHeatForw2d$, with the following differences. The third argument specifies only the initial time $t = t_0$. An optional seventeenth argument can be entered. If it is omitted, the output is a procedure that permits the evaluation of the solution at arbitrary values of t. If there is a seventeenth argument $t = t_m$, the output is an array of grid values at the time $t = t_m$. Maple's default numerical solution method is used to solve the system of ODEs. As is the case for the procedure $NumHeatLines$ that applies the method of lines to the heat equation in one dimension and was discussed in Section 11.2, additional optional arguments can be introduced to force Maple to use alternative numerical methods to solve the system.

Example 11.43. A Neumann Problem for the Heat Equation. We reconsider the initial and boundary value problem presented in Example 11.41. The x and y intervals are both subdivided into 10 equal parts and the solution is to be found at $t = 5$. The solution is expected to have reached a steady state at that time. We invoke the procedure $NumHeatLines2d(1, 0, t = 0, xy^2, x = 0..1, 0, 0, 0, 0, y = 0..1, 0, 0, 0, 0, 10, 10, t = 5)$. The solution values at the grid points (with $t = 5$) are either 0.167 or 0.168. They correspond to the steady-state values of the solution, as shown in Example 11.41. If the final argument $t = 5$ in $NumHeatLines2d$ is removed, the output is a procedure that displays the numerical solution values at various times t. It can be used to determine the rate at which the numerical solution approaches a steady state.

The foregoing results were obtained by Maple using its built-in numerical ODE solver. By inserting optional arguments in the procedure alternative numerical solutions methods with fixed time increments can be used. For example, if we replace $t = 5$ by *method* = *classical[foreuler]*, *stepsize* = 0.001, Maple uses Euler's method to solve the system of ODEs with the time increment given as $k = 0.001$. That is, the time derivatives in the system are replaced by forward difference quotients. The resulting difference scheme is stable and the output for $t = 5$ is again equal to the steady state-values. However, if we increase the value of k and put *stepsize* = 0.01, the difference scheme is unstable and the output is undefined.

Explicit Forward Difference Method for the Two-Dimensional Wave Equation

The two-dimensional nonhomogeneous wave equation for $u(x, y, t)$ is given as

$$u_{tt}(x, y, t) - c^2 \left(u_{xx}(x, y, t) + u_{yy}(x, y, t)\right) = F(x, y, t). \qquad (11.9.16)$$

Let $h_x > 0$ and $h_y > 0$ be increments in x and y, respectively, and $k > 0$ be an increment in t. We replace $u_{tt}(x, y, t)$, $u_{xx}(x, y, t)$ and $u_{yy}(x, y, t)$ by *centered second order difference quotients*. With $r_x = ck/h_x$ and $r_y = ck/h_y$, we obtain the *explicit forward difference scheme*

$$u(x, y, t + k) = 2(1 - r_x^2 - r_y^2)\, u(x, t) + r_x^2 \left(u(x + h_x, y, t) + u(x - h_x, y, t)\right)$$

$$+ r_y^2 \left(u(x, y + h_y, t) + u(x, y - h_y, t)\right) - u(x, y, t - k) + k^2 F(x, y, t), \quad (11.9.17)$$

after some simplification. The *truncation error* is $O(k^2 + h_x^2 + h_y^2)$ so that (11.9.17) is *consistent* with the wave equation (11.9.16).

The points $(x_p, y_j, t_m) = (a + ph_x, b + jh_y, t_0 + mk)$ with $p = 0, \pm 1, \pm 2, \ldots$, $j = 0, \pm 1, \pm 2, \ldots$, and $m = 0, 1, 2, \ldots$, and a, b, and t_0 as arbitrary constants represent the *grid points* for the difference scheme (11.9.17). Let $u_{p,j,m}$ represent the (numerical) approximation of the exact solution $u(x, y, t)$ of a given boundary value problem at the point (x_p, y_j, t_m), then (11.9.17) becomes

$$u_{p,j,m+1} = 2(1 - r_x^2 - r_y^2)u_{p,j,m} + r_x^2(u_{p+1,j,m} + u_{p-1,j,m}) + r_y^2(u_{p,j+1,m} + u_{p,j-1,m}) - u_{p,j,m-1} + k^2 F_{p,j,m}, \quad (11.9.18)$$

with $F_{p,j,m} = F(x_p, y_j, t_m)$.

The *von Neumann stability condition* for the difference scheme is obtained by inserting $u_{p,j,m} = \alpha^m \exp(i(\beta p + \sigma j))$ into (11.9.18) with $F = 0$. This yields the quadratic equation $\alpha^2 + \left(4r_x^2 \sin^2(\beta/2) + 4r_y^2 \sin^2(\sigma/2) - 2\right)\alpha + 1 = 0$. The two roots can be given as $\alpha = 1 - 2r_x^2 \sin^2(\beta/2) - 2r_y^2 \sin^2(\sigma/2) \pm 2\sqrt{(r_x^2 \sin^2(\beta/2) + r_y^2 \sin^2(\sigma/2) - 1)(r_x^2 \sin^2(\beta/2) + r_y^2 \sin^2(\sigma/2))}$.

If we set $\beta = \sigma = \pi$ in the above, and compare with (11.5.18) and the ensuing discussion for the one-dimensional wave equation, we conclude that if $r_x^2 + r_y^2 > 1$, the scheme is unstable. If $r_x^2 + r_y^2 \leq 1$, the roots are complex conjugates. Since the quadratic equation shows that the product of the roots equals 1, we conclude, for each root, that $|\alpha| = 1$, and this signifies stability. Consequently, the explicit forward difference scheme is *conditionally stable* with

$$r_x^2 + r_y^2 = \frac{c^2k^2}{h_x^2} + \frac{c^2k^2}{h_y^2} \leq 1 \quad (11.9.19)$$

as the *von Neumann stability condition*.

To determine the *CFL stability condition* for the difference scheme, we recall that the domain of dependence of a point (ξ, η) at the time t at the initial time $t = 0$ is given by the disk $c^2t^2 \leq (x-\xi)^2 + (y-\eta^2)$. Now, if $t = mk$, the grid values whose distance from (ξ, η) at $t = 0$ is maximal are given as $(\xi \pm mh_x, \eta \pm mh_y)$. For the domain of dependence of the wave equation to be contained within the numerical domain of dependence, we must require that $c^2k^2m^2 \leq m^2h_x^2 + m^2h_y^2$. This yields the *CFL stability condition* $1 \leq 1/r_x^2 + 1/r_y^2$. We note that if the von Neumann condition (11.9.19) is satisfied, so is the CFL condition, but the converse is not always true. If $r_x = r_y = r$, (11.9.19) gives $r^2 \leq 1/2$ while the CFL condition yields $r^2 \leq 2$. Additionally, if $r_x^2 = 1/2$, $r_y^2 = 3/4$, for example, the CFL condition is met but not the von Neumann condition.

We consider an initial and boundary value problem for the wave equation (11.9.16) in the region $a < x < b$, $c < y < d$, $t > t_0$ with the initial conditions

$$u(x, y, t_0) = f_1(x, y), \; u_t(x, y, t_0) = f_2(x, y). \quad (11.9.20)$$

Boundary conditions of all three kinds can be prescribed on the four sides of the rectangular region, as was done for the heat equation. The explicit difference scheme is a two-step method, and the initial values at the first two time steps are prescribed as in the one-dimensional case. The boundary conditions used in the difference scheme are determined as for the two-dimensional heat equation.

NumWaveForw2d$(c, F(x,y,t), t = t_0..t_f, f_1(x,y), f_2(x,y), x = a..b, lbcfactor, g(y,t), rbcfactor, h(y,t), y = c..d, lobcfactor, r(x,t), ubcfactor, s(x,t), n_x, n_y, k)$ solves the foregoing initial and boundary value problems for the wave equation (11.9.16) numerically, using the explicit forward difference scheme. Except for the

addition of the fifth argument that specifies $u_t(x, y, t_0)$, the arguments are the same as those defined for the procedure $NumHeatForw2d$.

Example 11.44. A Dirichlet Problem for the Wave Equation. We use the procedure $NumWaveForw2d(1, 2x^2y^2 - 2y^2t^2 - 2x^2t^2, t = 0..1, 0, 0, x = 0..1, dirichlet, 0, dirichlet, y^2t^2, y = 0..1, dirichlet, 0, dirichlet, x^2t^2, 4, 4, k)$. It solves the Dirichlet problem for the wave equation (11.9.16) with $F(x, y, t) = 2x^2y^2 - 2y^2t^2 - 2x^2t^2$ and homogeneous initial conditions. The exact solution of the foregoing initial value problem is $u(x, y, t) = x^2y^2t^2$. We consider the region $0 < x, y < 1$ and impose Dirichlet conditions on $x = 0$, $x = 1$, $y = 0$, $y = 1$ for $t > 0$ that equal the values of the exact solution on the boundaries. The x and y intervals are subdivided into four equal parts. The t increment k is to be specified, and the numerical solution is to be found at $t = 1$.

First, we put $k = 0.25$, so that $r_x^2 = r_y^2 = 1$, $r_x^2 + r_y^2 = 2$, and $1/r_x^2 + 1/r_y^2 = 2$. Consequently, the scheme satisfies the *CFL* condition but is not von Neumann stable. The solution values at the interior grid points with $y = 0.75$ are given as $SolWF(.25, .75) = -.360$, $SolWF(.50, .75) = 1.50$, $SolWF(.75, .75) = -1.49$, on using the global variable $SolWF$. The values of the exact solution at these points are $u(.25, .75, 1) = .035$, $u(.50, .75, 1) = .141$, $u(.75, .75, 1) = .316$. The results are far apart, indicating an instability. Next, we set $k = 0.005$. Then $r_x^2 + r_y^2 = 0.0008$, so that the difference scheme is von Neumann stable. The numerical solution values at the interior grid points with $y = 0.75$ are now given as $SolWF(.25, .75) = .034$, $SolWF(.50, .75) = .139$, $SolWF(.75, .75) = .317$.

Implicit Backward Difference Method for the Two-Dimensional Wave Equation

Whereas an entire family of implicit difference schemes was introduced for the one-dimensional wave equation, we restrict ourselves to presenting a single implicit backward difference scheme for the two-dimensional case. Centered difference quotients are again used to approximate all second derivative terms in the wave equation. The approximation of $u_{tt}(x, y, t)$ involves the time values $t + k$, t, $t - k$, but $u_{xx}(x, y, t)$ and $u_{yy}(x, y, t)$ are both approximated by the average of their centered differences taken at the times $t + k$ and $t - k$. The same is done for the nonhomogeneous term $F(x, y, t)$. As a result, the values of $u(x, y, t + k)$ must be determined implicitly. In the notation introduced above, the difference equation takes the form

$$\begin{aligned} u_{p,j,m+1} = {} & 2u_{p,j,m} - u_{p,j,m-1} + \frac{r_x^2}{2}(u_{p+1,j,m+1} - 2u_{p,j,m+1} + u_{p-1,j,m+1}) \\ & + \frac{r_y^2}{2}(u_{p,j+1,m+1} - 2u_{p,j,m+1} + u_{p,j-1,m+1}) + \frac{k^2}{2}F_{p,j,m+1} \\ & + \frac{r_x^2}{2}(u_{p+1,j,m-1} - 2u_{p,j,m-1} + u_{p-1,j,m-1}) \\ & + \frac{r_y^2}{2}u_{p,j+1,m-1} - 2u_{p,j,m-1} + u_{p,j-1,m-1}) + \frac{k^2}{2}F_{p,j,m-1}. \end{aligned} \quad (11.9.21)$$

This is the *implicit backward difference scheme* for the two-dimensional wave equation (11.9.16). The *truncation error* is $O(k^2 + h_x^2 + h_y^2)$ so that (11.9.21) is *consistent* with the wave equation (11.9.16). The difference scheme can be shown to be *unconditionally stable.*

The implicit scheme can be used to solve initial and boundary value problems for the wave equation, as formulated above for the explicit scheme. However, as for the implicit schemes for the two-dimensional heat equation, it is necessary to solve a system of linear equations to determine the numerical solution at each time step. Alternating direction methods were introduced above to simplify this process for the heat equation. Related techniques have been developed for the wave equation, but they are not presented here. The Maple procedure $NumWaveBackw2d$ solves the foregoing initial and boundary value problems for the wave equation (11.9.16) numerically using the implicit backward difference scheme. The arguments are the same as those defined for the procedure $NumWaveForw2d$.

We apply $NumWaveBackw2d$ to the problem of Example 11.44 with $k = 0.25$, which led to an unstable forward difference scheme in the example. Using the global variable $SolWB$ yields $SolWB(.25, .75) = .0021, SolWB(.50, .75) = -.00608, SolWB(.75, .75) = -.0056$.

Method of Lines for the Two-Dimensional Wave Equation

We refer to our discussion of the method of lines for (11.9.1) and present the corresponding method for (11.9.16). We deal only with the rectangular region in the (x, y)-plane, given as $0 \leq x \leq l$, $0 \leq y \leq L$. Initial conditions (11.9.20) are given at $t = 0$, and boundary conditions of all three kinds can be given at $x = 0$, $x = l$, $y = 0$, $y = L$. The x and y intervals are divided into n and m equal parts, respectively. The t variable is not discretized, so that the grid is given as $x_i = il/n$, $y_j = jL/m$ with $i = 0, 1, 2, \ldots, n$, $j = 0, 1, 2, \ldots, m$. Thus $h_x = l/n$ and $h_y = L/m$. The approximate solution is to be found at a specified time $t > 0$. The semidiscrete difference scheme is then given as

$$u''_{i,j}(t) = \frac{c^2}{h_x^2}(u_{i+1,j}(t) - 2u_{i,j}(t) + u_{i-1,j}(t))$$

$$+\frac{c^2}{h_y^2}(u_{i,j+1}(t) - 2u_{i,j}(t) + u_{i,j-1}(t)) + F_{i,j}(t), \qquad (11.9.22)$$

where $u_{i,j}(t)$ is the approximation to the exact solution $u(x, y, t)$ of the problem at the point $(x_i, y_j) = (il/n, jL/m)$ at the time t, and where $F_{i,j}(t) = F(il/n, jL/m, t)$. As before, the grid points at which the solution is to be found depend on the boundary conditions prescribed. Numerical solution values at boundary points are determined as above. The use of the system of ODEs to solve a given initial and boundary value problem for the wave equation numerically is the two-dimensional version of the *method of lines.*

The procedure $NumWaveLines2d(c, F(x,y,t), t = t_0, f_1(x,y), f_2(x,y), x = a..b, lbcfactor, g(y,t), rbcfactor, h(y,t), y = c..d, lobcfactor, r(x,t), ubcfactor, s(x,t), n_x, n_y)$ applies the method of lines to determine numerical solutions of the foregoing initial and boundary value problems. The arguments of the procedure are defined as in $NumWaveForw2d$ with the following differences. The third argument specifies only the initial time $t = t_0$. There is an optional eighteenth argument. If it is omitted, the output is a procedure that permits the evaluation of the solution at arbitrary values of t. If the eighteenth argument is $t = t_m$, the output is an array of grid values at the time $t = t_m$. Maple's default numerical solution method is used to solve the system of ODEs. As is the case for the procedure $NumWaveLines$ that applies the method of lines to the wave equation in one dimension and was discussed in Section 11.3, additional optional arguments can be introduced to force Maple to use alternative numerical methods to solve the system.

If we apply $NumWaveLines2d$ (using the built-in Maple ODE solver) to the problem of Example 11.44, the global variable obtains the numerical solution values at $t = 1$: $SolWL(.25, .75) = .0352, SolWL(.50, .75) = .141, SolWL(.75, .75) = .316$. There is excellent agreement with the exact solution.

Method of Lines for the Three-Dimensional Heat Equation

We consider initial and boundary value problems for the three-dimensional nonhomogeneous heat equation

$$u_t(x,y,z,t) - c^2\nabla^2 u(x,y,z,t) = F(x,y,z,t). \tag{11.9.23}$$

The initial condition is $u(x,y,z,t_0) = f(x,y,z)$. The spatial region (for any fixed value of t) is bounded by the planes $x = a_1$, $x = b_1$, $y = a_2$, $y = b_2$, $z = a_3$, $z = b_3$, and boundary conditions of all three kinds can be assigned on each of the bounding planes.

To determine the numerical solution of a problem, the three second order spatial derivatives in (11.9.23) are replaced by centered difference quotients with increments h_x, h_y and h_z, and t is left intact. The grid points are given as $(x_i, y_j, z_l) = (\alpha + ih_x, \beta + jh_y, \gamma + lh_z)$ with $i = 0, \pm1, \pm2, \ldots$, $j = 0, \pm1, \pm2, \ldots$, $l = 0, \pm1, \pm2, \ldots$, (where α, β, γ are arbitrary constants). Then, $u_{i,j,l}(t)$ represents the (numerical) solution of an initial and boundary value problem for (11.9.23) at (x_i, y_j, z_l, t). The resulting system of ODEs

$$\frac{du_{i,j,l}(t)}{dt} = F_{i,j,l}(t) + \frac{c^2}{h_x^2}\left(u_{i+1,j,l}(t) - 2u_{i,j,l}(t) + u_{i-1,j,l}(t)\right) \tag{11.9.24}$$

$$+\frac{c^2}{h_y^2}\left(u_{i,j+1,l}(t) - 2u_{i,j,l}(t) + u_{i,j-1,l}(t)\right) + \frac{c^2}{h_z^2}\left(u_{i,j,l+1}(t) - 2u_{i,j,l}(t) + u_{i,j,l-1}(t)\right),$$

where $F_{i,j,l}(t) = F(x_i, y_j, z_l, t)$, represents the *method of lines* for the numerical solution of (11.9.23).

The Maple procedure $NumHeatLines3d(c^2, F(x,y,z,t), t = t_0, f(x,y,z), x = a_1..b_1, lxf, g_1(y,z,t), rxf, g_2(y,z,t), y = a_2..b_2, lyf, r_1(x,z,t), ryf, r_2(x,z,t), z = a_3..b_3, lzf, s_1(x,y,t), rzf, s_2(x,y,t), n_x, n_y, n_z)$ applies the method of lines to determine numerical solutions of the foregoing initial and boundary value problems. The arguments of the procedure are as in $NumHeatLines2d$, except for the additional arguments relating to the z variable. Thus, if $lzf = dirichlet$, the boundary condition $u(x,y,a_3) = s_1(x,y,t)$ is prescribed. If $lzf = 0$, the Neumann boundary condition $u_z(x,y,a_3) = -s_1(x,y,t)$ is specified. If $lzf = \lambda$, the Robin boundary condition $-u_z(x,y,a_3) + \lambda u(x,y,a_3) = s_1(x,y,t)$ is prescribed. There is an optional twenty-third argument. If it is omitted, the output is a procedure that permits the evaluation of the solution at arbitrary values of t. If the twenty-third argument is $t = t_m$, the output is a list of arrays of grid values at the time $t = t_m$. Maple's default numerical solution method is used to solve the system of ODEs. As before, additional optional arguments can be introduced to force Maple to use alternative numerical methods to solve the system.

Example 11.45. A Dirichlet Problem for the Heat Equation. The function $u(x,y,z,t) = x^2yze^{-t}$ is a solution of the heat equation (11.9.23) with $c^2 = 1$ and $F(x,y,z,t) = -(x^2+2)yze^{-t}$. We construct an initial and boundary value problem for (11.9.23) in the unit cube $0 < x, y, z < 1$ with the initial condition $u(x,y,z,0) = x^2yz$ and Dirichlet boundary conditions determined by the values of $u(x,y,z,t) = x^2yze^{-t}$ on the boundary. The x, y, z intervals are each divided into four equal parts. We apply $NumHeatLines3d$ to find the numerical solution of this problem at $t = 1$. The arguments are given in terms of $u(x,y,z,t)$ as defined above. For example, we put $f(x,y,z) = x^2yz$, $a_1 = 0$, $b_1 = 1$, $lxf = dirichlet$, and $g_1(y,z,t) = 0$.

Adding a twenty-third argument $t = 1$ in the procedure, yields the output

$$\left[\begin{array}{cccccc} t = 1.0 & & & & & \\ z = 0.5 & x & 0.0 & 0.25 & 0.50 & 0.75 & 1.0 \\ y & & & & & & \\ 0.0 & & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.25 & & 0.0 & 0.00287 & 0.0115 & 0.0259 & 0.0460 \\ 0.50 & & 0.0 & 0.00575 & 0.0230 & 0.0517 & 0.0920 \\ 0.75 & & 0.0 & 0.00862 & 0.0345 & 0.0776 & 0.138 \\ 1.0 & & 0.0 & 0.0115 & 0.0460 & 0.103 & 0.184 \end{array}\right], \qquad (11.9.25)$$

which represents the numerical solution values at the time $t = 1$ and at the grid points $(x_i, y_j, 0.5)$. Four additional arrays that give numerical solution values for $z = 0,\ 0.25,\ 0.75,\ 1.$ are also displayed, but we do not exhibit them here. By invoking the global variable $SolH3$, the numerical solution at a given grid point (with $t = 1$) can be displayed. For example, $SolHL3(.50, .25, .25) = 0.005748$ and the exact solution value is $u(.50, .25, .25, 1) = 0.005748$. The global variable $SH3(t)$ exhibits the numerical solution for various values of t. Thus, $SH3(1)$ yields a list grid values at $t = 1$. One of the values in this list is $u_{1,2,2}(t) = 0.005748$, and this corresponds to the value of the exact solution $u(.50, .25, .25, 1) = 0.005748$.

If the twenty-third argument in $NumHeatLines3d$ is omitted, Maple's output is a procedure. Then the numerical solution can be displayed for various values of t.

Method of Lines for the Three-Dimensional Wave Equation

We consider initial and boundary value problems for the three-dimensional nonhomogeneous wave equation

$$u_{tt}(x,y,z,t) - c^2\nabla^2 u(x,y,z,t) = F(x,y,z,t). \tag{11.9.26}$$

The initial conditions are $u(x,y,z,t_0) = f_1(x,y,z)$, $u_t(x,y,z,t_0) = f_2(x,y,z)$. The spatial region (for any fixed value of t) is bounded by the planes $x = a_1$, $x = b_1$, $y = a_2$, $y = b_2$, $z = a_3$, $z = b_3$ and boundary conditions of all three kinds can be assigned on each of the bounding planes.

To solve a problem numerically, we replace second order spatial derivatives in (11.9.26) by centered difference quotients with increments h_x, h_y and h_z, with t left intact. The grid points are $(x_i, y_j, z_l) = (\alpha + ih_x, \beta + jh_y, \gamma + lh_z)$ with $i = 0, \pm1, \pm2, \ldots, j = 0, \pm1, \pm2, \ldots, l = 0, \pm1, \pm2, \ldots$ (where α, β, γ are arbitrary constants). $u_{i,j,l}(t)$ is the (numerical) solution of an initial and boundary value problem for (11.9.26) at (x_i, y_j, z_l, t). The system of ODEs

$$\frac{d^2u_{i,j,l}(t)}{dt^2} = F_{i,j,l}(t) + \frac{c^2}{h_x^2}\left(u_{i+1,j,l}(t) - 2u_{i,j,l}(t) + u_{i-1,j,l}(t)\right) \tag{11.9.27}$$

$$+\frac{c^2}{h_y^2}\left(u_{i,j+1,l}(t) - 2u_{i,j,l}(t) + u_{i,j-1,l}(t)\right) + \frac{c^2}{h_z^2}\left(u_{i,j,l+1}(t) - 2u_{i,j,l}(t) + u_{i,j,l-1}(t)\right),$$

where $F_{i,j,l}(t) = F(x_i, y_j, z_l, t)$, represents the *method of lines* for the numerical solution of (11.9.26).

$NumWaveLines3d(c, F(x,y,z,t), t = t_0, f_1(x,y,z), f_2(x,y,z), x = a_1..b_1, lxf, g_1(y,z,t), rxf, g_2(y,z,t), y = a_2..b_2, lyf, r_1(x,z,t), ryf, r_2(x,z,t), z = a_3..b_3, lzf, s_1(x,y,t), rzf, s_2(x,y,t), n_x, n_y, n_z)$ applies the method of lines to find numerical solutions of the foregoing initial and boundary value problems. The arguments are as in $NumWaveLines2d$, except for additional arguments relating to the z variable. (Arguments relating to boundary conditions are as in $NumHeatLines3d$.) If an optional twenty-fourth argument is omitted, the output is a procedure that permits the evaluation of the solution at arbitrary values of t. If the twenty-fourth argument is $t = t_m$, the output is a list of arrays of grid values at the time $t = t_m$. Maple's default numerical solution method is used to solve the system of ODEs. More optional arguments can be entered to cause Maple to use other numerical methods to solve the system.

Example 11.46. A Neumann Problem for the Wave Equation. The function $u(x,y,z,t) = x^2y^2z^2t^2$ is a solution of the wave equation (11.9.26) with $c^2 = 1$ and $F(x,y,z,t) = 2(x^2y^2z^2 - y^2z^2t^2 - x^2z^2t^2 - y^2x^2t^2)$. We construct an

initial and boundary value problem for (11.9.26) in the unit cube $0 < x, y, z < 1$ with the initial conditions $u(x, y, z, 0) = 0$ and $u_t(x, y, z, 0) = 0$. Neumann conditions are determined from derivatives of $x^2y^2z^2t^2$ on the boundary. The x, y, z intervals are divided into four equal parts. We use $NumWaveLines3d$ to find the numerical solution of this problem at $t = 1$, with arguments given in terms of $x^2y^2z^2t^2$. For example, we put $f_1(x, y, z) = f_2(x, y, z) = 0$, $a_1 = 0$, $b_1 = 1$, $rxf = 0$, $g_2(y, z, t) = 2y^2z^2t^2$. Adding a twenty-fourth argument $t = 1$ generates five arrays that exhibit the numerical solution values at the time $t = 1$ and at the grid points $(x_i, y_j, 0.5)$ for $z = 0, 0.25, 0.5, 0.75, 1$. They are not exhibited here. Invoking the global variable $SolWL3$ the numerical solution at a given grid point (with $t = 1$) is displayed. For example, $SolWL3(.50, .25, .25) = 0.000977$, in agreement with the exact solution value. The global variable $SW3(t)$ exhibits the numerical solution for various values of t. Thus, $SW3(1)$ yields a list grid values at $t = 1$. If the twenty-fourth argument is omitted, the output is a procedure that can find the numerical solution for various values of t.

Difference Methods for the Three-Dimensional Laplace and Poisson Equations

Poisson's equation in three dimensions is given as

$$\nabla^2 u(x, y, z) = F(x, y, z). \tag{11.9.28}$$

We consider the numerical solution of boundary value problems for (11.9.28) in rectangular regions in the (x, y, z)-plane. They are defined as in our discussion of the three-dimensional heat and wave equations. The partial derivatives in (11.9.28) are replaced by centered difference quotients with increments h_x, h_y, and h_z. The *truncation error* for the difference approximation is $O(h_x^2 + h_y^2 + h_z^2)$, so that it is *consistent* with Poisson's equation.

The grid points are given (in general) as $(x_i, y_j, z_k) = (\alpha + ih_x, \beta + jh_y, \gamma + kh_z)$ with $i = 0, \pm 1, \pm 2, \ldots, j = 0, \pm 1, \pm 2, \ldots, k = 0, \pm 1, \pm 2, \ldots$, (where α, β, γ are arbitrary constants). Then, $u_{i,j,k}$ represents the (numerical) solution of a boundary value problem for (11.9.28) at (x_i, y_j, z_k) for an appropriate range of subscript values. The difference scheme can be given as

$$u_{i,j,k} = \frac{u_{i-1,j,k} + u_{i+1,j,k} + (h_x^2/h_y^2)(u_{i,j-1,k} + u_{i,j+1,k})}{2 + 2h_x^2/h_y^2 + 2h_x^2/h_z^2}$$

$$+ \frac{(h_x^2/h_z^2)(u_{i,j,k-1} + u_{i,j,k+1}) - hx^2 F_{i,j,k}}{2 + 2h_x^2/h_y^2 + 2h_x^2/h_z^2}. \tag{11.9.29}$$

The boundary values of all three kinds are assigned as for the heat and wave equations. This determines the grid points (x_i, y_j, z_k) at which $u_{i,j,k}$ is to be found. The difference equations yield a system of simultaneous linear equations for the $u_{i,j,k}$.

Since a large number of equations must be solved in general, (11.9.29) has been given in a form suitable for solution by iteration.

To solve the problem numerically, we assume, as for two dimensions, that a set of values $u_{i,j,k}$ is prescribed on the relevant grid. This is the initial guess. Then, we let i, j, k range through all the grid values. We find $u_{1,1,1}, u_{1,1,2}, \ldots, u_{1,1,k_m}, u_{1,2,1}, u_{1,2,2}, \ldots, u_{i_m,j_m,k_m}$, where i_m, j_m, k_m represent the extreme values of i, j, k in the grid. In the *Gauss-Seidel method,* once a new value of $u_{i,j,k}$ is determined, it is used immediately in all ensuing calculations. In the *SOR* method, with relaxation parameter ω, once a new value of the numerical solution at i, j, k is determined—we call it $\hat{u}_{i,j,k}$— we replace it by $u_{i,j,k} = \omega\, \hat{u}_{i,j,k} + (1-\omega)\, u_{i,j,k}$ as the (new) grid point value, where the coefficient of $1-\omega$ is the old value at i, j, k. Then $u_{i,j,k}$ is used in all further calculations and entered on the right side of (11.9.29). In the *Jacobi method,* solution values on the right side of (11.9.29) are not updated with newly determined values until the next step in the iteration process begins.

NumLaplace3d$(F(x,y,z), x = a_1..b_1, lxf, g_1(y,z), rxf, g_2(y,z), y = a_2..b_2, lyf, r_1(x,z), ryf, r_2(x,z), z = a_3..b_3, lzf, s_1(x,y), rzf, s_2(x,y), n_x, n_y, n_z, numits, inguess, err, par)$ solves the foregoing problems for Poisson's equation numerically. The arguments correspond to those of $NumLaplace$ in Section 11.4, with the boundary conditions given as in the method of lines for the three-dimensional heat equation, say. The last argument, par, must satisfy $0 < par < 2$. If $par = 1$, the Gauss-Seidel method is used. Otherwise, the SOR method is used. The Jacobi method is not carried out.

As an example we consider a problem for Laplace's equation in the unit cube $0 < x, y, z < 1$ with the Dirichlet conditions $u(0,y,z) = 1$ and $u(x,y,z) = 0$ on the rest of the boundary. We put $n_x = n_y = n_z = 2$, so that $h_x = h_y = h_z = 1/2$. Then, $2 + 2h_x^2/h_y^2 + 2h_x^2/h_z^2 = 6$ and (11.9.29) characterizes the steady random walk of a particle with probability $1/6$ of taking a step from the single interior point $(.5, .5, .5)$ in any of six directions. After one step the boundary is reached. The problem asks for the probability that the particle reaches the boundary point on $x = 0$ before it reaches any of the other five boundary points where it is absorbed. The output of the procedure $NumLaplace3d$ with the appropriate arguments determines that $u_{1,1,1} = .1666667$ no matter what iteration method is used. This approximates the exact value of the probability $u(.5, .5, .5) = 1/6$.

Exercises 11.9

11.9.1. Show that the truncation error for (11.9.2) is $O(k + h_x^2 + h_y^2)$.

11.9.2. Verify the von Neumann stability condition (11.9.5).

11.9.3. Use the procedure $NumHeatForw2d$ to verify the results of Example 11.41. Consider the stable and unstable cases.

11.9.4. Verify the von Neumann stability condition (11.9.7).

11.9.5. Apply the procedure $NumHeatBackw2d$ to the Neumann problem of Example 11.41. Invoke the implicit backward scheme and the Crank-Nicolson scheme and obtain the steady-state solution.

11.9.6. Use $NumHeatPRADI2d$ and $NumHeatDRADI2d$ to obtain the results (11.9.13) and (11.9.14) obtained in Example 11.42.

11.9.7. Reproduce the results of Example 11.43 using $NumHeatLines2d$.

11.9.8. Verify the von Neumann stability condition (11.9.19).

11.9.9. Use the procedure $NumWaveForw2d$ to verify the results of Example 11.44. Consider the stable and unstable cases.

11.9.10. Apply $NumWaveBackw2d$ to the problem of Example 11.44 with $k = 0.25$ and obtain the results given by $SolWB$.

11.9.11. Apply $NumWaveLines2d$ to the problem of Example 11.44 and obtain the results given by $SolWL$.

11.9.12. Use $NumHeatLines3d$ to reproduce the results of Example 11.45.

11.9.13. Apply $NumWaveLines3d$ to reproduce the results of Example 11.46.

11.9.14. Use the procedure $NumLaplace3d$ to solve Laplace's equation in the unit cube $0 < x, y, z < 1$ with the Dirichlet boundary conditions $u(0, y, z) = 1$ and $u(x, y, z) = 0$ on the rest of the boundary. Put $n_x = n_y = n_z = 4$ in the procedure and use the Gauss-Seidel method and the SOR method (with two choices for the relaxation parameter ω) to solve the problem.

11.9.15. Reconsider the problem of Exercise 11.9.14 with the Dirichlet conditions $u(x, y, z) = 0$ replaced by Neumann conditions $\partial u(x, y, z)/\partial n = 0$.

11.10 MAPLE FINITE DIFFERENCE METHODS FOR PARABOLIC AND HYPERBOLIC PDEs

The basic built-in Maple procedure for the solution of PDEs is *pdsolve*. As indicated previously, the procedure finds exact solutions of the PDEs or shows how to obtain exact solutions by separation of variables or other means, if that is possible. Whereas Maple's *dsolve* procedure treats initial and boundary value problems for ODEs, *pdsolve* does not construct analytical solutions of initial or initial and boundary value problems for PDEs. However, in the case of time-dependent problems with one space variable, so that the PDEs contain two independent variables, Maple does obtain numerical solutions of initial and boundary value problems. If the PDE and appropriate initial and boundary conditions are entered in *pdsolve*, together with the option *type=numeric* or simply *numeric*, Maple employs finite difference methods for the solution of these problems. The basic PDE treated by the procedure contains only one time derivative. Problems with a higher-order time derivative are reduced (by Maple) to equivalent first order systems of PDEs. Maple's *default* solution method uses a second order (in space and time) centered, implicit finite difference scheme. Boundary conditions of all three kinds can be handled.

The procedure takes the form $pdsolve(PDE, ICBCs, numeric, options)$. The first and second arguments give the PDE and a set of initial and boundary conditions for the PDE. The third argument tells Maple to solve the problem numerically. In the default, Maple divides the (finite) x interval into 20 equal parts, and the time increment

k equals the spatial step size h. Optional arguments can be added to control the values of h and k. If the PDE is of first order in time, the finite difference method used to solve the problem can also be specified. If the *method* option is omitted, Maple uses its default difference scheme to solve the problem. Only problems given over a finite x interval can be solved. For problems with x derivatives of order two or higher, the interval is determined by the location of the boundary conditions. If the PDE is of first order in x (so that only one boundary condition is prescribed), an argument that gives the interval over which the solution is to be found must be added. The boundary condition for the problem is specified at one endpoint of this interval and an argument that prescribes a *numerical* boundary condition at the other endpoint of the interval must be provided.

The output is a (Maple) *module* that permits the solution to be evaluated at all points within the specified interval, or to be plotted and animated. If a single PDE in two independent variables, first order in time is considered, one of 11 specific numerical methods, such as the Lax-Friedrichs or Crank-Nicolson method, can be prescribed for the numerical solution of the problem. A full description of the procedure and its implementation is given in the Maple help facility. Plots of numerical solutions and animations of plots can be generated by the module.

Example 11.47. A Mixed Problem for the Heat Equation. We reconsider the problem presented in Examples 11.3 and 11.5. It is given as $u_t(x,t) - u_{xx}(x,t) = 2tx^2 - 2t^2$, $0 < x < 1$, $0 < t$, with the initial and mixed boundary conditions $u(x,0) = 0$, $0 < x < 1$, $u_x(0,t) = 0$, $u_x(1,t) + u(1,t) = 3t^2$, $0 < t$. The exact solution of the problem is $u(x,t) = x^2t^2$. We obtain the numerical solution of the problem using *pdsolve* with the *numeric* option. For each difference scheme used, the x interval is divided into four equal parts, so that $h = 0.25$ and we put $k = 0.001$. The solution values are found at $t = 1$.

We invoke $pdsolve(PDE, ICBCs, numeric, spacestep = 0.25, timestep = 0.001)$, where PDE and ICBCs represent the given PDE and auxiliary conditions, respectively. Maple uses its *default implicit difference scheme* to solve the problem, with the output **module**(), **export** $plot, plot3d, animate, value, settings$; **end**. The five listed quantities exported by the *module* permit the solution $u(x,t)$ to be plotted or evaluated at various values of x and times t. If we assign a name *smod* to the module, the Maple command $xval := smod : -value(t = 1)$ permits the evaluation of the numerical solution for all x within $0 \le x \le 1$ with $t = 1$. (Maple uses interpolation to evaluate the solution at points x that are not in the grid.) For example, $xval(.5)$ yields $[x = 0.5, t = 1., u(x,t) = .2500]$.

The addition of the argument $method = Euler$ in the procedure causes Maple to use the *explicit forward difference scheme* presented in Section 10.2 to solve the problem numerically using the specified values of h and k. If we now define *smod* and $xval$ as above, we obtain for $xval(.5) = [x = 0.5, t = 1., u(x,t) = .2498]$. If, instead, we enter $method = CrankNicholson$, we obtain, on proceeding as before, $xval(.5) = [x = 0.5, t = 1., u(x,t) = .2500]$. Maple uses the *implicit Crank-Nicolson method* to solve the problem. (Maple uses a variant spelling of Nicolson.) The introduction of $method = BackwardEuler$ causes the *implicit backward dif-*

ference scheme to be used and gives $xval(.5) = [x = 0.5, t = 1., u(x,t) = .2502]$. The use of the *DuFort-Frankel scheme*, which is invoked by entering $method = DuFortFrankel$, yields $xval(.5) = [x = 0.5, t = 1., u(x,t) = .2500]$. These results differ from the values obtained previously, because of differences in evaluation methods.

Finally, we consider an initial and boundary value problem for a linear first order PDE. To solve the problem numerically using *pdsolve*, an x interval must be specified. An initial condition is prescribed and a boundary condition is given at one endpoint of the interval. Since the time derivative is of first order, the *pdsolve* numerical solution can be obtained using a number of the difference schemes introduced above. Depending on the difference scheme used to solve the problem, it may be necessary to include an additional boundary condition at the other endpoint of the x interval to determine the numerical solution, as indicated in Section 11.7. Exact solution values at the other endpoint are unavailable, and these values cannot be prescribed arbitrarily, as this would (generally) give rise to an ill-posed problem for the PDE.

We assume that the grid points in the x interval are $x_1, \ldots, x_n$ and that the boundary condition is assigned at the left endpoint x_1. (A similar approach can be used if the reverse is true.) An indirect method is used to assign a solution value at the right endpoint x_n. The simplest approach is to equate the numerical solution value at x_n to that at x_{n-1}. Other methods relate the solution value at x_n to that at x_{n-1}, and possibly other values at an earlier time, by replacing derivatives by difference quotients in the PDE at x_n. The effect of each of these methods is to eliminate points x_i that lie outside the x grid and thereby generate a solvable finite number of linear equations for the numerical solution values at $x_1, \ldots, x_n$. The additional condition placed on the endpoint value is not a part of the original formulation of the initial and boundary value problem, so it is called a *numerical boundary condition.*

Example 11.48. An Initial and Boundary Value Problem for a First Order PDE. We examine the problem considered in Example 11.26. The first order PDE to be solved is $u_t(x,t) + 4\,u_x(x,t) = -u(x,t) + 8xt - 8x + 4 + x^2t - 4t$, $0 < x < 1$, and the initial and boundary conditions are $u(x,0) = -x^2 + \exp(-x/4) + 8$, $u(0,t) = 9 - 4t$. The exact solution of the problem is $u(x,t) = (t-1)x^2 + \exp(-x/4) + 8 - 4t$. We put $h = 1/5$ and $k = 1/30$, and the numerical solution is to be found at $t = 1$.

First we apply the procedure $pdsolve(PDE, IBC, numeric, time = t, range = 0..1, spacestep = 1/5, timestep = 1/30)$, where PDE and IBC refer to the given PDE and the above initial and boundary values, respectively. Maple uses its default method to solve the problem numerically. As before, the output is a module. If we assign the name *pds* to the module, the Maple command $xval := pds : -value(t = 1)$ permits the evaluation of the numerical solution at all points x with $0 \le x \le 1$ and with $t = 1$. Then, if we enter $[seq(rhs(xval(i/5)[3]), i = 0..5)]$, the output is $[5.000, 4.951, 4.904, 4.859, 4.817, 4.777]$, for the numerical solution values at $x = 0, 0.2, 0.4, 0.6, 0.8, 1$, respectively, with four decimals retained. The exact solution values at these points are $[5.000, 4.951, 4.905, 4.861, 4.819, 4.779]$.

The optional argument $method = ForwardTime1Space[backward]$ in the procedure causes Maple to use the explicit *forward-backward difference scheme* (as defined in Example 11.26) to solve the problem. Proceeding as above yields the Maple output $[5.000, 4.951, 4.902, 4.854, 4.806, 4.760]$, which agrees with the result in Example 11.26. With the optional argument $method = BackwardTime1Space[backward]$, Maple uses the implicit *backward-backward difference scheme* to solve the problem. Again proceeding as above yields the same Maple output as before. For each case, if the x derivative is approximated by a forward difference, the procedure fails.

The *Lax-Friedrichs method* can be applied to this problem only if a numerical boundary condition is given at $x = 1$. The optional arguments *method* = *LaxFriedrichs, numericalbcs* = $u[1, n] - u[1, n-1]$, causes Maple to solve the problem using the Lax-Friedrichs method with the added numerical boundary condition that u at $x = 1$ has the same value as at $x = 0.8$. Proceeding as before yields the Maple output $[5.000, 4.950, 4.900, 4.853, 4.816, 4.816]$. The solution values at $x = 0.8$ and $x = 1$ are equal because of the assigned numerical boundary condition. With the exception of the value at $x = 1$, there is good agreement with the foregoing results.

The numerical boundary condition *numericalbcs* $= [BackwardTime1Space[forward], n]$ uses the *backward-forward method* to assign a value to u at $x = 1$. Now we obtain as the numerical solution values $[5., 4.95, 4.899, 4.849, 4.8, 4.75]$, which is an improvement over the foregoing. Then, the use of *numericalbcs* $= [Backward Time1Space[backward], n]$ which uses the *backward-backward method*, yields essentially the same result as before.

The *Lax-Wendroff method* also cannot be applied directly to solve this problem without introducing a numerical boundary condition at $x = 1$. The optional arguments *method=LaxWendroff, numericalbcs* = $u[1, n] - u[1, n-1]$ cause Maple to solve the problem using the Lax-Wendroff method with the added numerical boundary condition that u at $x = 1$ has the same value as at $x = 0.8$. Proceeding as before yields $[5.000, 4.951, 4.904, 4.863, 4.811, 4.811]$. With the exception of the solution value at $x = 1$, there is an improved agreement with the foregoing results. The use of the numerical boundary conditions that correspond to the *backward-forward* and *backward-backward methods* yields essentially the same results as for the exact solution.

Exercises 11.10

11.10.1. Invoke *help* for *pdsolve* with the *numeric* option to determine precisely what PDEs can be solved numerically and to find several examples of how this numerical procedure can be applied.

11.10.2. Reproduce the results of Example 11.47.

11.10.3. Reproduce the results of Example 11.48.

CHAPTER 12

FINITE ELEMENT METHODS IN TWO DIMENSIONS

12.1 INTRODUCTION

It was shown in Chapters 6 and 8 that the *Rayleigh-Ritz* and *Galerkin methods* can be used to construct approximate solutions to boundary value problems for Laplace's and Poisson's equations, in particular, and linear elliptic PDEs of second order in general. The Rayleigh-Ritz method applies to formally self-adjoint PDEs, while the Galerkin method can deal with more general PDEs. In each method, the PDE and the boundary conditions are replaced by integral relations with the property that they are equivalent to the given PDE with its boundary condition(s) if the solutions are smooth, (at least) twice continuously differentiable functions. The integral relations contain derivatives of the solution of at most first order. As a result, a meaning can be attached to solutions of the given boundary value problem that are at most piecewise continuously differentiable. In the terminology introduced previously, they are referred to as *weak solutions* of the problem.

Because of its greater generality, we will deal exclusively with the Galerkin integral representation and method. Our discussion is restricted to problems in two spatial dimensions. The Galerkin integral representation for boundary value problems for second order elliptic PDEs is derived below, as are corresponding Galerkin integral

Partial Differential Equations of Appplied Mathematics, Third Edition. By Erich Zauderer

representations for initial and boundary value problems for second order hyperbolic and parabolic PDEs. The possibility of constructing weak solutions of each of the problems will be exploited and approximate solutions will be constructed using the *Galerkin method.* This method was introduced in Section 8.2 for Laplace's equation and is formulated below for each of the aforementioned problems.

The *Galerkin method* constructs an approximate solution of the given problem as an expansion of a finite number of *basis functions* with coefficients that are to be determined. Each basis function must satisfy an admissibility condition appropriate for the problem. The basis functions can be chosen to be weak solutions of the Galerkin integral representation associated with the given problem. In the *finite element method* as we formulate it, the basis functions are chosen to be piecewise linear functions related to an approximate decomposition of the given region into a collection of triangles, or a *triangulation of the region.* Unless the region has a polygonal boundary, the mesh formed by the union of the triangles will only approximate the region. These triangles are called *finite elements.* Each of the basis functions is chosen to be piecewise continuously differentiable. As will be seen, the number of basis functions corresponds to the number of vertices in the triangulation. (The vertices are often called *nodes.*) We will show how to triangulate a region, determine the corresponding basis functions, and use them to construct the Galerkin approximation to the solution of problems for PDEs of all three types. We have created a number of Maple procedures specifically for this purpose.

Alternative decompositions of the region (such as into rectangles) can be made with corresponding basis functions. The finite element method can also be extended to PDEs of higher order and with more than two spatial dimensions. Various modifications of the method have also been introduced. *Maple* does not have any built-in procedures related directly to the use of the finite element method. *Matlab* has a PDE Toolbox that includes the finite element method in two spatial variables. It carries out triangulations of various regions and uses the Galerkin method to approximate solutions. Many library packages are devoted to the finite element method.

12.2 THE TRIANGULATION OF A REGION

Let the bounded two-dimensional region in the (x, y)-plane over which the problem for a PDE of elliptic, hyperbolic, or parabolic type is given be denoted by G and its boundary by S. The first step in carrying out the *finite element method* for the solution of the problem is to triangulate the region G. This is done by subdividing the region G into a union of triangles. It must be carried out so that no triangle in the subdivision has a vertex on the side or edge of another triangle, unless it is also a vertex of the other triangle. The subdivision contains interior vertices (that lie in G) and boundary vertices (that lie on S). If S or a portion thereof is curved, it is replaced by a polygonal (piecewise linear) approximation determined by the placement of vertices on S. The full boundary S is approximated by a union of line segments that connect the chosen boundary vertices. (Some or a portion of these segments may be exterior to G.) Thus, the boundary S (if it is not already a polygon) is replaced by a polygon which we

denote by $\hat{S}$. The region G is thereby replaced by the interior of $\hat{S}$ (if S is not a polygon), and we denote this approximation of the region by $\hat{G}$. It is the polygonal region $\hat{G}$ that is triangulated.

Each triangle in the subdivision is identified or specified by a list of its three vertices. The triangulation of a region is generated by constructing a list of triads (groups of three) of vertex points of the triangles. Some of the vertices lie in $\hat{G}$, and some are on the boundary $\hat{S}$. The vertices on the boundary are just the vertices of the polygon $\hat{S}$. The union of all the triangles equals $\hat{G}$.

For any region G with boundary S, the triangulation can be carried out directly by prescribing a list of boundary and interior vertices and determining a list of triads for the triangles in the subdivision. This is generally a difficult task if a large-scale triangulation is needed, as is usually the case. Consequently, we have constructed a number of Maple procedures that automate the triangulation of regions based on the specification of interior and boundary vertices for these regions. Additionally, a number of procedures that generate useful and important information relating to the subdivision have also been constructed. This includes procedures that deal with the triangulation of rectangular, triangular, circular, or elliptical regions or portions thereof. They are not presented here, but we consider, instead, a procedure that triangulates a large class of polygons. (As indicated, to triangulate the interior of any non-polygonal region G, its boundary S must first be approximated by a polygon $\hat{S}$ and G replaced by the resulting polygonal region $\hat{G}$.)

Triangulation of a Polygon and Its Refinement

The procedure $PolygonTriang$ constructs a simple triangulation of convex polygons and various arbitrarily shaped polygons. In the default, it finds the centroid of the polygon and constructs a triangulation of the polygonal region by connecting all the vertices of the polygon to the centroid. (The x and y coordinates of the centroid are the averages of the x and y coordinates of the vertices of the polygon.) The method works if the centroid lies in the interior or is a boundary vertex point of the polygonal region, and all the triangles lie in $\hat{G}$. The procedure is given as $PolygonTriang([BP_1, BP_2, \ldots, BP_n])$, where BP_i, $i = 1, \ldots, n$ represents a vertex $[x_i, y_i]$ on the polygonal boundary $\hat{S}$. (All vertices of the polygon must be listed.) If a second argument $[x_0, y_0]$ (with $[x_0, y_0] \in \hat{G}$) is added to the procedure, all the vertices of the polygon are connected to the point $[x_0, y_0]$ rather than the centroid. The number of triangles in the subdivision equals the number of vertices of the polygon. The output of the procedure is a list of triads of the vertices of the triangles in the subdivision. The global variables $VertL$ and $BVertL$ reproduce the list of triads and boundary points, respectively. The procedure $Vertexlist(VertL, BVertL)$ determines and lists the number of interior and boundary vertices, with the list of interior vertices preceding that of the boundary vertices.

Once an initial triangulation of the region $\hat{G}$ is constructed, the procedure $Refine\ Triang$ refines the triangulation by replacing each triangle by four subtriangles whose vertices comprise the vertices and the midpoints of each of the edges of the original triangle. Thus, the procedure quadruples the number of triangles in the original

triangulation. It takes the form *RefineTriang(VertTriadList,BVertlist,NumIts)*, where the first and second arguments are a list of triads and boundary vertices, respectively, for the original triangulation. The refinement process can be iterated (as part of the procedure), and the third argument of the procedure must be a positive integer that prescribes the number of times the triangulation is refined. The output is a list of triads of the vertices of the triangles in the refined subdivision.

We begin with an example that illustrates the uses of the foregoing procedures.

Example 12.1. The Triangulation of a Circle. We consider a circle with center at $(0, 0)$ and radius 1 and construct a triangulation of its interior. First, the circle is divided into four equal parts with the boundary vertices chosen as $BPList = [[1, 0], [0, 1], [-1, 0], [0, -1]]$. As a result, the circle is approximated by a diamond-shaped quadrilateral. This represents the polygon $\hat{S}$, and it is its interior $\hat{G}$ that is triangulated.

PolygonTriang$(BPList)$ gives the initial triangulation as [[1, 0], [0, 1], [0, 0], [0, 1], [-1, 0], [0, 0], [-1, 0], [0, -1], [0, 0], [0, -1], [1, 0], [0, 0]], which we denote as $VertL$. The 12 elements in the list represent a list of four triads of vertices of the four triangles in the subdivision. This list determines the triangulation. The four triangles are constructed by connecting the vertices of the quadrilateral with the centroid of the quadrilateral, which is the origin $(0, 0)$. [If a second argument $[x_0, y_0]$ is added in the procedure, the triangles are constructed by connecting the vertices of the quadrilateral with the point (x_0, y_0).] $Vertexlist(VertL, BPList)$ yields $InteriorVertices = 1, BoundaryVertices = 4$, $[[0, 0], [1, 0], [0, 1], [-1, 0], [0, -1]]$. The origin is the only interior vertex, and there are four boundary vertices. These are all displayed in the output.

Next, we refine the triangulation by increasing the number of triangles in the subdivision and invoke the procedure $RefineTriang(VertL, BPList, 2)$. The third argument of the procedure indicates that we are carrying out a double refinement. The output, which we do not display, is a list of 64 triads that specify the vertices of the 64 triangles in the refined subdivision. The global variables $MListMod$ and $BPListMod$ exhibit the (new) list of triads and boundary points, respectively. Then, $Vertexlist(MListMod, BPListMod)$ states that there are 25 interior vertices and 16 boundary vertices in the refined triangulation and lists these vertices. This approach has increased the number of triangles, but it is unsatisfactory in this case, because it retains the quadrilateral as an approximation to the circle, and this is a poor approximation. (*RefineTriang* increases the number of boundary points, but they lie on the quadrilateral and not on the circle, as would be desirable.)

To improve the approximation to the unit disk by triangulation, we must increase the number of vertices on the circle in the initial use of $PolygonTriang$. To do so, we prescribe a list of n (equally spaced) vertices on the circle $[\cos(2\pi i/n), \sin(2\pi i/n)]$, $i = 0, \ldots, n-1$, and enter it as the argument of $PolygonTriang$. Then, if we put $n = 16$, for example, the initial triangulation obtained from $PolygonTriang(BPList)$, where $BPList$ is the foregoing list of boundary vertices, yields a list of 16 triads of vertices (which we do not display) of the 16 triangles in the subdivision. Next, we refine the triangulation by increasing the number of triangles in the subdivision and

invoke the procedure *RefineTriang*($VertL$, $BPList$, 2) as before. The output is a list of 256 triads, denoted as $MListMod$, that specifies the vertices of the 256 triangles in the subdivision. The procedure $Vertexlist(MListMod, BPListMod)$ states that there are 97 interior vertices and 64 boundary vertices in the refined triangulation and lists these vertices.

Clearly it is possible to generate an extremely accurate triangulation of the disk using the procedures *PolyTriang* and *RefineTriang*.

Plots of Triangulations

If the boundary S of the region G is not polygonal, it must first be approximated by a polygonal boundary curve $\hat{S}$ before a triangulation can be introduced. The curves S and $\hat{S}$ can generally be plotted using built-in Maple procedures.

As an example, we consider the region G bounded by the parabolas $y = 1 - x^2$ and $y = x^2 - 1$. The points BV =[[1, 0],[1/2,3/4], [0,1],[-1/2,3/4],[-1,0], [-1/2,-3/4],[0,-1],[1/2,-3/4]] lie on the boundary S and determine the vertices of an approximating polygon $\hat{S}$. The boundary S and the polygon $\hat{S}$ can be plotted using the Maple procedures $plot([1 - x^2, x^2 - 1], x = -1..1)$ and $plot(BVM)$, respectively, where BVM represents the list BV supplemented by the point $[1, 0]$. If we assign the names $P1$ and $P2$ to the plots of S and $\hat{S}$, the procedure $display(\{P1, P2\})$ from the *plots* package exhibits both curves in a single display, as shown in Figure 12.1.

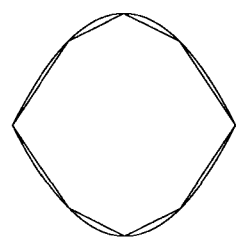

Figure 12.1 The boundaries S and $\hat{S}$.

The interior of the polygon $\hat{S}$ represents the region $\hat{G}$ that is to be triangulated. If a triangulation of $\hat{G}$ is constructed by any means and is represented by a list of triads $VList$, then $TriangPlot(VList)$ displays a plot of the triangulation. It uses the built-in Maple procedure *polygonplot* to plot the collection of polygons determined by the triads of vertices. Additionally, suppose that $BVList$ represents an ordered list of the vertices of the polygon $\hat{S}$, and the list of vertices given as the output of $Vertexlist(VList, BVList)$ is denoted by $VVList$. Then $TriangPlot(VList, Triangles, VVList)$ displays a plot of the triangulation in which the interior and boundary vertices are numbered and the n triangles $T1, T2, \ldots, Tn$ in the triangulation are identified. The numbering of the vertices and the naming of the triangles is carried out using the Maple procedure *textplot*.

To continue with our example, we construct an initial triangulation of $\hat{G}$ by using $PolygonTriang(BV)$ and denote the resulting triad list by $VList$. (All the vertices of the polygon are connected with the origin, which is the centroid of the polygon.) The procedure $TriangPlot(VList)$ generates a plot of the triangulation which is exhibited in Figure 12.2.

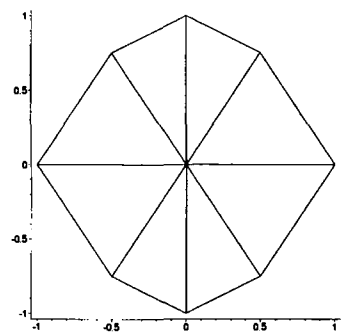

Figure 12.2 The triangulation of $\hat{G}$.

We refine the triangulation once by using *RefineTriang(VList,BV,1)* and denote the new triad list by *VListM* and the modified boundary vertex list by *BVM*. Then, *Vertexlist(VListM,BVM)* finds that there are nine interior vertices and 16 boundary vertices in the refined triangulation. With *VVListM* as the new vertex list, the procedure *TriangPlot(VListM,Triangles,VVListM)* displays the refined triangulation as shown in Figure 12.3.

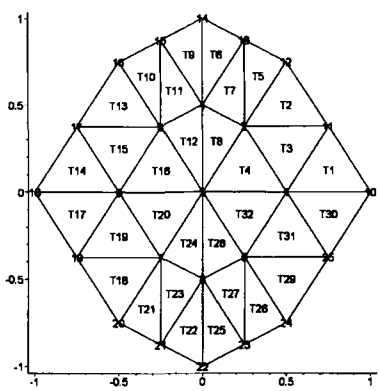

Figure 12.3 The refined triangulation of $\hat{G}$.

Maximum and Minimum Areas of the Triangles in a Triangulation

When constructing a triangulation of a region using the foregoing procedures, it is generally the case that the triangles in the subdivision are not (roughly) uniform in size. Some triangles may have much larger dimensions and areas than others and this can lead to a loss of precision for the numerical solution in the region covered by these triangles. Additionally, the given problem for the PDE may require that an especially fine subdivision be created only for certain parts of the given region G but not for remaining part of G. (It is true that in all cases, the finer the subdivision, the better the approximation, but it is not necessary to go to extremes throughout G.)

$TriangArea(TriadList)$ determines the variation in the dimensions of the triangles in a subdivision. The argument $TriadList$ is a list of the triads of vertices that comprise the triangulation. The output gives the number of triangles in the triangulation, as well as the mean and standard deviation of the areas of the triangles. Additionally, the minimum and maximum areas are displayed and the triangles that have these areas are identified. The triangles in the triad list are numbered consecutively in groups of three and denoted by $T1, T2, \ldots, Tn$, where n is the number of triangles in the triangulation. There are five global variables *AList, AListSort, AWList, TMaxList, TMinList* that display information about the areas of the triangles. Some of them are invoked in the following example.

Example 12.2. The Triangulation of a Trapezoid. The list $BTrapList = [[0,0],[1,0],[1,1],[0,2]]$ represents the vertices of a trapezoid, whose centroid is $(1/2, 3/4)$. We triangulate the trapezoid using $PolygonTriang(BTrapList)$. Its output, which we denote by *TrapList*, gives the four triangles $T1, T2, T3, T4$ (identified by their vertices) obtained by connecting the four vertices to the centroid. The triangle $T1$, for example, has the vertices $[0,0]$, $[1,0]$, $[1/2, 3/4]$. $Vertexlist$ $(TrapList, BTrapList)$, whose output we denote as $VTrapList$, lists the single interior vertex and the four boundary vertices in the triangulation. $TriangPlot$ $(TrapList, Triangles, VTrapList)$ plots the triangulation. The addition of the final two arguments in the procedure causes Maple to number the vertices and to identify the triangles in the triangulation, as shown in Figure 12.4.

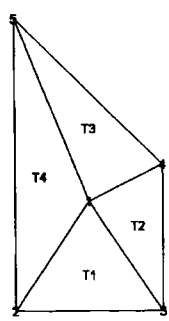

Figure 12.4 The triangulation of the trapezoid.

Next we invoke the procedure $TriangArea(TrapList)$. The output states that there are four triangles in the triangulation, with mean area 0.375 and standard deviation 0.08839. The triangle $T2$ has the minimum area 0.25, and $T4$ has the maximum area 0.5. The output of the global variable $AList$ is $[0.375, 0.250, 0.375, 0.5]$, which represents a list of the areas of $T1$, $T2$, $T3$, $T4$, respectively. It shows that $T1$ and $T3$ each have the area 0.375. The global variable $AListSort$ lists the areas in increasing order, while *TMinList* and *TMaxList* gives the numbers of the triangles that have the minimum and maximum areas, respectively.

Location of a Point in a Triangulated Region

The areas of the triangles in the triangulation of Example 12.2 are unequal and it may be of interest to construct a further triangulation of the triangle $T4$, whose area is maximal, while leaving the remaining three triangles untouched. Additionally, if a region G is triangulated and a particular portion of G, but not all of G, is deemed to require further triangulation, it is necessary to identify the triangles that comprise this portion. For a triangulation of a region $\hat{G}$, with a given list of triads and boundary vertices $TriadList$ and $BVList$, respectively, the procedure $FindTriangle(TriadList, BVList, [x_0, y_0])$ determines whether $[x_0, y_0]$ lies within a specific triangle, on the edge of one or more triangles of the subdivision, or outside the triangulated region.

Example 12.3. The Points in a Trapezoid. We consider the triangulation of the trapezoid obtained in Example 12.2 and apply the procedure $FindTriangle$ to determine the location of a number of points $P_0 = [x_0, y_0]$. The triad list and the boundary vertex list given in Example 12.2 are used here. First we set $P_0 = [.26, 4.73]$, and the output is $[.26, 4.73]$ *is not in the region.* That is, P_0 is exterior to the trapezoid. Next, we put $P_0 = [.77, .68]$ and obtain Triangle(2) = $[[1, 0], [1, 1], [1/2, 3/4]]$, and $[.77, .68]$ *is an interior point.* This means that P_0 is interior to the triangle $T2$ in the subdivision. (The vertices of $T2$ are displayed.) With $P_0 = [.5, .75]$ all four triangles are displayed together with their vertices, and the statement $[.5, .75]$ *is an interior vertex point.* That is, $[.5, .75]$ is a vertex for all four triangles. The choice $P_0 = [1, 0]$ results in the listing of $T1$ and $T2$ and $[1, 0]$ *is a boundary vertex point.* With $P_0 = [.77, 0]$ it is determined that the point is on the boundary edge of the triangle $T1$ and that the vertices at the ends of the edge are $[0, 0]$, $[1, 0]$. Finally, the point $P_0 = [1/4, 3/8]$ is found to lie on the (interior) edge common to the triangles $T1$ and $T4$, with the vertices at the ends of the edge given as $[0, 0]$, $[1/2, 3/4]$.

Partial Refinement of a Triangulation

Having determined which triangles are candidates for additional subdivision, $RefineTriangST$, $RefineTriangInterior$, and $RefineTriangBound$ generate a triangulation of a single triangle in varying forms.

RefineTriangST replaces a prescribed triangle by four triangles determined by connecting the midpoints of the sides. It also replaces each neighboring triangle that shares an edge with the given triangle, with two triangles constructed by connecting the opposite vertex to the midpoint of the edge. This is required to obtain a valid triangulation of the full region.

RefineTriangInterior creates three new (sub)triangles for one or more of a group of specified list of triangles of a given triangulation. This is done by finding the centroid of each triangle and connecting that point to each of the vertices of the triangle. If a list of triangles is omitted, each triangle in the full triangulation is modified. In contrast to the procedures *RefineTriang* and *RefineTriangST*, the boundary vertices and the original edges remain unchanged in the modified triangulation. Only new vertices and edges that are interior to existing triangles are added. As a result, neighboring triangles need not be modified.

RefineTriangBound creates two new triangles for each of the specified triangles in a given triangulation that contain a boundary edge, by finding the center of each boundary edge and connecting that point with the opposite vertex of the triangle. The edge is determined by specifying its boundary vertices. If a list of lists of boundary vertices and triangles is omitted, the full boundary is modified. Again, neighbors of the specified triangles need not be modified.

Example 12.4. The Modification of a Triangulated Trapezoid. We reconsider the triangulation of the trapezoid obtained in Example 12.2. It was found that the triangle $T4$ has the maximum area, and we construct an additional triangulation of $T4$. The vertices of $T4$ are $[0, 2]$, $[0, 0]$, $[1/2, 3/4]$. Since $T4$ has an edge on the boundary, each of the three procedures presented above can be used. The name $BTrapList$ represents the list of vertices of the trapezoid, and $TrapList$ is the list of triads for the triangulation given by $PolygonTriang$.

RefineTriangST(*TrapList,BTrapList,T4*) triangulates $T4$ (it is represented by a triad of its vertices) in the manner described above. The triangle $T4$ is replaced by four (sub)triangles, and the triangles $T1$ and $T3$ are replaced by two (sub)triangles. The list of vertices in the modified triangulation is [[1/4, 3/8], [1/2, 3/4], [1/4, 11/8], [0, 0], [1, 0], [1, 1], [0, 2], [0, 1]]. The first three points are the interior vertices and the last five are the boundary vertices, with $[1/4, 3/8]$, $[1/4, 11/8]$ as the new interior vertices and $[0, 1]$ as the added boundary vertex. The original triangle $T4$ is replaced by four triangles, $T1$ and $T3$ are replaced by two triangles each, and $T2$ is unchanged.

RefineTriangInterior($TrapList$, $VTrapList$, $BTrapList$, $[T4]$), with the vertex list of the original triangulation given by $VTrapList$, subdivides $T4$ into three triangles. The other three triangles and the boundary vertices are unchanged. The list of vertices in the modified triangulation of the trapezoid is [[1/2, 3/4], [1/6, 11/12], [0, 0], [1, 0], [1, 1], [0, 2]]. The first two points are the interior vertices and the last four are the boundary vertices, with $[1/6, 11/12]$ as the new interior vertex. If the fourth argument in the procedure is eliminated, all four triangles in the original subdivision are triangulated.

Finally, *RefineTriangBound*(*TrapList,VTrapList,BTrapList*, $[[[0, 0], [0, 2]]]$, $[T4]$) subdivides $T4$ into two triangles. The boundary edge of $T4$ that connects the vertices

$[0,0]$ and $[0,2]$ is divided into two equal parts and a new boundary vertex $[0,1]$ is introduced. The other three triangles are unchanged. The output of the procedure is a list of boundary vertices $BVTMod$. It is a modification of $BTrapList$ to which the vertex $[0,1]$ is added. The global variable $VListMod$ gives the modified triad list. Then *Vertexlist(VListMod,BVTMod)* produces the list of vertices in the modified triangulation of the trapezoid, $[[1/2,3/4],[0,0],[0,1],[1,0],[1,1],[0,2]]$, which we denote by MVT. The first point is the interior vertex and the last five are the vertices on the boundary, with $[0,1]$ as the new boundary vertex. The procedure *TriangPlot*$(VListMod, Triangles, MVT)$ displays a graph of the modified triangulation. It is shown in Figure 12.5.

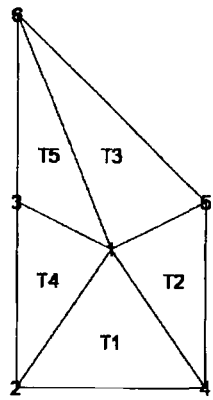

Figure 12.5 The modified triangulation of the trapezoid.

When the last two arguments in *RefineTriangBound* are removed, all four triangles in the original subdivision are triangulated.

Bounding Lines for Triangles Determined by Vertices

The triangles in the triangulation of a region have been prescribed above by lists of their vertices. It is of interest and is necessary for a number of operations relating to the triangles, such as graphics and integration, to describe them in terms of the three lines that constitute the bounds for each triangle. This is easily done by determining the equations of the lines that connect the vertices taken two at a time.

The procedure $ElementBounds(VList)$ determines the three bounding lines for a triangle determined by a list of its three vertices denoted by $VList$. Two sets of equations are given in the output. In one set the lines are given in the form $y = ax + b$ or $x = c$, and in the other set they are given as $x = \hat{a}y + \hat{b}$ or $y = \hat{c}$, with specified coefficients in each case. (The procedure solves the linear equations that determine the constants.)

Example 12.5. The Bounding Lines for the Triangulated Trapezoid. We reconsider the triangulation of the trapezoid obtained in Examples 12.2 and 12.4, and recall that $TrapList$ is the list of triads for the triangulation given by $PolygonTriang$. Then $TrapList[1..3]$ yields a list of the first three vertices in $TrapList$ and these are the vertices of the triangle $T1$. They are given as $(0,0)$, $(1,0)$, $(1/2,3/4)$. We determine the bounding lines of $T1$ by invoking $ElementBounds(TrapList[1..3])$. This yields the two sets of lines

$$\left\{y=0,\ y=\frac{3x}{2},\ y=-\frac{3x}{2}+\frac{3}{2}\right\},\ \left\{x=-\frac{2y}{3}+1,\ y=0,\ x=\frac{2y}{3}\right\}. \tag{12.2.1}$$

We observe that since the x-axis is one of the bounding lines for $T1$ and it cannot be represented in the form $x=f(y)$, the equation $y=0$ is given as a bounding line in the second set of (12.2.1).

Exercises 12.2

12.2.1. Consider the triangle with the vertices $[1,0]$, $[2,3]$, $[4,5]$. (a) Apply the procedure $PolygonTriang$ to construct a triangulation of the triangle and obtain a list of triads for the triangulation. (b) Use the procedure $Vertexlist$ to determine the four vertices of the triangulation. Show that the interior vertex is the centroid of the triangle. (c) Invoke the procedure $TriangPlot$ to obtain a plot of the triangulation that lists the vertices and the triangles in the triangulation. (d) Apply the procedure $RefineTriang$ with the last argument equal to 2, so that a double refinement of the triangulation is carried out. (e) Again use the procedure $Vertexlist$ to determine the interior and boundary vertices of the refined triangulation. (f) Use the procedure $TriangPlot$ to generate a plot of the refined triangulation where the vertices and triangles are not numbered.

12.2.2. Use the procedures $PolygonTriang$ and $RefineTriang$ to reproduce the results given in Example 12.1. Additionally, use $TriangPlot$ to generate a plot for each of the triangulations.

12.2.3. Consider the square with vertices $[0,0]$, $[1,0]$, $[1,1]$, $[0,1]$. Divide each side of the square into four equal parts and take the resulting 16 points to represent boundary vertices for the square. (a) Apply the procedure $PolygonTriang$ to construct a triangulation of the square (using the 16 boundary vertices) and obtain a list of triads for the triangulation. (b) Use the procedure $Vertexlist$ to determine the vertices of the triangulation. Show that the interior vertex is the centroid of the square. (c) Invoke the procedure $TriangPlot$ to obtain a plot of the triangulation that lists the vertices and the triangles in the triangulation. (d) Apply the procedure $RefineTriang$ with the last argument equal to 2, so that a double refinement of the triangulation is carried out. (e) Again use the procedure $Vertexlist$ to determine the interior and boundary vertices of the refined triangulation. (f) Use the procedure $TriangPlot$ to generate a plot of the refined triangulation where the vertices and triangles are not numbered.

12.2.4. Use the procedures referred to in Example 12.2 to obtain all the results of that example.

12.2.5. Apply $FindTriangle$ to verify the results of Example 12.3.

12.2.6. Reproduce the results of Example 12.4.

12.2.7. Determine all the bounding lines for the triangulation of the trapezoid in Example 12.5 by using the procedure $ElementBounds$.

12.2.8. Find all the bounding lines for the triangulation of the triangle obtained in part (a) of Exercise 12.2.1 by using the procedure $ElementBounds$.

12.2.9. Use the procedure $TriangArea$ to determine the areas of the triangles obtained in part (a) of Exercise 12.2.1.

12.2.10. Consider the points [1,1], [0,1], [0,1/2], [0,0], [1,0], [2,0], [2,1/2], [2,1], [2,2], [1,2]. (a) Form a list of lists of these points in which the point $[1, 1]$ is listed at the beginning and the end. Use *plot* to obtain a graph of the polygon with the points as vertices. (It is a backward L-shaped region.) (b) Apply $PolygonTriang$ to the list of boundary vertices and generate a triangulation of the polygonal region. (c) Invoke $Vertexlist$ to the triangulation and then $TriangPlot$ to generate a plot of the triangulation that numbers the vertices and the triangles in the triangulation. (d) Add $[1.3, .7]$ as the second argument of $PolygonTriang$ in part (b) and obtain a modified triangulation of the region. Use $TriangPlot$ to plot the modified triangulation. (e) Use $TriangArea$ to determine the smallest and largest areas of the triangles obtained in parts (b) and (d).

12.3 FINITE ELEMENT OPERATIONS

Having determined how to triangulate a region, we now show how to represent functions over triangulated regions $\hat{G}$. We assume that the definitions of all functions given over the region G can be extended to the region $\hat{G}$ if G is a proper subset of $\hat{G}$. Although prescribed continuous functions retain their values over all triangles in the triangulation of $\hat{G}$, they will generally be approximated by constants over individual triangles and boundary line segments of the triangulation. Additionally, some functions will be defined directly in piecewise form, with values prescribed over individual triangles. This means that we will be dealing, in general, with piecewise-defined functions over the full triangulation.

In the application of the *finite element Galerkin method* for the solution of boundary value and initial and boundary value problems for PDEs, it is necessary to carry out operations, such as multiplication, differentiation, and integration, on these piecewise defined functions, as well as to create such functions directly when constructing candidates for approximate solutions. Consequently, the finite element solutions that will be determined are piecewise-defined functions. Rather than represent the finite element functions in case form, where the function values are displayed (in equation form) for each triangle in the triangulation, we represent them as ordered lists. Each

member of the list determines a value of the function in a specific triangle of the triangulation.

As we have seen in Section 12.1, the fundamental components of the finite element Galerkin method are the basis functions associated with a given triangulation. We now show how to construct these basis functions and carry out the operations needed to find an approximate finite element solution of a problem via the Galerkin method. This will be done directly for some simple cases, but in general, specially constructed Maple procedures will be used.

Plane Elements for a Triangle

We begin by showing how to construct *plane elements* for each triangle in the triangulated subdivision of a region and for the entire triangulation. The plane elements have the general form $ax + by + c$, and they determine a plane $z = ax + by + c$. If the three vertices of a triangle are given as (x_1, y_1), (x_2, y_2), and (x_3, y_3), we determine three planes for which z successively assumes the value 1 at one of the vertices and vanishes at the other two vertices. For example, one of these planes passes through the points $(x_1, y_1, 1), (x_2, y_2, 0), (x_3, y_3, 0)$ in (x, y, z)-space. However, if we set $z = 0$ at all three vertices of the triangle, the corresponding plane element is 0 and the plane is given as $z = 0$. These plane elements will be used to construct the *piecewise linear basis functions* that occur in the *Galerkin method* for the solution of the given problem for the PDE using finite elements.

The procedure $PlaneElements([[x_1, y_1], [x_2, y_2], [x_3, y_3]])$ automates the solution process and obtains three nonzero plane elements for a triangle determined by the three listed vertices. Each element vanishes at two of the vertices and equals 1 at the third vertex. If one of the vertices is given as the second argument, only the plane element for the triangle that is associated with the given vertex is displayed. By applying the procedure $PlaneElements$ to each triad of a full triangulation for a region, the full set of plane elements can be determined.

This can be done by invoking $PlaneElementslist(TriadList)$. It constructs a complete list of triangles and their plane elements for a list of triads (i.e., $TriadList$) of vertices for full triangulation of a region. If a second argument $Triangles$ is added, the triangles and the corresponding plane elements are listed. Otherwise, only the plane elements are listed. The global variables $T||i$, $i = 1..n$, give the vertices of the n triangles in the triangulation, while the global variables $P||i$, $i = 1..n$, give the plane elements for each region.

Example 12.6. Plane Elements for the Triangulated Trapezoid. We refer to the triangulation of the trapezoid given in Example 12.2 and consider the triangle $T1$, whose vertices are $(0, 0), (1, 0), (1/2, 3/4)$. The general equation of a plane [not perpendicular to the (x, y)-plane] is $z = ax + by + c$. First we obtain the equation of the plane that passes through the points $(0, 0, 1), (1, 0, 0), (1/2, 3/4, 0)$. These values of x, y, and z are substituted into $z = ax + by + c$, and the resulting linear system of three equations is solved for a, b, c to get $z = -x - 2y/3 + 1$. Similarly, we obtain $z = x - 2y/3$ for the points $(0, 0, 0), (1, 0, 1), (1/2, 3/4, 0)$ and $z = 4y/3$ for the

points $(0,0,0),(1,0,0),(1/2,3/4,1)$. The same process can be carried out for the vertices in each of the remaining triangles $T2, T3, T4$.

$PlaneElements([[0,0],[1,0],[1/2,3/4]])$ automates the construction of the three nonzero plane elements for the triangle $T1$. The output is

$$[0,0], \quad -x-\frac{2y}{3}+1, \qquad [1,0], \quad x-\frac{2y}{3}, \qquad \left[\frac{1}{2},\frac{3}{4}\right], \quad \frac{4y}{3}. \tag{12.3.1}$$

Each vertex of $T1$ is listed together with the plane element that equals 1 at that vertex and vanishes at the other two vertices. The plane elements for the full triangulation of the trapezoid are given by $PlaneElementslist(TrapList, Triangles)$, where $TrapList$ lists the triads for the full triangulation of the trapezoid as given in Example 12.4. The output is

$$T1 = \left[[0,0],\ -x-\frac{2y}{3}+1, \quad [1,0],\ x-\frac{2y}{3}, \quad \left[\frac{1}{2},\frac{3}{4}\right],\ \frac{4y}{3}\right],$$

$$T2 = \left[[1,0],\ \frac{x}{2}-y+\frac{1}{2}, \quad [1,1],\ \frac{3x}{2}+y-\frac{3}{2}, \quad \left[\frac{1}{2},\frac{3}{4}\right], \quad -2x+2\right],$$

$$T3 = \left[[1,1],\ \frac{5x}{3}+\frac{2y}{3}-\frac{4}{3},\ [0,2],\ -\frac{x}{3}+\frac{2y}{3}-\frac{1}{3},\ \left[\frac{1}{2},\frac{3}{4}\right],\ -\frac{4x}{3}-\frac{4y}{3}+\frac{8}{3}\right],$$

$$T4 = \left[[0,2],\ -\frac{3x}{4}+\frac{y}{2}, \quad [0,0],\ -\frac{5x}{4}-\frac{y}{2}+1, \quad \left[\frac{1}{2},\frac{3}{4}\right],\ 2x\right]. \tag{12.3.2}$$

The added argument $Triangles$ in the procedure causes Maple to identify the triangles in the triangulation and to list the vertices and the plane elements for each triangle. The output for each triangle is interpreted as before.

Finite Element Basis Functions

Once the plane elements for a triangulation have been determined using the procedure $PlaneElementslist$, we are in a position to use these elements to construct the *basis functions* for the finite element method. A basis function is constructed for each vertex in the triangulation, be it an interior or aboundary vertex. For a specified vertex, the basis function is chosen to equal 1 at the vertex and to vanish at all the other vertices in the full triangulation. The basis function is a piecewise planar function whose value in each triangle of the subdivision is given by one of the three nonzero plane elements associated with that triangle, as found above, or by the plane element 0. The specific element selected depends on the vertex values of the basis function in that triangle. If the specified vertex is not a vertex in a specific triangle of the subdivision, the plane element for that triangle is 0.

$FEBasisFunc(PEList, P_0)$ constructs the piecewise planar finite element basis function defined over the entire domain associated with a prescribed list of plane elements $PEList$ and a vertex P_0. The function vanishes at the vertices of each

triangular element except at the vertex P_0, where its value is 1. The first argument $PEList$ is the output of the procedure $PlaneElementslist$ as applied for the given triangulation. The planar function values can be reproduced by invoking $F||i(P_0)$, where i is the triangle number.

Example 12.7. Basis Functions for the Triangulated Trapezoid. We refer to the triangulation of the trapezoid obtained in Example 12.2. There are five vertices, one in the interior and four on the boundary. As a result, five basis functions must be constructed. The plane elements for each of the four triangles in the triangulation are given in (12.3.2). To determine the basis function for the interior vertex $(1/2, 3/4)$, for example, we note that the last entry on the right side of each of the four equations in (12.3.2) gives the plane element associated with the vertex $(1/2, 3/4)$ for each of the four triangles. Let the list of vertices be given as $VPL = [(1/2, 3/4), (0, 0), (1, 0), (1, 1), (0, 2)]$ and the corresponding basis functions by $v_i(x, y)$, $i = 1, \ldots, 5$. Thus, for example, the basis function for the vertex $(1/2, 3/4)$ is

$$v_1(x, y) = \left[\frac{4y}{3},\ -2x + 2,\ -\frac{4x}{3} - \frac{4y}{3} + \frac{8}{3},\ 2x\right]. \tag{12.3.3}$$

The four entries in the list correspond to the plane elements for the four triangles $T1, \ldots, T4$, respectively, in the triangulation. That is, $v_1(x, y)$ is a piecewise-defined function with values in each triangle given by the elements in (12.3.3). Each element equals 1 at the vertex $(1/2, 3/4)$ and vanishes at the two other vertices in each triangle. The basis function can be expressed in piecewise form as

$$v_1(x, y) = \begin{cases} 4y/3, & (x, y) \in T1, \\ -2x + 2, & (x, y) \in T2, \\ -4x/3 - 4y/3 + 8/3, & (x, y) \in T3, \\ 2x, & (x, y) \in T4. \end{cases} \tag{12.3.4}$$

For example, the vertices in $T3$ are $(1, 1), (0, 2), (1/2, 3/4)$, and its plane element is $-4x/3 - 4y/3 + 8/3$. It vanishes at $(1, 1), (0, 2)$ and assumes the value 1 at $(1/2, 3/4)$.

As it is cumbersome to express basis functions and other finite element functions in piecewise form as in (12.3.4), especially when dealing with triangulations that contain a large number of triangles, we represent such functions in list form below, as was done in (12.3.3). Each element in the list is identified with a specific triangle in the triangulation.

The foregoing can be obtained from $FEBasisFunc(PEList, [1/2, 3/4])$, where $PEList$ is the output of $PlaneElementslist(TrapList)$. We do not display this result but apply instead $FEBasisFunc(PEList, [0, 2])$ which yields

$$Vertex = [0, 2],\ T1 = 0,\ T2 = 0,\ T3 = -x/3 + 2y/3 - 1/3,\ T4 = -3x/4 + y/2, \tag{12.3.5}$$

as the plane elements of the basis function $v_5(x, y)$ for the vertex $(0, 2)$. Since $T1$ and $T2$ do not contain the vertex $(0, 2)$, the plane elements for these triangles are zero. The plane elements for $T3$ and $T4$ assume the correct values at the vertices.

Plots of Basis Functions

The finite element basis functions $v_i(x, y)$ are piecewise planar functions whose graph $z = v_i(x, y)$ can be characterized as a portion of a three-dimensional polygon. It is possible to use the Maple procedure *plot3d* to plot $z = v_i(x, y)$ if it is represented as a Maple piecewise function. However, because it is a polygonal surface, it is simpler to use the Maple procedure *polygonplot3d* from the *plots* package to generate the plot. Given a list of three-dimensional points (the vertices of the polygon), the procedure plots a polygonal surface determined by these points. The procedure $ElementPlot(TriadList, V_i)$ plots the basis function $v_i(x, y)$. The first argument is a list of triads for a given triangulation, and V_i is the i-th vertex in the triangulation.

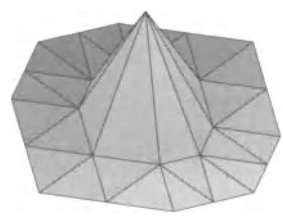

Figure 12.6 The basis function $v_3(x, y)$.

We refer to the triangulation displayed in Figure 12.3 and plot the basis function $z = v_3(x, y)$ that corresponds to the vertex at the origin $(0, 0)$. The procedure $ElementPlot(VListM, VVListM[3])$ generates this plot and the result is exhibited in Figure 12.6. The plot is characteristic of the shape of basis functions. The polygonal surface is continuous and piecewise continuously differentiable. We see that $v_3(0, 0) = 1$ and that $v_3(x, y) = 0$ at all other vertices. The projection of the surface on the (x, y)-plane reproduces the triangulation $\hat{G}$. The plane elements for the triangles that do not contain the vertex $(0, 0)$ are all 0. The plane elements for the triangles that contain the vertex $(0, 0)$ all vanish along the edge that is opposite to the vertex.

Full Set of Finite Element Basis Functions

For a triangulation of a region $\hat{G}$ with n vertices and m triangles, the procedure $FEBasisFuncFull(PEList, VList)$ determines all the finite element basis functions $v_i(x, y)$, $i = 1, \ldots, n$. The first argument $PEList$ is defined as for the procedure $FEBasisFunc$, while the second argument $VList$ is a list of all the vertices in the triangulation as given by the procedure $Vertexlist$, both of which were introduced above. The global variable $F||i(VP||j)$ gives the plane element in the triangle

$T||i$, which has the value 1 at the vertex $VP||j||$. The global variable $vt||i$ is a list whose components are the elements $F||i(VP1)$, $F||i(VP2)$, $F||i(VP3), \ldots$, for each vertex $VP||j$, $j = 1, \ldots, n$, while $vv||i$ is a list whose components are the elements $F1(VP||i), F2(VP||i), F3(VP||i), \ldots$ for each triangle $T||i$, $i = 1, \ldots, m$.

Example 12.8. The Full Set of Basis Functions for the Triangulated Trapezoid. We refer to the triangulation of the trapezoid in the preceding examples. The full list of vertices for initial triangulation is determined from the procedure $Vertexlist(TrapList, BTrapList)$ and is denoted as $VTrapList$. Then the procedure $FEBasisFuncFull(PEList, VTrapList)$ yields

$$\begin{cases} v_1(x,y) = [4y/3,\ -2x+2,\ -4x/3-4y/3+8/3,\ 2x]\,, \\ v_2(x,y) = [-x-2y/3+1,\ 0,\ 0,\ -5x/4-y/2+1]\,, \\ v_3(x,y) = [x-2y/3,\ x/2-y+1/2,\ 0,\ 0]\,, \\ v_4(x,y) = [0,\ 3x/2+y-3/2,\ 5x/3-2y/3-4/3,\ 0]\,, \\ v_5(x,y) = [0,\ 0,\ -x/3+2y/3-1/3,\ -3x/4+y/2]\,, \end{cases} \tag{12.3.6}$$

as the basis functions for the five vertices $(1/2, 3/4)$, $(0,0)$, $(1,0)$, $(1,1)$, $(0,2)$, respectively. Each basis function is a list of four plane elements for the four triangles $T1$, $T2$, $T3$, $T4$ in the triangulation. The addition of a third argument, $Triangles$, in the procedure, results in an output that identifies the vertices and triangles as in (12.3.5).

Now suppose that the n vertices in a given triangulation are (x_i, y_i), $i = 1, \ldots, n$, and that the corresponding piecewise-defined basis functions are denoted as $v_i(x,y)$, $i = 1, \ldots, n$. As we have seen, each basis function has the property that $v_i(x_j, y_j) = \delta_{ij}$, where δ_{ij} is the Kronecker delta. That is, $v_i(x,y)$ vanishes at all vertices (x_j, y_j), $j \neq i$ and $v_i(x_i, y_i) = 1$. Within each triangle of the triangulation, each basis function is linear in x and y, so it has derivatives of all orders. Each of the $v_i(x,y)$ is continuous across a common edge of two adjacent triangles of the triangulation, because the values of the plane elements at the vertices at the ends of each edge are the same for both triangles. Since the planes determined by the plane elements intersect at the two vertices, they must intersect along the edge determined by these vertices, so that the basis function is continuous across the edge. As a result, all the $v_i(x,y)$ are piecewise continuously differentiable functions in the full triangulated region and are integrable over the entire region. They are a linearly independent set because each $v_i(x,y) = 1$ at the vertex (x_i, y_i), while the remaining $v_j(x,y) = 0$, $j \neq i$ at (x_i, y_i).

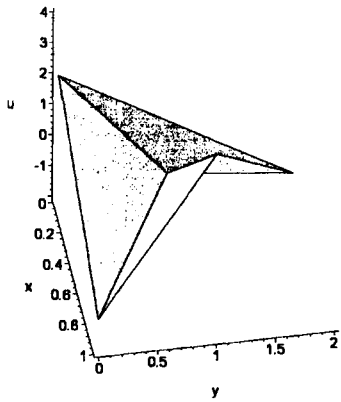

Figure 12.7 The plot of a finite element function.

Finite Element Representations in Terms of Basis Functions and Their Plots

In the finite element method (which we base on the Galerkin method), the approximate solution $\hat{u}(x, y)$ of a given problem for a (second order) PDE in a region G is represented by a linear combination of the basis functions for the triangulated region $\hat{G}$ in the form $\hat{u}(x, y) = \sum_{i=1}^{n} c_i v_i(x, y)$. Some of the c_i are prescribed and others are to be determined, depending on the boundary conditions. (We are assuming in this discussion that the problem is time-independent. Otherwise, both $\hat{u}$ and the c_i depend on t as well.) Once the c_i have been found, the values of $\hat{u}(x, y)$ at the vertices (x_i, y_i) equal c_i, as follows from the properties of the basis functions. The approximate solution $\hat{u}(x, y)$ is thus given as a continuous piecewise planar surface.

Alternatively, we see that the finite element solution determines a list of n points in space given as $(x_i, y_i, \hat{u}(x_i, y_i))$, $i = 1, \ldots, n$. To plot the approximate polygonal solution surface $z = \hat{u}(x, y)$, we can use the Maple procedure *polygonplot3d*, as was indicated above, using the list of n points. (It is unnecessary to use Maple's *plot3d* procedure to plot the surface.) To create this list of points for a given triangulation, we use the procedure $VertexSolution(VList, ValList)$, where $VList$ is the list of vertices in the triangulation and $ValList$ is a list of values to be assigned at the vertices. It combines both lists into a single list of three-component elements that represent the vertices and the values. For example, if we use the vertex list for the triangulated trapezoid of Example 12.7, and use $VertexSolution([[1/2, 3/4], [0, 0], [1, 0], [1, 1], [0, 2]], [c_1, c_2, c_3, c_4, c_5])$, we obtain $[[1/2, 3/4, c_1], [0, 0, c_2], [1, 0, c_3], [1, 1, c_4], [0, 2, c_5]]$.

Let the triad list for the triangulation and the list of n points in space be denoted as $TriadList$ and $PointList$, respectively. $SolutionPlot(TriadList, PointList)$ uses *polygonplot3d* to plot the polygonal surface determined by these points. For the triangulated trapezoid of Example 12.7, the list of triads is given as $TrapList$.

If we assign values for the c_i so that the foregoing list has the form $PL = [[1/2, 3/4, 1], [0, 0, 2], [1, 0, -1], [1, 1, 4], [0, 2, -2]]$, then $SolutionPlot(TrapList, PL)$ yields the plot shown in Figure 12.7. The polygonal surface is composed of four triangles and it is continuous and piecewise continuously differentiable. The vertex values are those assigned by the three-dimensional point list.

Representation of a Function Over a Triangulated Region and Its Evaluation

When solving a problem using the finite element method, not only the solution but also coefficients of the PDE and data for the problem may need to be represented as linear combinations of the basis functions $v_i(x, y)$. As we have seen, each basis function is given as a list of plane elements that correspond to its representation in each of the triangles in the triangulation of the region. Thus, the linear combination of basis functions has the same form. Each basis function can be thought of as a vector with m components, where m is the number of triangles in the triangulation. Thus, linear combinations of basis functions are constructed in the same manner as linear combinations of vectors. We note that for a given function $w(x, y)$, the foregoing linear combination takes on the values of the function at each vertex but has the form of a planar element in each triangle of the subdivision. $FEF(VertValues, BasisList, Vertexlist, [x, y])$ automates the construction of the linear combination of basis functions. The second argument of FEF is a list of basis functions $[v_1(x, y), \ldots, v_n(x, y)]$, and the third argument is a list of the n vertex points of the triangulation. The first argument can either be a list $[c_1, \ldots, c_n]$ of values to be assigned at vertices of the triangulation or a given function $w(x, y)$. The output is a list of plane elements.

Example 12.9. The Representation of a Function Over a Triangulated Trapezoid. We continue with our discussion of the triangulated trapezoid of Example 12.8. If we set $VertValues = [c_1, c_2, c_3, c_4, c_5]$ and $BasisList = [v_1(x, y), v_2(x, y), v_3(x, y), v_4(x, y), v_5(x, y)]$, the procedure $FEF(VertValues, BasisList, VTrapList, [x, y])$ yields the piecewise-defined function

$$\begin{aligned} F(x, y) = [&c_3(x - 2y/3) + c_2(1 - x - 2y/3) + c_1(4y/3), \\ &c_4(3x/2 + y - 3/2) + c_3(x/2 - y + 1/2) + c_1(2 - 2x), \\ c_5(-x/3 + 2y/3 - 1/3) &+ c_4(5x/3 + 2y/3 - 4/3) + c_1(-4x/3 - 4y/3 + 8/3, \\ &c_5(-3x/4 + y/2 + c_2(-5x/4 - y/2 + 1) + c_1(2x)]. \end{aligned} \tag{12.3.7}$$

In the triangle $T1$, for example, $F(x, y)$ has the form $c_3(x - 2y/3) + c_2(1 - x - 2y/3) + c_1(4y/3)$. If we substitute the vertices $(1/2, 3/2), (0, 0), (1, 0)$ of $T1$ into $F(x, y)$, we obtain c_1, c_2, c_3, respectively. Similar results obtain for the triangles $T2$, $T3$, $T4$.

Now, if we put $VertValues = w(x, y)$ as the first argument in the procedure, the output retains the form of (12.3.7), but with $c_1 = w(1/2, 3/4)$, $c_2 = w(0, 0)$, $c_3 = w(1, 0)$, $c_4 = w(1, 1)$, $c_5 = w(0, 2)$. The function $w(x, y)$ is evaluated at the vertices of the triangulated trapezoid. Within each triangle of the triangulation, $F(x, y)$ assumes the values of $w(x, y)$ when it is evaluated at the vertices. As a specific example, we put $w(x, y) = \exp[\cos(xy - 4) + x^3 y]$ and plot it and its finite element representation $F(x, y)$. The graphs of both functions are displayed in Fig. 12.8.

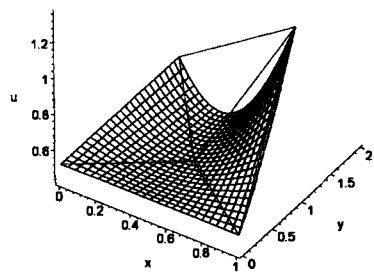

Figure 12.8 A function and its piecewise planar representation.

Once a piecewise planar finite element function has been prescribed and represented as a list of plane elements, if we want to evaluate it at a point in the triangulated region we must first determine which triangle the point lies in. $FindTriangle$ introduced in Section 12.2 does this. Then the plane element that corresponds to this triangle must be selected and evaluated at the given point. The procedure $FEFValue(FEFunction, TriadList, [x_0, y_0], [x, y])$ automates this process. Here $FEFunction$ represents the finite element function determined from the procedure FEF given above, and $TriadList$ is the list of the triads in the given triangulation. The point $[x_0, y_0]$ is where the function is to be evaluated.

Example 12.10. Function Values in a Triangulated Trapezoid. We consider the function $F(x, y)$ of (12.3.7) and enter it as the first argument in $FEFValue$. The second argument is $TrapList$. If we put $[x_0, y_0] = [2, 6]$, the output of the procedure is $[2, 6]$, *is not in the region.* The point $(2, 6)$ lies outside the trapezoid. With $[x_0, y_0] = [1/2, 3/4]$, the interior vertex point, the output of the procedure is c_1. Finally, if $[x_0, y_0] = [1/3, 1/6]$ we obtain the value $2/9\, c_1 + 5/9\, c_2 + 2/9\, c_3$. The point $(1/3, 1/6)$ lies in the triangle $T1$, within which $F(x, y) = c_3(x - 2y/3) + c_2(1 - x - 2y/3) + c_1(4y/3)$.

Finite Element Centroid- and Midpoint-Valued Functions

The finite element method requires the evaluation of double integrals of functions over the triangles of the triangulated region and line integrals of functions over the polygonal boundary, as will be seen. In general, the integrands are products of coefficients, nonhomogeneous terms, or data for the given PDE with piecewise planar finite element functions of the form introduced above. For a large-scale triangulation, this can give rise to a large number of integrals that most often must be evaluated numerically. Since the dimensions of the triangles and the lengths of the boundary segments are expected to be small, the aforementioned functions can be expressed in a simplified form without a significant loss of accuracy in the final approximation to the solution. One method for doing so is to replace these functions by piecewise linear finite element functions, as was done above. This has the effect of introducing integrands that are, at worst, polynomials of degree four and can always be integrated in closed form. Yet the common practice is to simplify the integrands even further by approximating the coefficients, nonhomogeneous terms, or data by piecewise constant functions, that is, they assume constant values in the triangles of the triangulation and on the sides of the polygonal boundary. This generally has the effect of replacing continuous by piecewise continuous functions, but it does not affect their integrability properties.

The approximation of a function over a triangle is made by replacing it by its value at the *centroid* of the triangle. This is done for each triangle in the triangulation. The approximation of a function over a side of polygonal boundary is made by replacing it by its value at the *midpoint* of the line segment that represents the side. This is done for each side of the boundary. The procedures $FECentroid$ and $FEBoundMidP$ carry out this process for a full triangulated region. The first procedure takes the form $FECentroid(f(x,y), TriadList, [x,y])$ and the second procedure is given as $FEBoundMidP(f(x,y), BPList, [x,y])$. As before, $TriadList$ is a list of triads of vertices that determine the triangulation, and $BPList$ is a list of the vertices of the polygon that bounds the triangulated region. To get the correct result for $FEBoundMidP$, the first boundary vertex point must be listed twice, at the beginning and end of the list.

Example 12.11. Approximations for a Triangulated Trapezoid. We begin with the centroid approximation for the function $f(x,y)$ over the triangulation of the trapezoid introduced in the preceding examples. $TrapList$ is defined as before and we use $FECentroid(f(x,y), TrapList, [x,y])$. The output is $[f(1/2,1/4), f(5/6,7/12), f(1/2,5/4), f(1/6,11/12)]$. It lists the values of $f(x,y)$ at the centroids of the four triangles $T1$, $T2$, $T3$, $T4$, respectively.

$FEBoundMidP(f(x,y), [[0,0],[1,0],[1,1],[0,2],[0,0]], [x,y])$ has the output $[f(1/2,0), f(1,1/2), f(1/2,3/2), f(0,1)]$, which is a list of the values of $f(x,y)$ at the midpoints of the four sides of the trapezoid.

Integral of a Finite Element Function Over a Triangulated Region

The next step in the finite element method, after the region is triangulated and the basis functions and related piecewise defined functions are constructed, involves the integration of these functions and their products over the triangulated region. This includes double integration over individual triangles and line integration over the polygonal boundary. We begin with a discussion of double integration over a triangle.

Let $T1, T2, \ldots, Tm$ be the triangles in the triangulation of the region. To carry out a double integration of a function $F(x, y)$ over a triangle Ti, we must determine the lines that bound the triangle and express the double integral as an iterated integral. The procedure $ElementBounds$ introduced above determines these lines. Then the iterated integral(s) can be expressed in $dx\,dy$ form or $dy\,dx$ form. The procedure $FEIntegral(f(x,y), Ti, [x,y])$ automates the integration process. The integrals are always expressed in $dy\,dx$ form.

Example 12.12. Integration Over a Triangle in a Triangulated Trapezoid. It was found in Example 12.5 that the vertices of the triangle $T1$ of the triangulated trapezoid are $(0,0)$, $(1,0)$, $(1/2, 3/4)$. One set of bounding lines was obtained in (12.2.1) as $\{y = 0,\ y = 3x/2,\ y = -3x/2 + 3/2\}$. Thus,

$$\int_{T1} f(x,y)\,dA = \int_0^{1/2}\int_0^{3x/2} f(x,y)\,dy\,dx + \int_{1/2}^{1}\int_0^{-3x/2+3/2} f(x,y)\,dy\,dx. \tag{12.3.8}$$

The output of $FEIntegral(f(x,y), T1, [x,y], blank)$ is identical with the right side of (12.3.8). If the fourth argument in the procedure is omitted, the vertices of $T1$ and the bounding lines for $T1$ are printed out.

We cannot expect, in general, that the foregoing integrals can be evaluated (easily) in closed form. Numerical integration methods will have to be used. This can be a lengthy process when carried out over a large triangulation. A significant simplification is achieved if the coefficients, nonhomogeneous terms, or data for the given PDE are replaced by centroid approximations, as was done above. This has the effect that all integrations over the triangles of the triangulation involve only constant multiples of the linear elements of basis functions and their products. As a result, the integrands are multivariate polynomials, and the integrals can always be evaluated in closed form.

In fact, it is not even necessary to evaluate these integrals directly (using antiderivatives) because the following formula can be used. Consider the triangle Ts of a given triangulation with the vertices $(x_i, y_i), (x_j, y_j), (x_k, y_k)$, and let the corresponding linear elements of the basis functions in Ts be $v_{is}(x,y)$, $v_{js}(x,y)$, $v_{ks}(x,y)$ [i.e., $v_{is}(x_i, y_i) = 1$, $v_{is}(x_j, y_j) = 0$, $v_{is}(x_k, y_k) = 0$, with corresponding results for $v_{js}(x,y)$ and $v_{ks}(x,y)$]. Then

$$\int_{Ts} v_{is}(x,y)^n v_{js}(x,y)^m v_{ks}(x,y)^r\,dA = \frac{2\,n!\,m!\,r!\,A(Ts)}{(2+n+m+r)!}, \tag{12.3.9}$$

where n, m, r are positive integers or zero and $A(Ts)$ is the area of the triangle Ts. The procedure $TriangArea(Ts)$ determines the area of Ts if we set $Ts = [[x_i, y_i], [x_j, y_j], [x_k, y_k]]$.

This result (12.3.9) can be verified by transforming the integral using an (affine) coordinate transformation $\hat{x} = a_1 x + b_1 y + c_1$, $\hat{y} = a_2 x + b_2 y + c_2$, which takes the vertices $(x_i, y_i), (x_j, y_j), (x_k, y_k)$ into the points $(0,0), (1,0), (0,1)$, respectively. (It transforms the triangle Ts into a right triangle with the latter points as vertices.) This requirement on the vertices yields a system of six linear equations for the six unknowns. The system has a unique solution. The Jacobian of the transformation equals twice the area of Ts, and we obtain

$$\int_{Ts} v_{is}(x,y)^n v_{js}(x,y)^m v_{ks}(x,y)^r \, dA = 2A(Ts) \int_0^1 \int_0^{1-\hat{x}} (1-\hat{x}-\hat{y})^n \hat{y}^m \hat{x}^r \, d\hat{y} \, d\hat{x}, \tag{12.3.10}$$

where the domain of integration on the right is the right triangle. The linear function $1 - \hat{x} - \hat{y}$ in the integrand, vanishes at the vertices $(1,0)$ and $(0,1)$ and equals 1 at the vertex $(0,0)$. A related result is valid for the other two linear functions $\hat{x}$ and $\hat{y}$ in the integrand. Repeatedly integrating by parts in the iterated integral of (12.3.10) yields (12.3.9).

Example 12.13. Integrals of Products of Linear Elements of Basis Functions. The vertices of the triangle $T1$ of the triangulated trapezoid are $(0,0)$, $(1,0), (1/2, 3/4)$. Let the basis functions for these vertices be denoted as $v_1(x,y)$, $v_2(x,y), v_3(x,y)$, respectively. It was shown in (12.3.1) that the linear elements of the three basis functions in $T1$ are $v_{11}(x,y) = -x - 2y/3 + 1$, $v_{21}(x,y) = x - 2y/3$, $v_{31}(x,y) = 4y/3$. Then, on using $FEIntegral$ or by direct integration, we obtain

$$\int_{T1} v_{11}(x,y)^4 v_{21}(x,y)^3 v_{31}(x,y)^2 \, dA = 1/184,800 = A(T1)/69,300. \tag{12.3.11}$$

The last result follows from (12.3.9). Invoking $TriangleArea$ shows that $A(T1) = 3/8$, so that (12.3.11) is verified.

Given a triangulation of a region into the triangles $T1, T2, \ldots, Tm$, we consider a function $\hat{f}(x,y)$ represented as a list $[f_1(x,y), \ldots, f_m(x,y)]$ of m-elements. $f_i(x,y)$ is the representation of $\hat{f}(x,y)$ in the triangle Ti. Then the double integral of $\hat{f}(x,y)$ over the full triangulated region is the sum of the integrals of the $f_i(x,y)$ over the triangles Ti for $i = 1, 2, \ldots, m$. $FEIntegralFull(\hat{f}(x,y), TriadList, [x,y])$, where $TriadList$ is a list of triads for the triangulation, automates this integration process.

Example 12.14. Integration Over the Full Triangulated Trapezoid. The triad list for the triangulated trapezoid is given by $TrapList$. Then $FEIntegralFull([f_1(x,y), f_2(x,y), f_3(x,y), f_4(x,y)], TrapList, [x,y])$ yields

$$\int_0^{1/2}\int_0^{3/2\,x} f_1(x,y)\,dy\,dx + \int_{1/2}^{1}\int_0^{-3/2\,x+3/2} f_1(x,y)\,dy\,dx$$
$$+ \int_{1/2}^{1}\int_{-3/2\,x+3/2}^{1/2+1/2\,x} f_2(x,y)\,dy\,dx + \int_0^{1/2}\int_{2-5/2\,x}^{2-x} f_3(x,y)\,dy\,dx$$
$$+ \int_{1/2}^{1}\int_{1/2+1/2\,x}^{2-x} f_3(x,y)\,dy\,dx + \int_0^{1/2}\int_{3/2\,x}^{2-5/2\,x} f_4(x,y)\,dy\,dx. \qquad (12.3.12)$$

The iterated integrals represent integration over the four triangles in the triangulation.

Again, if centroid approximations are used in the finite element method, each of the $f_i(x,y)$ is given as a product of powers of linear elements of basis functions, as indicated above. As a result, each of the integrals can be evaluated in closed form (because the integrands are multivariate polynomials) or by using (12.3.9).

Line Integral of a Finite Element Function Over a Full or Partial Boundary

Finally, we consider line integration over the polygonal boundary of the triangulated region. Depending on the form of the boundary conditions for the problem, it may be necessary to integrate certain functions over the full boundary or only a part of it. In any case, the line integral is broken up into a sum of line integrals over the line segments that comprise the domain of line integration. Each line segment is determined by two adjacent vertices, say P_1 and P_2, on the polygon. If the integral extends from P_1 to P_2, the line segment can be expressed parametrically as $x = (x_1 - x_0)t + x_0$, $y = (y_1 - y_0)t + y_0$, $0 \le t \le 1$, with $P_1 = (x_0, y_0)$ and $P_2 = (x_1, y_1)$. The line integral of a function $f(x,y)$ is

$$\int_{P_1}^{P_2} f(x,y)\,ds = L\int_0^1 f((x_1 - x_0)t + x_0, (y_1 - y_0)t + y_0)\,dt, \qquad (12.3.13)$$

where L is the length of the line segment connecting P_1 to P_2..

In general it is necessary to evaluate these integrals numerically. However, if the data or coefficients for the given problem that occur in the line integrals are replaced by midpoint approximations, as was done above, the integrals are greatly simplified. Then the functions $f((x_1 - x_0)t + x_0, (y_1 - y_0)t + y_0)$ arise from linear elements of basis functions and their products, and they are all polynomials in t. Consequently, all the line integrals can be evaluated exactly.

As was the case for integration over the triangles in the triangulation, there is a simple formula that can be used to determine the values of the latter line integrals

without finding antiderivatives of the integrands. Let Tr represent the triangle and Sr the boundary edge of Tr over which the line integration is carried out. With (x_i, y_i) and (x_j, y_j) as the two boundary vertices that comprise the end points of Sr, let $v_{ir}(x, y)$ and $v_{jr}(x, y)$ be the corresponding linear elements of the basis functions for the triangle Tr. [Thus, $v_{ir}(x_i, y_i) = 1$, $v_{ir}(x_j, y_j) = 0$, $v_{jr}(x_i, y_i) = 0$, $v_{jr}(x_j, y_j) = 1$.] Then

$$\int_{Sr} v_{ir}(x,y)^n v_{jr}(x,y)^m \, ds = \frac{n!\, m! L(Sr)}{(1+n+m)!}, \tag{12.3.14}$$

where $L(Sr) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$ is the length of the line segment, n and m are positive integers, and the integration is taken in the positive direction with respect to the full polygonal boundary. The result (12.3.14) follows by noting that on proceeding as was done for the double integral (12.16), the transformed integral in this case assumes the form

$$\int_{Sr} v_{ir}(x,y)^n v_{jr}(x,y)^m \, ds = L(Sr) \int_0^1 (1-t)^n \, t^m \, dt. \tag{12.3.15}$$

Integrating by parts repeatedly yields (12.3.14).

In correspondence with $FEIntegral$ and $FEIntegralFull$ presented above for integrating functions over specific triangles or over the full triangulation of a region, $FEBoundaryIntegral$ and $FEBoundaryIntegralFull$ integrate finite element functions over only a portion of or over the full polygonal boundary of the triangulated region. $FEBoundaryIntegral$ is basically a subprocedure used in carrying out $FEBoundaryIntegralFull$, so we discuss only the latter procedure. It has the form $FEBoundaryIntegralFull(F(x,y), BPointList, TriadList, [x,y])$, where $F(x,y) = [f_1(x,y), \ldots, f_m(x,y)]$ with m as the number of triangles in the triangulation. The second argument is a list of (at least two) consecutive vertices of the polygon. The line integration is carried out over the polygonal segments determined by these vertices. To obtain the line integral over the full polygon, the first vertex in the list must be appended at the end of the list of boundary points. As before, $TriadList$ determines the triangulation of the region.

Example 12.15. Integration Over the Boundary of the Triangulated Trapezoid. *TrapList* determines the triad list for the triangulation of the trapezoid, and then the procedure *FEBoundaryIntegralFull*($[f_1(x,y), f_2(x,y), f_3(x,y), f_4(x,y)], [[0,0],[1,0],[1,1],[0,2],[0,0]], Traplist, [x,y]$) yields the line integral over the full polygonal boundary in the form

$$\int_0^1 f_1(t,0)\, dt + \int_0^1 f_2(1,t)\, dt + \sqrt{2} \int_0^1 f_3(1-t, 1+t)\, dt + 2 \int_0^1 f_4(0, 2-2t)\, dt. \tag{12.3.16}$$

Each integral represents integration over one of the four sides of the trapezoid and has the form given in (12.3.13). In each case the integral is multiplied by the length of the side being integrated over. To obtain only the line integral over the side determined

by the vertices $(1,1)$, $(0,2)$, we replace the second argument in the procedure by $[[1,1],[0,2]]$, and the output is the third integral in (12.3.16).

Next we integrate a prescribed function $F(x,y)$ over the side of the trapezoid determined by the vertices $(1,0)$ and $(1,1)$. The corresponding triangle in the triangulation has the additional vertex $(1/2, 3/4)$. The (basis) plane elements $x/2 - y + 1/2$ and $3x/2 + y - 3/2$ correspond to the vertices $(1,0)$ and $(1,1)$. Indeed, they both vanish at $(1/2, 3/4)$ and at $(1,1)$ and $(1,0)$, respectively. At the remaining vertex they equal 1. We put $F(x,y) = [0, (x/2-y+1/2)^5(3x/2+y-3/2)^2, 0, 0]$ in the first argument of the foregoing procedure and enter $[[1,0],[1,1]]$ as the second argument. The output of the procedure is $1/168$. The line integral corresponds to that in (12.3.14) with Sr as the side determined by the two vertices $(x_i, y_i) = (1,1)$ and $(x_j, y_j) = (1,0)$, and with $v_{ir} = x/2 - y + 1/2$, $v_{jr} = 3x/2 + y - 3/2$ and $n = 5$, $m = 2$. The length of Sr equals 1. The line integral is $\int_0^1 (1-t)^5 t^2 \, dt = 1/168$.

This completes our discussion of the triangulation of regions and the construction of finite element basis functions, finite element functions, and operations on such functions. The foregoing results will be used in the application of the finite element method to the solution of boundary value problems for elliptic PDEs and initial and boundary value problems for hyperbolic and parabolic PDEs over finite regions in two dimensions.

Exercises 12.3

12.3.1. Apply the procedure *PlaneElements* to determine the plane elements for the four triangles $T1$, $T2$, $T3$, $T4$ of the triangulation of the trapezoid of Example 12.6.

12.3.2. Use the procedure *PlaneElementslist* to obtain the plane elements given in (12.3.2).

12.3.3. Use *PlaneElementslist* to obtain the plane elements for the full triangulation obtained in Exercise 12.2.1(a).

12.3.4. Construct the five basis functions that correspond to the vertices in the triangulation of the trapezoid obtained in Example 12.2 by using the procedure *FEBasisFunc*. [The basis function for the vertex $[0,2]$ is given in equation (12.3.5) of Example 12.7.]

12.3.5. Use *ElementPlot* to plot each of the basis function obtained in Exercise 12.3.4.

12.3.6. Use the procedure *FEBasisFuncFull* to obtain the set of basis functions (12.3.6) given in Example 12.8.

12.3.7. Apply *FEBasisFuncFull* to obtain the full set of basis functions for the triangulation obtained Exercise 12.2.3(a).

12.3.8. Use $ElementPlot$ to plot the first five basis function obtained in the solution of Exercise 12.3.7.

12.3.9. Apply the procedures $VertexSolution$ and $SolutionPlot$ to generate the plot of the finite element function given in Figure 12.7.

12.3.10. Consider the triangulation constructed in Exercise 12.2.1(a). Use the procedures $Vertexlist$ and $VertexSolution$ to assign an arbitrarily chosen set of values at the vertices. Then use $SolutionPlot$ to generate the plot of the resulting finite element function.

12.3.11. Consider the triangulation constructed in Exercise 12.2.3(d). Use the procedures $Vertexlist$ and $VertexSolution$ to assign an arbitrarily chosen set of values at the vertices. (The procedure $randvector$ can be used to generate a set of vertex values.) Then use $SolutionPlot$ to generate the plot of the resulting finite element function.

12.3.12. Reproduce the results of Example 12.9 using the procedure FEF. Find the values of $F(x, y)$ [i.e (12.3.7)] in each of the four triangles of the triangulation. Obtain the plot given in Figure 12.8.

12.3.13. Consider the function $F(x, y)$ given in (12.3.7). Verify the results of Example 12.10 using the procedure $FEFValue$.

12.3.14. Reproduce the results of Example 12.11.

12.3.15. Carry out approximations of Example 12.11 if the triangulated trapezoid is replaced by the triangulated triangle of Exercise 12.2.1(a).

12.3.16. Carry out approximations of Example 12.11 if the triangulated trapezoid is replaced by the triangulated square of Exercise 12.2.3(a).

12.3.17. Use the procedure $FEintegral$ to obtain the result (12.2.8) in Example 12.12 and obtain corresponding integrals for the remaining three triangles in the triangulated trapezoid.

12.3.18. Verify the results of Exercise 12.3.16 by using $FEIntegralFull$ as in Example 12.14 and obtaining the result (12.3.12).

12.3.19. Confirm the validity of (12.3.10) in the special case (12.3.11).

12.3.20. Reproduce the results of Example 12.15.

12.4 THE FINITE ELEMENT METHOD FOR ELLIPTIC EQUATIONS IN TWO DIMENSIONS

Galerkin Integrals for Elliptic Equations

We consider the elliptic PDE in the two-dimensional region G:

$$-(pu_x(x,y))_x - (pu_y(x,y))_y + \beta u_x(x,y) + \gamma u_y(x,y) + qu(x,y) = F. \quad (12.4.1)$$

It is assumed that p, q, β, γ, F are functions of (x, y) and that $p(x, y) > 0$ in G.

The PDE has the form $-\nabla \cdot (p(x,y)\nabla u(x,y)) + q(x,y)u(x,y) + [\beta(x,y), \gamma(x,y)] \cdot \nabla u(x,y) = F(x,y)$. It is formally *self-adjoint* if $\beta(x,y) = 0$ and $\gamma(x,y) = 0$. If, in addition, $p(x,y) = 1$ and $q(x,y) = 0$, it reduces to *Poisson's equation*. Then with $F(x,y) = 0$, it becomes *Laplace's equation*. *Dirichlet, Neumann, Robin, or mixed boundary conditions* are prescribed for $u(x,y)$ on the boundary S of the region G. As in Section 4.1, the boundary is comprised of the union of the arcs S_1, S_2, and S_3, with Dirichlet conditions given on S_1, Neumann conditions on S_2, and Robin conditions on S_3. Either S_1, S_2, or S_3 or combinations of two of them may be empty. The boundary values are

$$u(x,y)\big|_{S_1} = f,\ u_n(x,y)\big|_{S_2} = g,\ u_n(x,y) + \alpha u(x,y)\big|_{S_3} = h, \qquad (12.4.2)$$

where f, g, α, h are functions of (x,y).

To determine the *Galerkin integral relation* for the problem, we multiply across in (12.4.1) by an arbitrary *admissible function* $v(x,y)$, apply the divergence theorem (or Green's theorem in the plane) as in Chapter 7, and use the (mixed) boundary conditions given above. A function $v(x,y)$ is admissible if it satisfies the condition $v(x,y)\big|_{S_1} = 0$ when S_1 is nonempty. Its values are unrestricted on S_2 and S_3. As a result, the line integral over S_1 that arises in the divergence theorem vanishes. This yields the following *Galerkin relation* for the mixed problem for the elliptic equation (12.4.1);

$$\int_G p(u_x v_x + u_y v_y) + quv + v(\beta u_x + \gamma u_y)\, dA + \int_{S_3} p\alpha uv\, ds$$

$$= \int_G vF\, dA + \int_{S_2} pvg\, ds + \int_{S_3} pvh\, ds. \qquad (12.4.3)$$

(The integrand in the first integral on the left can be given as $p\nabla u \cdot \nabla v + quv + v[\beta, \gamma] \cdot \nabla u$.) Each of the functions in the integral depends on (x,y).

Now, if $u(x,y)$ is a solution of the given boundary value problem for (12.4.1), the Galerkin relation (12.4.3) is satisfied for all admissible $v(x,y)$. We reverse the process and use the fact that if the Galerkin relation is satisfied for all admissible $v(x,y)$, the original boundary value problem is satisfied by $u(x,y)$, in view of the arbitrariness of $v(x,y)$. The advantage of using this approach is that in contrast to the original formulation of the problem, the solution $u(x,y)$ in the Galerkin relation need not be twice differentiable. The same is true for the admissible functions. The Galerkin relation can be satisfied even if they have only piecewise continuous first derivatives. Thereby the concept of a solution to a given boundary value problem is extended and weakened, as has been done on a number of times in the book in various contexts. However, rather than constructing a weak solution and showing that it satisfies the Galerkin relation, we use this approach to construct an approximate weak solution for the given boundary value problem using the *Galerkin method*.

The *Galerkin method* determines an approximate solution to the foregoing boundary value problem as follows. We select a set of m linearly independent admissible basis functions $v_j(x,y)$, $j = 1, \ldots, m$. As such, each $v_j(x,y)$ must vanish on the

portion S_1 of the boundary of G but is otherwise arbitrary. In addition, we choose a function $w(x, y)$ that is required to satisfy the boundary condition on S_1 for the given problem [i.e., $w(x, y) = f(x, y)$ on S_1] but is otherwise arbitrary. Then we represent the (approximate) solution as

$$u(x, y) = w(x, y) + \sum_{i=1}^{m} c_i v_i(x, y), \tag{12.4.4}$$

with arbitrary constant coefficients c_i that are to be determined. This expansion is inserted into the Galerkin relation (12.4.3) where we consecutively set $v(x, y) = v_j(x, y)$, $j = 1, \ldots, m$. This yields the following m simultaneous equations for the m constants c_i, for $j = 1, \ldots, m$,

$$\int_G p\left[\left[\frac{\partial w}{\partial x} + \sum_{i=1}^{m} c_i \frac{\partial v_i}{\partial x}\right]\frac{\partial v_j}{\partial x} + \left[\frac{\partial w}{\partial y} + \sum_{i=1}^{m} c_i \frac{\partial v_i}{\partial y}\right]\frac{\partial v_j}{\partial y}\right] + q\left[w + \sum_{i=1}^{m} c_i v_i\right] v_j \, dA$$

$$+ \int_G \beta\left[\frac{\partial w}{\partial x} + \sum_{i=1}^{m} c_i \frac{\partial v_i}{\partial x}\right] v_j + \gamma\left[\frac{\partial w}{\partial y} + \sum_{i=1}^{m} c_i \frac{\partial v_i}{\partial y}\right] v_j \, dA \tag{12.4.5}$$

$$+ \int_{S_3} p\alpha\left[w + \sum_{i=1}^{m} c_i v_i\right] v_j \, ds = \int_G F v_j \, dA + \int_{S_2} p v_j g \, ds + \int_{S_3} p v_j h \, ds.$$

On solving these equations and substituting the result in the expansion (12.4.4) for $u(x, y)$, we obtain an approximate solution to the boundary value problem. This represents *Galerkin's method* for solving the foregoing boundary value problem approximately. Clearly, if the $v_i(x, y)$ are members of a complete set of functions, we expect that the infinite series representation of $u(x, y)$ in terms of these functions will yield a (possibly weak) solution of the boundary value problem. This solution is approximated by the finite sum (12.4.4).

Finite Element Method for Elliptic Equations

The *finite element method* for the solution of (12.4.1)–(12.4.2) employs the Galerkin method with the $v_j(x, y)$ chosen as the finite element basis functions associated with a specific triangulation of the region G constructed as above. As a result, both the solution and the region undergo an approximation. We distinguish between interior and boundary vertices in the triangulation and between points on the boundary at which a Dirichlet condition is assigned and points that determine boundary segments where Neumann or Robin conditions are given. Each vertex gives rise to a piecewise planar basis function and we express the approximate solution $u(x, y)$ of the boundary value problem as a linear combination of the basis functions. A basis function that corresponds to a Dirichlet boundary point is multiplied by $f(x, y)$ evaluated at that point, whereas if it corresponds to an interior point, a Neumann condition, or a Robin condition boundary point, it is multiplied by an arbitrary constant.

Suppose that the triangulation has n vertices given as (x_i, y_i), $i = 1, \ldots, n$, of which $n - m$ (with $i = m + 1, \ldots, n$) are Dirichlet boundary point vertices. We express the approximate solution as

$$u(x, y) = \sum_{i=m+1}^{n} f(x_i, y_i) v_i(x, y) + \sum_{i=1}^{m} c_i v_i(x, y). \tag{12.4.6}$$

Since $v_i(x_i, y_i) = 1$ and $v_i(x, y) = 0$ at all other vertices, we find that $u(x_j, y_j) = f(x_j, y_j)$ if (x_j, y_j) is a Dirichlet vertex and $u(x_j, y_j) = c_j$ if (x_j, y_j) is any other vertex. In the resulting system (12.4.5), we set $w(x, y) = \sum_{i=m+1}^{n} f(x_i, y_i) v_i(x, y)$ and the region G is replaced by the triangulated region $\hat{G}$, while S_2 and S_3 represent the line segments on the polygonal boundary on which Neumann and Robin conditions are prescribed. The $v_i(x, y)$ are piecewise-defined functions composed of plane elements, each of which is continuous and piecewise continuously differentiable over $\hat{G}$. The double integrals over $\hat{G}$ are a sum of integrals over the triangles that comprise the triangulation, and the line integrals over S_2 and S_3 are taken over appropriate line segments of the polygonal boundary of $\hat{G}$.

Each basis function $v_i(x, y)$ is a linear function of x and y in each of the triangles. Thus the x and y derivatives of $w(x, y)$ and the $v_i(x, y)$ are constants in each triangle. Exact integrations in (12.4.5) of terms that involve the products of the $v_i(x, y)$ and their first derivatives, with the functions $p(x, y)$, $q(x, y)$, $\beta(x, y)$, $\gamma(x, y)$, $\alpha(x, y)$, $g(x, y)$, $h(x, y)$, and $F(x, y)$, may be difficult to carry out. However, if these functions are replaced by their centroid approximations in the double integrals over triangles and by their midpoint approximations in the line integrals over the boundary, the resulting integrals contain polynomial integrands and can all be evaluated exactly. In fact, the integration formulas (12.3.10) and (12.3.14) can be used for their evaluation.

Thus, to solve a boundary value problem for the elliptic equation (12.4.1) using the finite element method, we must first triangulate the region, determine the basis functions, evaluate the relevant double and line integrals in the Galerkin relation, and solve the resulting linear system for the constants that determine the approximate solution. This is a rather lengthy process in general. We have created Maple procedures that automate the triangulation of a region, determine basis functions, and carry out additional required steps in setting up the system (12.4.5).

Given a boundary value problem for an elliptic PDE of the form (12.4.1), the procedure $NumEllipticFEMCM$ obtains an approximate solution of the problem via the finite element method. The list of triads and vertices for the triangulation of the approximating polygonal region $\hat{G}$ must be entered in the procedure, together with arguments that prescribe the coefficients, inhomogeneous term, and data for boundary conditions of the first, second, or third kind. The procedure determines centroid and midpoint values as required. The system (12.4.5) is set up, the double and line integrals are evaluated, and the resulting linear system is solved for the unknown c_i in (12.4.6). As stated above, the c_i represent the values of the approximate finite element solution at the corresponding vertex points (x_i, y_i) in the triangulation. Combined with prescribed Dirichlet boundary point values, we obtain (approximate) solution values at all vertex points. This corresponds to the construction of finite difference

solutions at a prescribed set of grid points. In the finite element method, however, the approximate solution (12.4.6) yields a piecewise planar approximation to the exact solution surface. Additionally, solution values in $\hat{G}$ at points that are not vertices can easily be found from (12.4.6).

To see what is involved in the process of solving a boundary value problem with the finite element method, we consider a simple example in some detail.

Example 12.16. A Mixed Problem for an Elliptic PDE. We consider a boundary value for the elliptic PDE

$$-u_{xx}(x,y) - u_{yy}(x,y) + xy\,u(x,y) = xy(x+y) \tag{12.4.7}$$

in the triangular region G bounded by the lines $y = 0$, $x + y = 1$, $x = 0$. The three sides determined by these lines are denoted by S_1, S_2, S_3, respectively. Mixed boundary conditions

$$u(x,y)\big|_{S_1} = x, \quad u_n(x,y)\big|_{S_2} = \sqrt{2}, \quad u_n(x,y) + u(x,y)\big|_{S_3} = y - 1, \tag{12.4.8}$$

are prescribed. This means that Dirichlet, Neumann, and Robin boundary conditions are assigned on S_1, S_2, S_3, respectively. The exact solution is $u(x,y) = x + y$, and we construct an approximate finite element solution.

First, we triangulate the region G. It does not undergo any approximation because of its triangular boundary. The bounding (right) triangle for G has the vertex list $BVList = [[0,0],[1,0],[0,1]]$. We use $PolygonTriang(BVList)$ to triangulate G. The triangulation yields three triangles, whose triad list is $VList = [[0,0],[1,0],[1/3,1/3],[1,0],[0,1],[1/3,1/3],[0,1],[0,0],[1/3,1/3]]$. $Vertexlist$ $(VList, BVList)$ then determines that there is one interior vertex and three boundary vertices in the triangulation. The vertex list is $VVList$ =[[1/3, 1/3], [0, 0], [1, 0], [0, 1]], with [1/3, 1/3] as the interior vertex. It is the centroid of the given triangle. The three triangles in the triangulation are given as, in terms of their vertices, $T1$ =[[0, 0], [1, 0], [1/3, 1/3]], $T2$ =[[1, 0], [0, 1], [1/3, 1/3]], and $T3$ =[[0, 1], [0, 0], [1/3, 1/3]]. The plot of the triangulation is given by $TriangPlot(VList, Triangles, VVList)$ and is displayed in Figure 12.9.

The procedure $PlaneElementslist(VList, VVlist)$ determines the plane elements for each of the three vertices in each of the three triangles. We denote the output by $VPEList$. Then $FEBasisFuncFull(VPEList, VVList)$ finds the four basis functions for the four vertices in the triangulation. These are the $v_i(x,y)$, $i = 1,\ldots,4$, for the problem. We order them according to the vertex list $VVList$ given above, so that $i = 1$ corresponds to the interior vertex $[1/3, 1/3]$, for example. They are given as

$$v_1(x,y) = [3y, -3x - 3y + 3, 3x],\ v_2(x,y) = [-x - 2y + 1, 0, -2x - y + 1],$$

$$v_3(x,y) = [x - y, 2x + y - 1, 0],\ v_4(x,y) = [0, x + 2y - 1, -x + y]. \tag{12.4.9}$$

We see that each $v_i(x,y)$ contains three plane elements that correspond to the three triangles in the triangulation. For example, $v_4(x,y)$ corresponds to the boundary

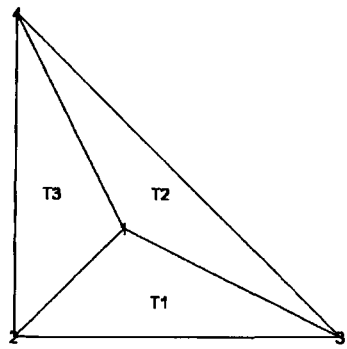

Figure 12.9 The triangulation of the triangle.

vertex $[0,1]$, which is a vertex of the triangles $T2$ and $T3$. As a result, the first linear element, which corresponds to $T1$, is 0. The second linear element, for $T2$, is $x+2y-1$ and its value at $[0,1]$ is 1. The third linear element, for $T3$, is $-x+y$ and its value at $[0,1]$ is 1.

The boundary segments S_1, S_2, S_3, correspond to the line segments determined by the pairs of points $[(0,0),(1,0)]$, $[(1,0),(0,1)]$, $[(0,1),(0,0)]$, respectively. The Dirichlet condition $u(x,y) = x$ on S_1 implies that $u(0,0) = 0$ and that $u(1,0) = 1$. The value of $u(x,y)$ at the remaining two vertices of the triangulation must be determined by the finite element method. Thus, in accordance with (12.4.6), the approximate solution is represented as $u(x,y) = c_1 v_1(x,y) + v_3(x,y) + c_4 v_4(x,y)$. We note that the vertex numbers are not ordered as in (12.4.6). [There is no $v_2(x,y)$ term since the Dirichlet condition is 0 there.] In piecewise form we have

$$u(x,y) = \begin{cases} 3c_1 y + (x-y), & (x,y) \in T1, \\ c_1(-3x-3y+3) + (2x+y-1) + c_4(x+2y-1), & (x,y) \in T2, \\ 3c_1 x + c_4(-x+y), & (x,y) \in T3. \end{cases} \tag{12.4.10}$$

It is easily seen that $u(x,y)$ has the correct values at the vertex points.

The foregoing expression for $u(x,y)$ is inserted into (12.4.5), appropriately modified for the present problem. This yields the following two equations for c_1 and c_4:

$$\int_G \left[\left(\frac{\partial v_3}{\partial x} + c_1 \frac{\partial v_1}{\partial x} + c_4 \frac{\partial v_4}{\partial x} \right) \frac{\partial v_j}{\partial x} + \left(\frac{\partial v_3}{\partial y} + c_1 \frac{\partial v_1}{\partial y} + c_4 \frac{\partial v_4}{\partial y} \right) \frac{\partial v_j}{\partial y} \right] dA$$

$$+ \int_G xy \left(v_3 + c_1 v_1 + c_4 v_4 \right) v_j \, dA + \int_{S_3} \left(v_3 + c_1 v_1 + c_4 v_4 \right) v_j \, ds$$

$$= \int_G xy(x+y) v_j \, dA + \int_{S_2} \sqrt{2}\, v_j \, ds + \int_{S_3} (y-1) v_j \, ds, \; j = 1, 4. \tag{12.4.11}$$

Each integral over G is the sum of three double integrals over the triangles T_1, T_2, T_3, in each of which the appropriate linear element of the v_i must be used. The line integrals also require that the correct linear element of the v_i be used. It is possible, for this problem, to carry out all integrations exactly without introducing any further simplifications. Nevertheless, we apply the centroid approximation to the functions xy and $xy(x+y)$, and the midpoint approximation to $y-1$. The centroid values of xy and $xy(x+y)$ for the triangles T_1, T_2, T_3 are [0.04938, 0.1975, 0.04938] and [0.02743, 0.1756, 0.02743], respectively. The midpoint value of $y-1$ on S_3 is -0.5. These values must be inserted in the appropriate integrals in (12.4.11). Although there are only three triangles to integrate over, it is tedious to carry out all the integrations.

As an example, we select the triangle $T1$ and carry out the integrations for $j=1$. From (12.4.9) we see that the plane element of $v_1(x,y)$ in $T1$ is $3y$. As a result, the corresponding elements of the first derivatives are $\partial v_1/\partial x = 0$ and $\partial v_1/\partial y = 3$. In the first integral in (12.4.11) over $T1$, these terms are multiplied by the x- and y-derivatives of the part of $u(x,y)$ in $T1$, as given in (12.4.10). The x-and y-derivatives equal 1 and $3c_1-1$, respectively. The integrand for the integral over $T1$ is $9c_1-3$, so the integral evaluates to $(9c_1-3)/6$, since the area of the triangle is $1/6$. On evaluating the other two double integrals over $T1$, xy and $xy(x+y)$ must be replaced by their centroid values over $T1$. The line integral over S_3 vanishes when $j=1$. This follows since S_3 is in the triangle $T3$ and (12.4.8) shows that the plane element of v_1 in that triangle equals $3x$. But S_3 lies on $x=0$, so that the plane element on S_3 is zero.

The process of determining the equations for the c_i is extremely tedious and lengthy even for small-scale triangulation. We do not carry out any additional integrations but solve the problem using $NumEllipticFEMCM$. In the general case that corresponds to the mixed boundary value problem (12.4.1)–(12.4.2), we have $NumEllipticFEMCM(TriadList, VertexList, p, q, [\beta,\gamma], F, [x,y], f, DBV, g, NBV, h, \alpha, TKBV)$. The first two arguments represent the triad list for the triangulation of $\hat{G}$ and the list of interior and boundary vertices, respectively. The arguments $DBV, NBV, TKBV$, are lists of boundary vertex points that determine the boundary line segments on which Dirichlet, Neumann, and Robin boundary conditions, respectively, are given. The other arguments are related in an obvious manner to the terms given in (12.4.1)–(12.4.2). If one of the boundary conditions is absent, the corresponding argument is $NONE$ and an empty list of points [ ] is entered. If there is no Robin boundary condition, $NONE$ is entered twice.

Example 12.17. A Mixed Problem for an Elliptic PDE. We use $NumElliptic FEMCM(VList, VVList, 1, xy, [0,0], xy(x+y), [x,y], x+y, BVList[1..2], \sqrt{2}, BVList[2..3], y-1, 1, [BVList[3], BVList[1]])$ to solve the problem of Example 12.16. Here $VList$, $VVList$, and $BVList$ represent the triad list, the vertex list, and the boundary vertex list for the triangulation that was constructed above. There are three boundary vertices, and $BVList[1..2]$ lists the first two of these vertices while $BVList[1]$ selects the first boundary vertex. The other arguments are given as

in the general case. The output is

$$[[1/3, 1/3, 0.6328],\ [0, 0, 0],\ [1, 0, 1],\ [0, 1, 0.8990]]. \tag{12.4.12}$$

The first two arguments in each term of the four terms in the list are the coordinates of the four vertex points, and the third argument is the value of the solution at these points. Since $(0, 0)$ and $(1, 0)$ are Dirichlet vertex points, the solution values 0 and 1 are exact at those points. The other two solution values are $c_1 = 0.6328$ and $c_4 = 0.8990$, and they correspond to the vertices $(1/3, 1/3)$ and $(0, 1)$. Since the exact solution is $u(x, y) = x + y$, the results are not so good. Improved results can be achieved if the triangulation is refined or if the integrations are carried out exactly rather than by using centroid or midpoint approximations. Such approximations lead to a poor result if there are large triangles in the triangulation, as is the case here.

The global variable $FEECM$ yields the piecewise linear representation of the finite element solution,

$$u(x, y) = \begin{cases} x + 0.8985y, & (x, y) \in T1, \\ 1.001x + 0.8995y - 0.00053, & (x, y) \in T2, \\ 0.9995x + 0.8990y, & (x, y) \in T3. \end{cases} \tag{12.4.13}$$

(If the results were exact, each element would be $x + y$.) We can use (12.4.13) to evaluate $u(x, y)$ at points other than the vertices in the trangulation. This can also be done by using $FEFValue(FEECM, VList, [x_0, y_0], [x, y])$. If the point (x_0, y_0) lies outside the triangulated region, the output declares this fact. Otherwise, $u(x_0, y_0)$ is displayed.

To improve our results, we refine the triangulation of the region. With $VList$ and $BVList$ defined as above, we apply $RefineTriang(VList, BVList, 2)$ and obtain a new triad list that we call $VList2$. The modified list of boundary vertices is $BPListMod$. Then $Vertexlist(VList2, BPListMod)$ determines a list of vertices that we call $VVList2$. It finds that the refined triangulation contains 19 interior vertices and 12 boundary vertices. $NumEllipticFEMCM(VList2, VVList2, 1, xy, [0, 0], xy(x + y), [x, y], x + y, BPListMod[1..5], \sqrt{2}, BPListMod[5..9], y - 1, 1, [op(BPListMod[9..12]), BPListMod[1]])$ obtains the corresponding finite element solution.

We do not display the full output of the procedure, but only compare results with those given at the two non-Dirichlet vertex points in (12.4.12). The new values are $[1/3, 1/3, 0.6655]$ and $[0, 1, 0.9883]$, a considerable improvement on the values given in (12.4.12). The exact solution values are $2/3$ and 1.

The global variable $VSolECM$ reproduces the list of points determined by $NumEllipticFEMCM$. $P1 = plot3d(x + y, y = 0..1 - x, x = 0..1)$ and $P2 = SolutionPlot(VList2, VSolECM)$, together with $display(\{P1, P2\})$, yield plots of the exact and finite element solutions shown in Figure 12.10.

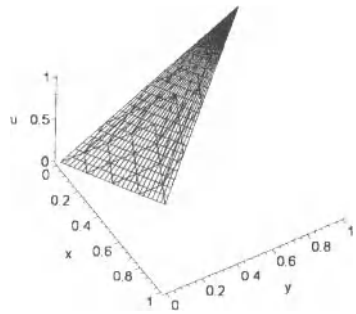

Figure 12.10 Plots of the exact and finite element solutions.

Exercises 12.4

12.4.1. Derive the Galerkin relation (12.4.3).

12.4.2. Verify that $u(x, y) = x + y$ is the exact solution of (12.4.7)–(12.4.8).

12.4.3. Carry out all the steps in Example 12.16.

12.4.4. Use the procedure $NumEllipticFEMCM$ to obtain the finite element solution (12.4.13) in Example 12.17.

12.4.5. Obtain the finite element solution of the boundary value problem of Example 12.17 using $NumEllipticFEMCM$ if a double refinement of the initial triangulation is introduced. Use $SolutionPlot$ to plot the solution.

12.4.6. Noting that $u(x, y) = x + y$ is an exact solution of (12.4.7), set up a Dirichlet problem in the triangular region G defined following (12.4.7) which has $u(x, y) = x + y$ as its exact solution. Proceed as in Exercises 12.4.3 and 12.4.4 to obtain the finite element solution of the Dirichlet problem.

12.4.7. Repeat Exercise 12.4.5 but replace the Dirichlet conditions by Robin boundary conditions with the parameter $h = 1$. (That is, set up the correct boundary conditions and solve the problem using the finite element method.)

12.4.8. The function $u(x, y) = \cos(xy)$ is a solution of the elliptic PDE $-(e^x u_x(x, y))_x - (e^x u_y(x, y))_y + \sinh(x-y)u_x(x, y) + u_y(x, y)/(1+x^4) + (1+x^2+y^2)u(x, y) = e^x y \sin(xy) + e^x y^2 \cos(xy) + e^x x^2 \cos(xy) - y \sinh(x-y) \sin(xy) - x \sin(xy)/(1+x^4) + (1 + x^2 + y^2) \cos(xy)$. Consider the square defined in Exercise 12.2.3 and prescribe $u(x, y) = \cos(xy)$ as the Dirichlet condition for the PDE on the square. Use the triangulation of Exercise 12.2.3(d) and $NumEllipticFEMCM$ to obtain a finite element solution of the problem. Plot the exact and finite element solutions using $plot3d$ and $SolutionPlot$.

12.4.9. Solve the problem (11.4.14)–(11.4.15) by the finite element method. Use the triangulation of Exercise 12.2.3(a) and introduce a single refinement via *RefineTriang*.

Then invoke $NumEllipticFEMCM$ and compare the finite element solution values at the points that correspond to those given in Example 11.10. Plot the finite element solution using $SolutionPlot$.

12.4.10. The function $u(x, y) = x + y + c$, where c is an arbitrary constant, is an exact solution of Laplace's equation. Consider the unit square with vertices $[0, 0]$, $[1, 0]$, $[1, 1]$, $[0, 1]$ and set up a Neumann problem for Laplace's equation in the unit square that has the foregoing function as its nonunique solution. Solve this problem by the finite element method. Use $PolygonTriang$ to triangulate the square using the four vertices of the square as the boundary points. (There will be only one interior point.) Then invoke $NumEllipticFEMCM$ to determine a nonunique finite element solution.

12.5 THE FINITE ELEMENT METHOD FOR PARABOLIC EQUATIONS IN TWO DIMENSIONS

Galerkin Integrals for Parabolic Equations

We consider the parabolic PDE in the two-dimensional spatial region G,

$$\rho u_t - (pu_x)_x - (pu_y)_y + qu + \beta u_x + \gamma u_y = \rho F, \tag{12.5.1}$$

where u, F are functions of (x, y, t), $\rho, p, q, \beta, \gamma$ are functions of (x, y), with $(x, y) \in G$ and $t > t_0$. It is assumed that $p(x, y) > 0$ and $\rho(x, y) > 0$ in G. If $\beta(x, y) = 0$, $\gamma(x, y) = 0$, $p(x, y) = c^2$, $\rho(x, y) = 1$, and $q(x, y) = 0$, (12.5.1) is the nonhomogeneous *heat or diffusion equation.* An *initial condition*

$$u(x, y, t_0) = r(x, y), \qquad (x, y) \in G, \tag{12.5.2}$$

is prescribed at $t = t_0$. *Dirichlet, Neumann,* or *Robin boundary conditions*, or *mixed boundary conditions* that combine the three types of boundary conditions, are prescribed for $u(x, y, t)$ on the boundary S of the region G for all $t > t_0$. As in Section 4.1, the boundary for each $t \geq t_0$ is comprised of the union of the arcs S_1, S_2, and S_3, with Dirichlet conditions given on S_1, Neumann conditions on S_2, and Robin conditions on S_3. Either S_1, S_2 or S_3 or combinations of two of them may be empty. The boundary values are

$$u(x, y, t)\big|_{S_1} = f,\ u_n(x, y, t)\big|_{S_2} = g,\ u_n(x, y, t) + \alpha(x, y)u(x, y, t)\big|_{S_3} = h, \tag{12.5.3}$$

where f, g, h are functions of (x, y, t).

To determine the *Galerkin integral relation* for the problem, we multiply across in (12.5.1) by an arbitrary *admissible function* $v(x, y)$, apply Green's theorem in the plane as was done above for the elliptic PDE, and use the (mixed) boundary conditions given above. It is again required that all admissible $v(x, y)$ satisfy the condition $v(x, y)\big|_{S_1} = 0$, if S_1 is nonempty, but their values are unrestricted on S_2

and S_3. As a result, the line integral over S_1 that arises in Green's theorem must vanish. This yields the following Galerkin relation for the mixed problem for the parabolic equation (12.5.1):

$$\int_G \rho v u_t + p(u_x v_x + u_y v_y) + quv + v(\beta u_x + \gamma u_y)\, dA + \int_{S_3} p\alpha uv\, ds$$

$$= \int_G v\rho F\, dA + \int_{S_2} pvg\, ds + \int_{S_3} pvh\, ds. \tag{12.5.4}$$

Now, if $u(x, y, t)$ is a solution of the initial and boundary value problem (12.5.1)–(12.5.3), the Galerkin relation (12.5.4) is satisfied for all admissible $v(x, y)$. We reverse the process and use the fact that if the Galerkin relation is satisfied for all admissible $v(x, y)$, the original initial and boundary value problem is satisfied by $u(x, y, t)$, in view of the arbitrariness of $v(x, y)$. Again, the Galerkin relation can be satisfied even if $u(x, y, t)$ and $v(x, y)$ have only piecewise continuous first derivatives in x and y. Thereby the concept of a solution to a given initial and boundary value problem is extended and weakened. However, rather than constructing a weak solution and showing that it satisfies the Galerkin relation, we use this approach to construct an approximate weak solution for the given initial and boundary value problem using the *Galerkin method.*

The *Galerkin method* determines an approximate solution to the foregoing initial and boundary value problem as follows. We select a set of m linearly independent admissible basis functions $v_j(x, y)$, $j = 1, \ldots, m$. Each $v_j(x, y)$ must vanish on the portion S_1 of the boundary of G but is otherwise arbitrary. In addition, we choose a function $w(x, y, t)$ that is required to satisfy the boundary condition on S_1 for the given problem [i.e., $w(x, y, t) = f(x, y, t)$ on S_1] but is otherwise arbitrary. We represent the (approximate) solution as

$$u(x, y, t) = w(x, y, t) + \sum_{i=1}^{m} c_i(t) v_i(x, y), \tag{12.5.5}$$

with arbitrary coefficients $c_i(t)$ that are to be determined. The expansion is inserted into the Galerkin relation (12.5.4), where we consecutively set $v(x, y) = v_j(x, y)$, $j = 1, \ldots, m$. This yields the following system of m simultaneous first order ODEs for the m functions $c_i(t)$:

$$\int_G \rho \left[w_t + \sum_{i=1}^{m} c_i' v_i \right] v_j + p \left[\left[w_x + \sum_{i=1}^{m} c_i \frac{\partial v_i}{\partial x} \right] \frac{\partial v_j}{\partial x} + \left[w_y + \sum_{i=1}^{m} c_i \frac{\partial v_i}{\partial y} \right] \frac{\partial v_j}{\partial y} \right] dA$$

$$+ \int_G q \left[w + \sum_{i=1}^{m} c_i v_i \right] v_j + \beta \left[\frac{\partial w}{\partial x} + \sum_{i=1}^{m} c_i \frac{\partial v_i}{\partial x} \right] v_j + \gamma \left[\frac{\partial w}{\partial y} + \sum_{i=1}^{m} c_i \frac{\partial v_i}{\partial y} \right] v_j\, dA$$

$$+ \int_{S_3} p\alpha \left[w + \sum_{i=1}^{m} c_i v_i \right] v_j\, ds = \int_G F\rho v_j dA + \int_{S_2} pv_j g ds + \int_{S_3} pv_j h\, ds, \tag{12.5.6}$$

for $j = 1, \ldots, m$. On solving these equations subject to appropriate initial conditions and substituting the result in the expansion (12.5.5) for $u(x, y, t)$, we obtain an approximate solution to the boundary value problem. This represents *Galerkin's method* for solving the foregoing initial and boundary value problem approximately.

Finite Element Method for Parabolic Equations

The *finite element method* for the (approximate) solution of (12.5.1)–(12.5.3) uses the Galerkin method with the $v_j(x, y)$ chosen to be the finite element basis functions associated with a specific triangulation of the region G constructed as before. As a result, both the solution and the region undergo an approximation. We again distinguish between interior and boundary vertices in the triangulation and between points on the boundary at which a Dirichlet boundary condition is assigned and points that determine boundary segments where Neumann or Robin boundary conditions are given. Each vertex gives rise to a piecewise planar basis function. The approximate solution $u(x, y, t)$ of the given initial and boundary value problem is given as a linear combination of the basis functions. Each basis function that corresponds to a Dirichlet boundary point is multiplied by $f(x, y, t)$ evaluated at that point. Each basis function that corresponds to an interior point, or a Neumann or Robin boundary point, is multiplied by an arbitrary function of t.

Suppose that there are n vertices in the triangulation given as (x_i, y_i), $i = 1, \ldots, n$ of which $n - m$ (with $i = m+1, \ldots, n$) are Dirichlet boundary point vertices. Then we express the approximate solution as

$$u(x, y, t) = \sum_{i=m+1}^{n} f(x_i, y_i, t) v_i(x, y) + \sum_{i=1}^{m} c_i(t) v_i(x, y). \tag{12.5.7}$$

Since $v_i(x_i, y_i) = 1$ and $v_i(x, y) = 0$ at all other vertices, we find that $u(x_j, y_j, t) = f(x_j, y_j, t)$ if (x_j, y_j) is a Dirichlet vertex and $u(x_j, y_j, t) = c_j(t)$ if (x_j, y_j) is any other vertex. At the initial time $t = t_0$, we have $u(x, y, t_0) = r(x, y)$, so that $u(x_j, y_j, t_0) = c_j(t_0) = r(x_j, y_j)$, if (x_j, y_j) is not a Dirichlet vertex. This determines a set of (initial) values $c_i(t_0) = r(x_i, y_i)$. In the system (12.5.6), we set $w(x, y, t) = \sum_{i=m+1}^{n} f(x_i, y_i, t) v_i(x, y)$ and the region G is replaced by the triangulated region $\hat{G}$, while S_2 and S_3 represent the line segments on the polygonal boundary on which Neumann and Robin boundary conditions are prescribed. The $v_i(x, y)$ are piecewise-defined functions composed of plane elements that are continuous and piecewise continuously differentiable over $\hat{G}$. The double integrals over $\hat{G}$ are given as a sum of integrals over the triangles that comprise the triangulation, and the line integrals over S_2 and S_3 are line integrals over appropriate line segments of the polygonal boundary of $\hat{G}$.

The first integral in (12.5.6), $\int_G \rho \left(\partial w/\partial t + \sum_{i=1}^{m} v_i dc_i/dt\right) v_j \, dA$, can be expressed as $d/dt \int_G \rho \left(w + \sum_{i=1}^{m} v_i \, c_i\right) v_j \, dA$. Each basis function $v_i(x, y)$ is a linear function of x and y in each of the triangles, so that the x and y derivatives of $w(x, y, t)$ and the $v_i(x, y)$ are independent of x and y in each triangle. Exact integrations in (12.5.6) of terms that involve the products of the $v_i(x, y)$ and their first derivatives with

$p(x,y), q(x,y), \beta(x,y), \gamma(x,y), \alpha(x,y), g(x,y,t), h(x,y,t)$, and $F(x,y,t)$ may be difficult to carry out. However, if these functions are replaced by their centroid approximations in the double integrals over triangles and by their midpoint approximations in the line integrals over the boundary, the resulting integrals contain polynomial integrands and can all be evaluated exactly. Again, the integration formulas (12.3.10) and (12.3.14) can be invoked for their evaluation. Even though some of the integrands may depend on t, all integrations involve only the x and y variables, so that all the results presented above can be used here.

Thus, to solve an initial and boundary value problem for the parabolic equation (12.5.1) using the finite element method, we must first triangulate the region, determine the basis functions, evaluate the relevant double and line integrals and solve the resulting initial value problem for the system of ODEs for the $c_i(t)$ that determine the approximate solution. As for the elliptic PDE considered above, the process that leads to the system of ODEs is rather lengthy in general. An additional difficulty that occurs here is that a system of linear ODEs must be solved rather than a system of linear equations.

For initial and boundary value problems for parabolic PDEs of the form (12.5.1), we have created a procedure $NumParabolicFEMCM$ that obtains an approximate solution of the problem via the finite element method. The list of triads and vertices for the triangulation of the approximating polygonal region $\hat{G}$ must be entered in the procedure, together with arguments that prescribe the coefficients, inhomogeneous term, and data for initial conditions and for boundary conditions of the first, second, or third kind. The procedure determines centroid and midpoint values as required. The system (12.5.6) is set up, the double and line integrals are evaluated, and the resulting initial value problem for the linear system of ODEs is solved numerically for the unknown $c_i(t)$ in (12.5.7). As stated above, the $c_i(t)$ represent the values of the approximate finite element solution at the corresponding vertex points (x_i, y_i) in the triangulation, at the time t. Combined with prescribed Dirichlet boundary point values, we obtain (approximate) solution values at all vertex points and for a range of values of t. In fact, the output of $NumParabolicFEMCM$ is a procedure that permits the evaluation of the $c_i(t)$ over a range of values of $t > t_0$ as determined by the numerical method used by Maple to solve the system. The finite element approach corresponds to the construction of finite difference solutions for time-dependent problems at a prescribed set of grid points using the method of lines. In the finite element method, the approximate solution (12.5.7) yields a piecewise planar approximation to the exact solution surface at each time t. Additionally, solution values in $\hat{G}$ at points that are not vertices can easily be found from (12.5.7).

In the general case, for the initial and boundary value problem (12.5.1)–(12.5.3), the procedure is $NumParabolicFEMCM(TriadList, VertexList, \rho, p, q, [\beta, \gamma], F, [x,y], t = t_0, r, f, DBV, g, NBV, h, \alpha, TKBV)$. The first two arguments represent the triad list for the triangulation of the polygonal region $\hat{G}$ and the list of interior and boundary vertices. The arguments $t = t_0$ and r prescribe the initial time and initial value. The arguments DBV, NBV, $TKBV$ are lists of boundary vertex points that determine the boundary line segments on which Dirichlet, Neumann, and Robin boundary conditions, respectively, are prescribed. The other arguments are related

in an obvious manner to the terms given in (12.5.1)–(12.5.3). If one of the boundary conditions is absent, the related argument in the procedure is $NONE$, and an empty list [ ] is entered. If there is no Robin boundary condition, $NONE$ is entered twice.

Example 12.18. An Initial and Boundary Value Problem for a Parabolic PDE. We consider an initial and boundary value for the parabolic PDE

$$u_t - (1+x)(u_{xx} + u_{yy}) - u_x + \frac{4u}{1+x} = 3(1+x)^2 yt^2 \qquad (12.5.8)$$

in the unit square $0 < x < 1,\ 0 < y < 1$, with u and its derivatives as functions of (x, y, t). An exact solution of (12.5.8) is $u(x, y, t) = (1+x)^2 yt^3$. We formulate an initial and Dirichlet boundary value problem for (12.5.8) based on this solution. At $t = 0$ we set $u(x, y, 0) = 0$. On the four sides of the square we have $u(0, y, t) = yt^3,\ u(1, y, t) = 4yt^3,\ u(x, 0, t) = 0,\ u(x, 1, t) = (1+x)^2 t^3$.

First, we triangulate the square by using a boundary vertex list that comprises the four corners of the square and then invoking the procedure $PolygonTriang$. This yields the four triangles determined by connecting the centroid $(1/2, 1/2)$ with the four corners of the square. Next, the triangulation is refined using the procedure $RefineTriang$ with one iteration. This results in a triangulation of the square with 16 triangles. There are five interior vertices and eight boundary vertices given as the vertex list, which we denote as VL, [[1/4, 1/4], [3/4, 1/4], [1/2, 1/2], [3/4, 3/4], [1/4, 3/4], [0, 0], [1/2, 0], [1, 0], [1, 1/2], [1, 1], [1/2, 1], [0, 1], [0, 1/2]].

NumParabolicFEMCM$(TL, VL, 1, 1+x, 4/(1+x), [0,0], 3(1+x)^2 yt^2, [x, y], t = 0, 0, (1+x)^2 yt^3, BVL, NONE, [\,], NONE, NONE, [\,])$ is then applied to determine the finite element solution of the problem. The first two arguments are the list of triads TL given by $RefineTriang$ and the vertex list VL given above. The argument BVL is a list composed of the last eight elements of VL. They represent the boundary points in the refined triangulation. The output is $\left[\frac{1}{4}, \frac{1}{4}, c_1(t)\right]$, $\left[\frac{3}{4}, \frac{1}{4}, c_2(t)\right]$, $\left[\frac{1}{2}, \frac{1}{2},\ c_3(t)\right]$, $\left[\frac{3}{4}, \frac{3}{4}, c_4(t)\right]$, $\left[\frac{1}{4}, \frac{3}{4}, c_5(t)\right]$, $[0,0,0]$, $\left[\frac{1}{2}, 0, 0\right]$, $[1,0,0]$, $\left[1, \frac{1}{2}, 2t^3\right]$, $\left[1, 1, 4t^3\right]$, $\left[\frac{1}{2}, 1, \frac{9}{4}t^3\right]$, $\left[0, 1, t^3\right]$, $\left[0, \frac{1}{2}, \frac{1}{2}t^3\right]$. The first five terms in the list represent the solution values at the five interior vertices, with the values of $u(x, y, t)$ at those points given by the $c_i(t),\ i = 1..5$. They are left unspecified because each of them represents a Maple procedure and can be evaluated at a specified time t by using the global variable $SParFECM(t)$. To determine the values at $t = 1$, for example, we enter $SParFECM(1)$ and obtain $[t = 1, c_1(t) = 0.3918, c_2(t) = 0.7667, c_3(t) = 1.135, c_4(t) = 2.297, c_5(t) = 1.174]$. The exact solution values of $u(x, y, 1)$ at the corresponding points are given as $[0.3906, 0.7656, 1.125, 2.297, 1.172]$, so there is fairly good agreement. The global variable $FESolParCM$ gives the piecewise planar finite element representation of the solution as a function of t. We do not exhibit it.

$VSolParTCM(1)$ yields the output of *NumParabolicFEMCM* but with each third argument in the list evaluated at the time $t = 1$, using the values of $SParFECM(1)$ for $c_i(t)$. Then the procedures $P1 = plot3d((1+x)^2 y, x = 0..1, y = 0..1)$, $P2 = SolutionPlot(VList1, VSolParTCM(1))$ and $plots[display](\{P1, P2\})$ generate the plots of the exact and finite element solutions at $t = 1$, as shown in Figure 12.11. These solution surfaces can also be animated.

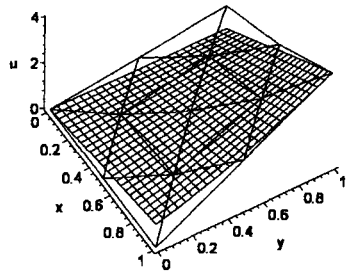

Figure 12.11 Plots of the exact and finite element solutions at $t = 1$.

Exercises 12.5

12.5.1. Derive the Galerkin relation (12.5.4).

12.5.2. Verify that $u(x, y, t) = (1 + x)^2 yt^3$ is the exact solution of the initial and Dirichlet problem formulated for (12.5.8) in Example 12.18.

12.5.3. Carry out all the steps in Example 12.18

12.5.4. Use the finite element method to solve the initial and boundary value problem for the heat equation $u_t(x, y, t) - u_{xx}(x, y, t) - u_{yy}(x, y, t) = -(x+y)e^{-t}$, $u(x, y, 0) = x + y$, $u(0, y, t) = ye^{-t}$, $u(1, y, t) = (1 + y)e^{-t}$, $u(x, 0, t) = xe^{-t}$, $u(x, 1, t) = (1 + x)e^{-t}$. Use $PolygonTriang$ to triangulate the unit square using only the boundary vertices $[0, 0]$, $[1, 0]$, $[1, 1]$, $[0, 1]$. (As a result, there is only one interior vertex.) Apply $NumParabolicFEMCM$ to determine the solution at the interior point as a function of t. Evaluate the solution at various values of t at the interior vertex point and compare with the values of the exact solution, which is given as $u(x, y, t) = (x + y)e^{-t}$.

12.5.5. Reconsider the problem of Exercise 12.5.4 but with the Dirichlet conditions replaced by Neumann conditions. Choose these conditions so that $u(x, y, t) = (x + y)e^{-t}$ remains the exact solution of the problem. Triangulate the unit square as in Exercise 12.5.4 and obtain the finite element solution at all five vertices in the triangulation, for various values of t. Compare with the values of the exact solution.

12.5.6. Reexamine the problem of Exercise 12.5.4 but with the Dirichlet conditions replaced by Robin conditions with the parameter $h = 1$. Choose these conditions so that $u(x, y, t) = (x + y)e^{-t}$ remains the exact solution of the problem. Triangulate the unit square as in Exercise 12.5.4 and obtain the finite element solution at all five vertices in the triangulation, for various values of t. Compare with the values of the exact solution.

12.6 FINITE ELEMENT SOLUTIONS FOR HYPERBOLIC EQUATIONS IN TWO DIMENSIONS

Galerkin Integrals for Hyperbolic Equations

We consider the hyperbolic PDE in the two-dimensional spatial region G

$$\rho u_{tt} - (pu_x)_x - (pu_y)_y + qu + \beta u_x + \gamma u_y = \rho F, \tag{12.6.1}$$

where u, F are functions of (x, y, t) and $\rho, p, q, \beta, \gamma$ are functions of (x, y), with $(x, y) \in G$ and $t > t_0$. It is assumed that $p(x, y) > 0$ and $\rho(x, y) > 0$ in G. If $\beta(x, y) = 0$, $\gamma(x, y) = 0$, $p(x, y) = c^2$, $\rho(x, y) = 1$, and $q(x, y) = 0$, (12.6.1) is the nonhomogeneous *wave equation*. The *initial conditions*

$$u(x, y, t_0) = r_1(x, y), \quad u_t(x, y, t_0) = r_2(x, y), \qquad (x, y) \in G, \tag{12.6.2}$$

are prescribed at $t = t_0$. *Dirichlet, Neumann, Robin*, or *mixed boundary conditions* that combine the three types of boundary conditions, are prescribed for $u(x, y, t)$ on the boundary S of the region G for all $t > t_0$. As in Section 4.1, the boundary, for each $t \geq t_0$, is comprised of the union of the arcs S_1, S_2, and S_3, with Dirichlet conditions given on S_1, Neumann conditions on S_2, and Robin boundary conditions on S_3. Either S_1, S_2, or S_3 or combinations of two of them may be empty. The boundary values are given as

$$u(x, y, t)\big|_{S_1} = f, \; u_n(x, y, t)\big|_{S_2} = g, \; u_n(x, y, t) + \alpha(x, y)u(x, y, t)\big|_{S_3} = h, \tag{12.6.3}$$

where f, g, h are functions of (x, y, t).

To determine the *Galerkin integral relation* for the problem, we multiply across in (12.6.1) by an arbitrary *admissible function* $v(x, y)$, apply Green's theorem in the plane, and use the (mixed) boundary conditions given above. It is again required that all admissible $v(x, y)$ satisfy the condition $v(x, y)\big|_{S_1} = 0$, if S_1 nonempty, but their values are unrestricted on S_2 and S_3. As a result, the line integral over S_1 that arises in the divergence theorem must vanish. This yields the following Galerkin relation for the mixed problem for the hyperbolic equation (12.6.1):

$$\int_G \rho v u_{tt} + p(u_x v_x + u_y v_y) + quv + v(\beta u_x + \gamma u_y)\, dA + \int_{S_3} p\alpha uv\, ds$$

$$= \int_G v\rho F\, dA + \int_{S_2} pvg\, ds + \int_{S_3} pvh\, ds. \tag{12.6.4}$$

Now, if $u(x, y, t)$ is a solution of the given initial and boundary value problem (12.6.1)–(12.6.3), the Galerkin relation (12.6.4) is satisfied for all admissible $v(x, y)$. We again reverse the process and use the fact that if the Galerkin relation is satisfied for all admissible $v(x, y)$, the original initial and boundary value problem is satisfied by $u(x, y, t)$, in view of the arbitrariness of $v(x, y)$. The Galerkin relation can be satisfied

even if $u(x, y, t)$ and $v(x, y)$ have only piecewise continuous first derivatives in x and y. Thereby the concept of a solution to a given initial and boundary value problem is extended and weakened. As before, rather than constructing a weak solution and showing that it satisfies the Galerkin relation, we use this approach to construct approximate weak solution for the given initial and boundary value problem using the *Galerkin method.*

The *Galerkin method* for the initial and boundary value problem (12.6.1)–(12.6.3) proceeds as for the initial and boundary value problem of Section 12.5. We do not repeat all the details but note that, as before, the (approximate) solution takes the form

$$u(x, y, t) = w(x, y, t) + \sum_{i=1}^{m} c_i(t) v_i(x, y), \tag{12.6.5}$$

with arbitrary coefficients $c_i(t)$ that are to be determined. This expansion is inserted into the Galerkin relation (12.6.4), where we consecutively set $v(x, y) = v_j(x, y)$, $j = 1, \ldots, m$. This yields the following system of m simultaneous second order ODEs for the m functions $c_i(t)$:

$$\int_G \rho \left[w_{tt} + \sum_{i=1}^{m} c_i'' v_i \right] v_j + p \left[\left[w_x + \sum_{i=1}^{m} c_i \frac{\partial v_i}{\partial x} \right] \frac{\partial v_j}{\partial x} + \left[w_y + \sum_{i=1}^{m} c_i \frac{\partial v_i}{\partial y} \right] \frac{\partial v_j}{\partial y} \right] dA$$

$$+ \int_G q \left[w + \sum_{i=1}^{m} c_i v_i \right] v_j + \beta \left[\frac{\partial w}{\partial x} + \sum_{i=1}^{m} c_i \frac{\partial v_i}{\partial x} \right] v_j + \gamma \left[\frac{\partial w}{\partial y} + \sum_{i=1}^{m} c_i \frac{\partial v_i}{\partial y} \right] v_j \, dA$$

$$+ \int_{S_3} p\alpha \left[w + \sum_{i=1}^{m} c_i v_i \right] v_j \, ds = \int_G F \rho v_j \, dA + \int_{S_2} p v_j g \, ds + \int_{S_3} p v_j h \, ds, \tag{12.6.6}$$

for $j = 1, \ldots, m$. On solving these equations subject to appropriate initial conditions and substituting the result in the expansion (12.6.5) for $u(x, y, t)$, we obtain an approximate solution to the boundary value problem. This represents *Galerkin's method* for solving the foregoing initial and boundary value problem approximately.

Finite Element Method for Hyperbolic Equations

The *finite element method* for the (approximate) solution of (12.6.1)–(12.6.3) uses the Galerkin method with the $v_j(x, y)$ chosen to be the finite element basis functions associated with a specific triangulation of the region G constructed as before. The discussion parallels that given above for parabolic equations and will not be repeated, except for the introduction of some modifications. The approximate solution is given as

$$u(x, y, t) = \sum_{i=m+1}^{n} f(x_i, y_i, t) v_i(x, y) + \sum_{i=1}^{m} c_i(t) v_i(x, y). \tag{12.6.7}$$

At the initial time $t = t_0$, we have $u(x, y, t_0) = r_1(x, y)$ and $u_t(x, y, t_0) = r_2(x, y)$, so that $u(x_j, y_j, t_0) = c_j(t_0) = r_1(x_j, y_j)$ and $u_t(x_j, y_j, t_0) = c_j'(t_0) = r_2(x_j, y_j)$

if (x_j, y_j) is not a Dirichlet vertex. This determines the set of initial values $c_i(t_0)$ and $c_i'(t_0)$. In the system (12.6.6), we set $w(x, y, t) = \sum_{i=m+1}^{n} f(x_i, y_i, t)v_i(x, y)$, and the region G is replaced by the triangulated region $\hat{G}$, while S_2 and S_3 represent the line segments on the polygonal boundary on which Neumann and Robin boundary conditions are prescribed.

The first integral in (12.6.6), $\int_G \rho \left(\partial^2 w/\partial t^2 + \sum_{i=1}^{m} v_i c_i''\right) v_j \, dA$, can be expressed as $d^2/dt^2 \int_G \rho \left(w + \sum_{i=1}^{m} v_i c_i\right) v_j \, dA$. Each basis function $v_i(x, y)$ is a linear function of x and y in each of the triangles, so that the x and y derivatives of $w(x, y, t)$ and the $v_i(x, y)$ are independent of x and y in each triangle. Exact integrations in (12.6.6) of terms that involve the products of the $v_i(x, y)$ and their first derivatives with $p(x, y), q(x, y), \beta(x, y), \gamma(x, y), \alpha(x, y), g(x, y, t), h(x, y, t)$, and $F(x, y, t)$ may be difficult to carry out. If these functions are again replaced by their centroid approximations in the double integrals over triangles and by their midpoint approximations in the line integrals over the boundary, the resulting integrals contain polynomial integrands and can all be evaluated exactly. The integration formulas (12.3.10) and (12.3.14) can be invoked for their evaluation. Even though some of the integrands may depend on t, all integrations involve only the x and y variables, so that all the results presented above can be used here.

Thus, to solve an initial and boundary value problem for the hyperbolic equation (12.6.1) using the finite element method, we must first triangulate the region, determine the basis functions, evaluate the relevant double and line integral and solve the resulting initial value problem for the system of ODEs for the $c_i(t)$ that determine the approximate solution. Everything proceeds as for the parabolic PDE considered above, except that the system of linear ODEs is of second order rather than first order.

For an initial and boundary value problem for a hyperbolic PDE of the form (12.6.1), we have created a procedure $NumHyperbolicFEMCM$ that obtains an approximate solution of the problem via the finite element method. The list of triads and vertices for the triangulation of the approximating polygonal region $\hat{G}$ must be entered in the procedure, together with arguments that prescribe the coefficients, inhomogeneous term, and data for initial conditions and for boundary conditions of the first, second, or third kind. The procedure determines centroid and midpoint values as required. The system (12.6.6) is set up, the double and line integrals are evaluated, and the resulting initial value problem for the linear system of ODEs is solved numerically for the unknown $c_i(t)$ in (12.6.7). The $c_i(t)$ represent the values of the approximate finite element solution at the corresponding vertex points (x_i, y_i) in the triangulation, at the time t. Combined with prescribed Dirichlet boundary point values, we obtain (approximate) solution values at all vertex points and for a range of values of t. The output of $NumHyperbolicFEMCM$ is a procedure that permits the evaluation of the $c_i(t)$ over a range of values of $t > t_0$ as determined by the numerical method used by Maple to solve the system. The foregoing finite element approach corresponds to the construction of finite difference solutions for time-dependent problems at a prescribed set of grid points using the method of lines. In the finite element method, the approximate solution (12.6.7) yields a piecewise planar approximation to the exact

solution surface at each time t. Additionally, solution values in $\hat{G}$ at points that are not vertices can easily be found from (12.6.7).

In the general case that corresponds to problem (12.6.1)–(12.6.3), we have *Num HyperbolicFEMCM(TriadList,VertexList,* $\rho, p, q, [\beta, \gamma], F, [x, y], t = t_0, r_1, r_2, f, DBV, g, NBV, h, \alpha, TKBV$). The meanings of the arguments in the procedure (and their modification for other boundary conditions) are the same as for the procedure $NumParabolicFEMCM$ of Section 12.5, with the exception that two initial values r_1 and r_2 must be entered.

Example 12.19. An Initial and Boundary Value Problem for a Hyperbolic PDE. We consider an initial and boundary value problem for the hyperbolic PDE

$$u_{tt} - (1+x)(u_{xx} + u_{yy}) - u_x + \frac{4u}{1+x} = 6(1+x)^2 yt \tag{12.6.8}$$

in the (right) triangle with vertices $(0,0), (1,0), (0,1)$, with u and its derivatives given as functions of (x, y, t). The initial conditions are

$$u(x, y, 0) = 0, \quad u_t(x, y, 0) = 0. \tag{12.6.9}$$

Let the three sides of the triangle $x+y=1$, $y=0$, $x=0$ be denoted by S_1, S_2, S_3, respectively. The mixed boundary conditions are

$$u\big|_{S_1} = (1+x)^2 yt^3,\ u_n\big|_{S_2} = -(1+x)^2 t^3,\ u_n + u\big|_{S_3} = -yt^3 \tag{12.6.10}$$

with u and its derivatives given as functions of (x, y, t). This problem has the exact solution $u(x, y, t) = (1+x)^2 yt^3$.

We triangulate the triangle using a boundary vertex list $[[0,0], [1,0], [0,1]]$ in the procedure $PolygonTriang$. This yields the three triangles determined by connecting the centroid $(1/3, 1/3)$ with the three vertices of the triangle. Next, the triangulation is refined using the procedure *RefineTriang* with one iteration. This results in a modified triangulation with twelve triangles. There are four interior vertices and six boundary vertices given as VL, $[\frac{1}{6}, \frac{1}{6}], [\frac{2}{3}, \frac{1}{6}], [\frac{1}{3}, \frac{1}{3}], [\frac{1}{6}, \frac{2}{3}], [0,0], [\frac{1}{2}, 0], [1,0], [\frac{1}{2}, \frac{1}{2}], [0,1], [0, \frac{1}{2}]$.

NumHyperbolicFEMCM($TL, VL, 1, 1+x, 4/(1+x), [0,0], 6(1+x)^2 yt, [x, y], t = 0, 0, 0, (1+x)^2 yt^3, BVL[3..5], -(1+x)^2 t^3, BVL[1..3], -yt^3, 1, [op(BVL[5..6]), BVL[1]]$) determines the finite element solution of the problem. The first two arguments are the list of triads TL given by *RefineTriang* and the vertex list VL. The argument BVL is a list composed of the last six elements of VL, which represents the boundary points in the refined triangulation. The terms $BVL[i..j]$ represent a list of boundary vertices extending from the ith to the jth vertex. The remaining arguments determine the coefficients, data and inhomogeneous terms of the PDE. The output of the procedure is $\left[\frac{1}{6}, \frac{1}{6}, c_1(t)\right], \left[\frac{2}{3}, \frac{1}{6}, c_2(t)\right], \left[\frac{1}{3}, \frac{1}{3}, c_3(t)\right], \left[\frac{1}{6}, \frac{2}{3}, c_4(t)\right], [0, 0, c_5(t)], \left[\frac{1}{2}, 0, c_6(t)\right], [1, 0, 0], \left[\frac{1}{2}, \frac{1}{2}, \frac{9t^3}{8}\right], [0, 1, t^3], \left[0, \frac{1}{2}, c_{10}(t)\right]$. The first four terms in the list represent the solution values at the four interior vertices.

The values of $u(x, y, t)$ at those points are given by the $c_i(t)$, $i = 1..4$. The terms $c_5(t)$, $c_6(t)$, $c_{10}(t)$ represent solution values at vertices on the boundaries where Neumann and Robin boundary values are assigned. They are left unspecified because each of them represents a Maple procedure and can be evaluated at a specified time t by using the global variable *SHypFECM(t)*.

To determine the values at $t = 0.5$, we enter *SHypFECM(0.5)* and obtain

$$[t = 0.5, c_1(t) = 0.02856, c_2(t) = 0.05569, c_3(t) = 0.06630, c_4(t) = 0.1164,$$
$$c_5(t) = -0.003699, c_6(t) = 0.01650, c_{10}(t) = 0.06496]. \qquad (12.6.11)$$

The derivatives $c_i'(0.5)$ are also approximated and displayed by the procedure, but we omit these values. The exact solution values of $u(x, y, 0.5)$ at the corresponding points are given as $[0.02836, 0.05787, 0.07407, 0.1134, 0, 0, 0.06250]$. We observe that the results do not agree as nicely as in the parabolic problem of Example 12.18, where a larger value of t was used. The hyperbolic problem does not have the smoothing properties of the parabolic problem. The global variable $FESolHypCM$ gives the piecewise planar finite element representation of the solution as a function of t.

VSolHypTCM(0.5) yields the output of *NumHyperbolicFEMCM* but each third argument in the list evaluated at the time $t = 0.5$, using the values displayed in (12.6.11) for the $c_i(t)$. Finally, the procedures $P1 = plot3d((1+x)^2 y(.5)^3, y = 0..1-x, x = 0..1)$, $P2 = SolutionPlot(TL, VSolHypTCM(0.5))$, and $plots[display](\{P1, P2\})$ yield the plots of the exact and finite element solutions at $t = 0.5$. They are not displayed. The solution surfaces can also be animated.

Exercises 12.6

12.6.1. Derive the Galerkin relation (12.6.4).

12.6.2. Reproduce the results of Example 12.19.

12.6.3. Replace the heat equation of Exercise 12.5.4 by the nonhomogeneous wave equation $u_{tt}(x, y, t) - u_{xx}(x, y, t) - u_{yy}(x, y, t) = (x + y)e^{-t}$, and add the initial condition $u_t(x, y, 0) = -x - y$, while retaining the initial and boundary conditions given in that exercise. This problem again has the exact solution $u(x, y, t) = (x + y)e^{-t}$. Proceed as in that exercise and obtain the finite element solution of foregoing problem by using the procedure $NumHyperbolicFEMCM$.

12.6.4. Solve the problem of Exercise 12.5.5, replacing the heat equation by the wave equation of Exercise 12.6.3.

12.6.5. Solve the problem of Exercise 12.5.6, replacing the heat equation by the wave equation of Exercise 12.6.3.

12.7 FINITE ELEMENT SOLUTIONS FOR PDE EIGENVALUE PROBLEMS IN TWO DIMENSIONS

Galerkin Integrals for PDE Eigenvalue Problems

We consider the eigenvalue problem for an elliptic PDE in the two-dimensional region jG

$$-(p(x,y)u_x(x,y))_x - (p(x,y)u_y(x,y))_y + q(x,y)u(x,y) = \lambda\rho(x,y)u(x,y), \tag{12.7.1}$$

with λ as the eigenvalue parameter, and $p(x,y) > 0$ and $\rho(x,y) > 0$ in G. [The PDE has the formally self-adjoint form $-\nabla \cdot (p(x,y)\nabla u(x,y)) + q(x,y)u(x,y) = \lambda\rho(x,y)u(x,y)$ and corresponds to the two-dimensional version of (8.1.1).] Homogeneous *Dirichlet, Neumann, Robin,* or *mixed boundary conditions* are prescribed for $u(x,y)$ on the boundary S of the region G. As in Section 4.1, the boundary is comprised of the union of the arcs S_1, S_2 and S_3, with Dirichlet conditions given on S_1, Neumann conditions on S_2, and Robin boundary conditions on S_3. Either S_1, S_2, or S_3 or combinations of two of them may be empty. The boundary values are given as

$$u(x,y)\big|_{S_1} = 0,\ u_n(x,y)\big|_{S_2} = 0,\ u_n(x,y) + \alpha(x,y)u(x,y)\big|_{S_3} = 0. \tag{12.7.2}$$

To determine the *Galerkin integral relation* for the problem, we multiply across in (12.7.1) by an arbitrary *admissible function* $v(x,y)$ and apply Green's theorem in the plane and the (mixed) boundary conditions. It is required that all admissible $v(x,y)$ satisfy the condition $v(x,y)\big|_{S_1} = 0$, if S_1 nonempty, but their values are unrestricted on S_2 and S_3. As a result, the line integral over S_1 that arises in the divergence theorem must vanish. This yields the following Galerkin relation for the eigenvalue problem (12.7.1)–(12.7.2):

$$\int_G p(u_x v_x + u_y v_y) + quv\, dA + \int_{S_3} p\alpha uv\, ds = \lambda \int_G \rho vu\, dA. \tag{12.7.3}$$

Each of the functions in the integrals depends on (x,y).

Now, if λ is an eigenvalue and $u(x,y)$ is a (corresponding) eigenfunction for the eigenvalue problem (12.7.1)–(12.7.2), the Galerkin relation (12.7.3) is satisfied for all admissible $v(x,y)$. We again reverse the process and use the fact that if the Galerkin relation is satisfied for all admissible $v(x,y)$, then λ is an eigenvalue and $u(x,y)$ is an eigenfunction for the problem, owing to the arbitrariness of $v(x,y)$. However, the Galerkin relation can be satisfied even if the eigenfunction $u(x,y)$ is a weak solution that has only piecewise continuous first derivatives. We use this approach to construct approximate eigenvalues and (weak) eigenfunctions for the eigenvalue problem using the *Galerkin method.*

The *Galerkin method* determines approximate eigenvalues and eigenfunctions as follows. We select a set of m linearly independent admissible basis functions $v_j(x,y)$, $j = 1,\ldots,m$. As such, each $v_j(x,y)$ must vanish on the portion S_1 of

the boundary of G but is otherwise arbitrary. Then we represent the (approximate) solution as $u(x,y) = \sum_{i=1}^{m} c_i v_i(x,y)$, with arbitrary constant coefficients c_i that are to be determined. This expansion is inserted into the Galerkin relation (12.7.3), where we consecutively set $v(x,y) = v_j(x,y)$, $j = 1,\ldots,m$. This yields the m linear simultaneous equations for the m constants c_i,

$$\sum_{i=1}^{m} c_i \left[\int_G p\nabla v_i \cdot \nabla v_j + q v_i v_j \, dA + \int_{S_3} p\alpha v_i v_j \, ds \right] = \lambda \sum_{i=1}^{m} c_i \int_G \rho v_i v_j \, dA, \tag{12.7.4}$$

for $j = 1,\ldots,m$. The foregoing system corresponds to a *generalized matrix eigenvalue problem* of the form $K\mathbf{c} = \lambda M\mathbf{c}$ for the $m \times m$ matrices K and M whose elements K_{ij} and M_{ij} are given as

$$K_{ij} = \int_G p\nabla v_i \cdot \nabla v_j + q v_i v_j \, dA + \int_{S_3} p\alpha v_i v_j \, ds, \tag{12.7.5}$$

$$M_{ij} = \int_G \rho v_i v_j \, dA, \tag{12.7.6}$$

with λ as the eigenvalue and the column vector $\mathbf{c} = [c_1, c_2, \ldots, c_m]^T$ as the eigenvector. [In engineering literature terminology, if $q = \alpha = 0$ in (12.7.4), K is referred to as the *stiffness matrix*. Additionally, if ρ is a density, M is the *mass matrix*.] The matrices K and M are symmetric and M is positive definite. They correspond to the matrices A and B defined in (8.2.3)–(8.2.4). These matrix eigenvalue problems were discussed in Section 8.2.

Each eigenvalue λ of $K\mathbf{c} = \lambda M\mathbf{c}$ yields an *approximate eigenvalue* for the PDE eigenvalue problem (12.7.1)–(12.7.2). On substituting the components c_i of the corresponding algebraic eigenvector into $u(x,y) = \sum_{i=1}^{m} c_i v_i(x,y)$, we obtain an *approximate eigenfunction*. This represents *Galerkin's method* for solving the PDE eigenvalue problem approximately. It yields m approximate eigenvalues (counted with their multiplicities) and m approximate eigenfunctions. By increasing the number of basis functions, additional approximate eigenvalues and eigenfunctions can be found.

Finite Element Method for the PDE Eigenvalue Problem

The *finite element method* for the (approximate) solution of the eigenvalue problem (12.7.1)–(12.7.2) uses the Galerkin method with the $v_j(x,y)$ chosen to be the finite element basis functions associated with a specific triangulation of the region G. (As a result, not only the solution but also the region undergo an approximation, in general.) The process of setting up the matrix eigenvalue problem parallels that for the construction of the finite element solution of an elliptic PDE as presented in Section 12.4. The approximate eigenfunction $u(x,y)$ is represented as a linear combination of the basis functions for the triangulation. Since the eigenfunction must vanish at all Dirichlet boundary points, only basis functions that correspond to an interior point or a Neumann or Robin condition boundary point occur in the representation.

We suppose that there are n vertices in the triangulation of the region, given as (x_i, y_i), $i = 1, \ldots, n$ of which $n-m$ (with $i = m+1, \ldots, n$) are Dirichlet boundary point vertices. Then we express the approximate eigenfunction as

$$u(x,y) = \sum_{i=1}^{m} c_i v_i(x,y), \tag{12.7.7}$$

with the constant coefficients c_i to be determined. Since $v_i(x_i, y_i) = 1$ and $v_i(x, y) = 0$ at all other vertices, we find that $u(x_j, y_j) = 0$ if (x_j, y_j) is a Dirichlet vertex and $u(x_j, y_j) = c_j$ if (x_j, y_j) is any other vertex. In the resulting system (12.7.4), the region G is replaced by the triangulated region $\hat{G}$, while S_3 represents the line segments on the polygonal boundary on which Robin boundary conditions are prescribed. The $v_i(x, y)$ are piecewise-defined functions composed of plane elements that are continuous and piecewise continuously differentiable over $\hat{G}$. The double integrals over $\hat{G}$ are given as a sum of integrals over the triangles that comprise the triangulation, and the line integral over S_3 is a line integral over appropriate line segments of the polygonal boundary of $\hat{G}$.

Again, each basis function $v_i(x, y)$ is a linear function of x and y in each of the triangles, so that the x and y derivatives of the $v_i(x, y)$ are constants in each triangle. Exact integrations in (12.7.5) of terms that involve the products of the $v_i(x, y)$ and their first derivatives with $p(x, y)$, $q(x, y)$, $\rho(x, y)$, and $\alpha(x, y)$ may be difficult to carry out. If these functions are replaced by their centroid approximations in the double integrals over triangles and by their midpoint approximations in the line integrals over the boundary, the resulting integrals contain polynomial integrands and can be evaluated exactly, using the integration formulas (12.3.10) and (12.3.14).

Thus, to solve eigenvalue problem (12.7.1)–(12.7.2) using the finite element method, we must first triangulate the region, determine the basis functions, and evaluate the relevant double and line integrals. (The process of evaluating and combining the values of the integrals over the triangles and boundary segments of the triangulation is referred to as *assembling* in the engineering literature.) The resulting linear system represents a matrix eigenvalue problem that yields approximate eigenvalues and eigenfunctions. This is a rather lengthy process in general.

The procedure $NumEigenvalFEMCM$ automates the triangulation of a region, determines the necessary basis functions, and carries out additional steps required to set up the system (12.7.5). Then built-in Maple procedures for the determination of algebraic eigenvalues and eigenvectors are used to complete the finite element solution of the PDE eigenvalue problem. The dimension of the matrices in the algebraic eigenvalue problem equals the number of non-Dirichlet vertices in the triangulation. Typically, large matrices can occur and it may be of interest to determine only the leading eigenvalues for the problem. There are approximate methods for doing so.

For the eigenvalue problem with the mixed boundary conditions (12.7.2), the procedure has the form $NumEigenvalFEMCM(TriadList, VertexList, p, q, [x, y], \lambda, 0, DBV, 0, NBV, 0, \alpha, TKBV)$. The first two arguments represent the triad list for the triangulation of the polygonal region $\hat{G}$ and the list of interior and boundary vertices. The arguments DBV, NBV, and $TKBV$, are lists of boundary vertex

points that determine the boundary line segments on which Dirichlet, Neumann, and Robin boundary conditions, respectively, are prescribed. The three arguments 0 in the procedure signify that only zero Dirichlet, Neumann, or Robin boundary conditions can be specified. If one of the boundary conditions is absent, the corresponding argument in the procedure is $NONE$ and an empty list of points [ ] is entered. If there is no Robin boundary condition, $NONE$ is entered twice. The remaining arguments relate in an obvious way to the coefficients in (12.7.1).

In the following example we apply the finite element method to determine the (approximate) leading eigenvalues and eigenfunctions for the eigenvalue problem for the Laplacian operator in the unit square with Neumann boundary conditions. A simple triangulation of the square is introduced, and various aspects of the solution method are brought out. Problems that involve other equations, boundary conditions, and triangulations can be treated similarly.

Example 12.20. An Eigenvalue Problem in the Unit Square. We apply the finite element method to the eigenvalue problem for the PDE

$$-u_{xx}(x,y) - u_{yy}(x,y) = \lambda u(x,y) \tag{12.7.8}$$

in the unit square $0 < x < 1$, $0 < y < 1$ with homogeneous Neumann conditions $u_n(x,y) = 0$ assigned on the boundary. The exact eigenvalues for this problem are $\lambda_{nm} = \pi^2(n^2+m^2)$, $n, m = 0, 1, 2, \ldots$, and a corresponding set of (unnormalized) eigenfunctions is $u_{nm}(x,y) = \cos(\pi nx)\cos(\pi my)$, $n, m = 0, 1, 2, \ldots$. As has been shown previously, $\lambda = 0$ is an eigenvalue, with $u(x,y) = 1$ as a corresponding eigenfunction. The eigenfunctions $\cos(\pi x)$ and $\cos(\pi y)$ correspond to the double eigenvalue $\lambda = \pi^2$, and so on.

We proceed as in Example 12.18 and triangulate the square using the boundary vertex list BSL =[0,0],[1,0],[1,1],[0,1] and $PolygonTriang$. This yields the four triangles obtained by connecting the centroid $(1/2, 1/2)$ with the four corners of the square. The triad list is VSL =[[0, 0], [1, 0], [.5, .5], [1, 0], [1, 1], [.5, .5], [1, 1], [0, 1], [.5, .5], [0, 1], [0, 0], [.5, .5]]. $Vertexlist(VSL, BSL)$ determines the list of vertices $VVSL$ =[[0.5,0.5],[0, 0], [1, 0], [1, 1], [0, 1]], of which the first is an interior vertex and the remaining four are boundary vertices. As we are dealing with a Neumann problem, (12.7.7) takes the form $u(x,y) = \sum_{i=1}^{5} c_i v_i(x,y)$. Consequently, K and M for this problem are both 5×5 matrices. Five eigenvalues will be found by the finite element method.

On using $NumEigenvalFEMCM(VSL, VVSL, 1, 0, 1, [x,y], \lambda, NONE, [\ ], 0, [BSL, [0,0]], NONE, NONE, [\ ])$ we obtain the output

$$\lambda_1 = -0.3785 \times 10^{-14}, \lambda_2 = 12.0, \lambda_3 = 12.0, \lambda_4 = 24.0, \lambda_5 = 72.09. \tag{12.7.9}$$

The first eigenvalue λ_1 clearly corresponds to exact eigenvalue $\lambda = 0$ for the problem. The double eigenvalue $\lambda_2 = \lambda_3 = 12$ approximates the exact double eigenvalue $\lambda = \pi^2 \approx 9.872$, while $\lambda_4 = 24$ approximates the exact eigenvalue $\lambda = 2\pi^2 \approx 19.74$. It is seen that the approximations to the higher eigenvalues undergo a rapid decrease in

accuracy. (This is consistent with the results given by the Rayleigh-Ritz method.) A refinement of the triangulation increases the number of basis functions and eigenvalues that are determined. Then the leading algebraic eigenvalues approximate the exact eigenvalues more closely.

For example, a double refinement of the foregoing triangulation obtained from the procedure $RefineTriang$ yields 25 interior vertices and 16 boundary vertices. As a result, 41 eigenvalues are found by $NumEigenvalFEMCM$. The first four eigenvalues are

$$\lambda_1 = -0.485 \times 10^{-13},\ \lambda_2 = 10.135,\ \lambda_3 = 10.135,\ \lambda_4 = 20.714. \quad (12.7.10)$$

The improvement in the approximation to the exact eigenvalues is apparent.

Next, we examine the eigenvectors determined by $NumEigenvalFEMCM$ for the initial triangulation with five vertices. The global variable $EVecList$ lists the eigenvectors determined by Maple for each of the five eigenvalues obtained above. For technical reasons, the eigenvectors are not listed in the same order as the eigenvalues in (12.7.9), but each eigenvector for the eigenvalue λ_i is identified as $EVec_{\lambda_i}$. For example, we obtain

$$EVec_{\lambda_1} = [-0.4474,\ -0.4474,\ -0.4474,\ -0.4474,\ -0.4474]. \quad (12.7.11)$$

This means that the eigenvector for $\lambda_1 = -0.3785 \times 10^{-14}$ (i.e., the zero eigenvalue) is a constant multiple of the vector $[1, 1, 1, 1, 1]$, which we choose as the eigenvector. The components of the eigenvector determine the values of the c_i in (12.7.7). They all equal 1, so that the approximate eigenfunction is just $u(x, y) = 1$. This is an exact eigenfunction for the PDE eigenvalue problem.

For the eigenvector that corresponds to $\lambda_4 = 24.0$, we have

$$EVec_{\lambda_4} = [-0.44491 \times 10^{-16},\ -0.50,\ 0.50,\ -0.50,\ 0.50]. \quad (12.7.12)$$

As a result, we can effectively set $c_1 = 0$ and $c_i = 0.5(-1)^{i-1}$, $i = 2, 3, 4, 5$, in (12.7.7). The use of the procedures $PlaneElementslist$ and FEF determines the finite element piecewise planar representation of the eigenfunction to be

$$EFunc_4 = [x - 0.5,\ 0.5 - y,\ 0.5 - x,\ y - 0.5]. \quad (12.7.13)$$

It has the values $[0, -0.5, 0.5, -0.5, 0.5]$ at the vertices $[[.5, .5], [0, 0], [1, 0], [1, 1], [0, 1]]$, respectively. The eigenfunction (12.7.13) is an approximation to the exact eigenfunction $u_{11}(x, y) = -0.5\cos(\pi x)\cos(\pi y)$, whose exact eigenvalue is $2\pi^2$. The values of the approximate and exact eigenfunctions agree at the five vertices of the triangulation.

The output of the global variable $VSol$ of $NumEigenvalFEMCM$ is the list $[[x_1, y_1, c_1], \ldots, [x_n, y_n, c_n]]$, where the (x_i, y_i), $i = 1, \ldots, n$, are the coordinates of the vertices in the triangulation of the region. In addition, the global variable EV lists the eigenvalues, their multiplicity, and the eigenvectors of the related algebraic

eigenvalue problem. Then the procedure $EFPlot(VSL, VSol, i, EV)$ plots the finite element eigenfunction determined by the ith eigenvalue and eigenvector. For our triangulation we have $n = 5$ and we put $i = 4$. The eigenvalue is $\lambda_4 = 24$ and the eigenvector is given as in (12.7.12). The corresponding exact eigenvalue is $2\pi^2$, and the eigenfunction is $u_{11}(x, y)$ as defined above. The plot of $u_{11}(x, y)$ is given by $plot3d(u_{11}(x, y), x = 0..1, y = 0..1)$, and we denote it as $P1$. With the plot of the finite element eigenfunction denoted by $P2$, $plots[display](\{P1, P2\})$ yields the plots of the exact and finite element eigenfunctions, as in Figure 12.12.

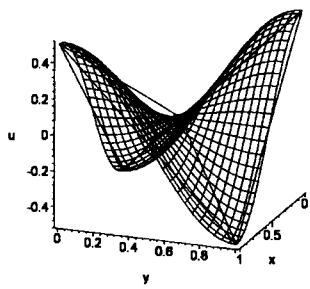

Figure 12.12 Plots of the exact and finite element eigenfunction.

Exercises 12.7

12.7.1. Derive the Galerkin relation (12.7.3).

12.7.2. Reproduce the results of Example 12.20.

12.7.3. Reconsider Example 12.20 if the Neumann conditions are replaced by homogeneous Dirichlet conditions.

12.7.4. Reconsider Example 12.20 if the Neumann conditions are replaced by homogeneous Robin conditions with the parameter $h = 1$.

12.7.5. Approximate the leading eigenvalue (see Example 8.2) for the Dirichlet eigenvalue problem for Laplace's equation in the unit circle $x^2 + y^2 = 1$ via the finite element method. Introduce the 100 boundary vertices $[\cos(2\pi i/n), \sin(2\pi i/n)]$, $i = 0, \ldots, 99$, and use $PolygonTriang$ as in Example 12.1. This yields an interior point at the center of the circle. Apply the procedure $NumEigenvalFEMCM$ and obtain an approximation to the leading eigenvalue.

12.7.6. The points $[0, 0], [1/2, 0], [1, 0], [1, 1/2], [1, 1], [1/2, 1/2]$ lie on the boundary of the right triangle with vertices $[[0, 0], [1, 0], [1, 1]]$. Use *PolygonTriang* to obtain an initial triangulation of the triangle and obtain a single and then a double refinement of the triangulation by applyng the procedure *RefineTriang*. Use these triangulations and $NumEigenvalFEMCM$ to obtain a finite element approximation to the leading eigenvalue for the Dirichlet eigenvalue problem for Laplace's equation in the triangle.

BIBLIOGRAPHY

CHAPTER 1

Barber, M. N., and B. W. Ninham, *Random and Restricted Walks: Theory and Applications*, Gordon and Breach, New York, 1970.

Berg, H. C., *Random Walks in Biology*, Princeton University Press, Princeton, N.J., 1983.

Courant, R., K. O. Friedrichs, and H. Lewy, Über die partiellen differenzengleichungen der mathematischen physik, *Math. Ann.*, 100, 1928, 32–45. (English translation in *IBM J. Res. Develop.*, 11, 1967, 217–234.)

Doyle, P. G., and J. L. Snell, *Random Walks and Electric Networks*, Carus Math. Monogr., No. 22, Mathematical Association of America, Washington, D.C., 1984.

Einstein, A., *Investigations on the Theory of the Brownian Movement*, R. Fürth, Ed., Dover, New York, 1952.

Feller, W., *An Introduction to Probability Theory and Its Applications*, Vol. 1, 3rd ed., Wiley, New York, 1968.

Fürth, R., Die Brownsche bewegung . . . , *Z. Phys.*, 2, 1920, 244–256.

Gardiner, C. W., *Handbook of Stochastic Methods*, Springer, Berlin, 1983.

Goldstein, S., On diffusion by discontinuous movements and the telegraph equation, *Q. J. Mech. Appl. Math.*, 4, 1951, 129–156.

Horsthemke, W., and R. Lefever, *Noise-Induced Transitions*, Springer, Berlin, 1984.

Kac, M., A stochastic model related to the telegrapher's equation, *Rocky Mt. J. Math.*, 4, 1974, 497–509.

Lin, C. C., and L. A. Segel, *Mathematics Applied to Deterministic Problems in the Applied Sciences*, Macmillan, New York, 1974.

Okubo, A., *Diffusion and Ecological Problems: Mathematical Models*, Springer, Berlin, 1980.

Risken, H., *The Fokker-Planck Equation*, Springer, Berlin, 1984.

Schuss, Z., *Theory and Applications of Stochastic Differential Equations*, Wiley, New York, 1980.

Taylor, G. I., Diffusion by continuous movements, *Proc. London Math. Soc.*, 20, 1920, 196–212.

Uhlenbeck, G. E., and L. S. Ornstein, On the theory of Brownian motion, *Phys. Rev.*, 36, 1930, 823–841.

Van Kampen, N. G., *Stochastic Processes in Physics and Chemistry*, NorthHolland, Amsterdam, 1981.

Wax, N., Ed., *Selected Papers on Noise and Stochastic Processes*, Dover, New York, 1954.

CHAPTER 2

Courant, R., *Methods of Mathematical Physics*, Vol. 2, Wiley-Interscience, New York, 1962.

Courant, R., and K. O. Friedrichs, *Supersonic Flow and Shock Waves*, Springer, New York, 1976.

Garabedian. P., *Partial Differential Equations*, 2nd ed., Chelsea, New York, 1986.

John, F., *Partial Differential Equations*, 4th ed., Springer, New York, 1991.

Rhee, H., R. Aris, and N. R. Amundson, *First Order Partial Differential Equations*, 2 Vols., Dover, New York, 2002.

Smoller, J., *Shock Waves and Reaction-Diffusion Equations*, Springer, New York, 1983.

Whitham, G. B., *Linear and Nonlinear Waves*, Wiley-Interscience, New York, 1974.

CHAPTER 3

Courant, R., *Methods of Mathematical Physics*, Vol. 2, Wiley-Interscience, New York, 1962.

Jeffrey, A., *Quasilinear Hyperbolic Systems and Waves*, Pitman, Boston, 1976.

CHAPTER 4

Abramowitz, M., and I. Stegun, *Handbook of Mathematical Functions*, Dover, New York, 1964.

Berg, P. W., and J. L. McGregor, *Elementary Partial Differential Equations*, Holden-Day, San Francisco, Calif., 1966.

Birkhoff, G., and G. C. Rota, *Ordinary Differential Equations*, 3rd ed., Wiley, New York, 1978.

Churchill, R. V., and J. W. Brown, *Fourier Series and Boundary Value Problems*, 4th ed., McGraw-Hill, New York, 1987.

Courant, R., *Methods of Mathematical Physics*, Vol. 2, Wiley-Interscience, New York, 1962.

Courant, R., and D. Hilbert, *Methods of Mathematical Physics*, Vol. 1, Interscience, New York, 1953.

Duff, G. F. D., and D. Naylor, *Differential Equations of Applied Mathematics*, Wiley, New York, 1966.

Haberman, R., *Elementary Applied Partial Differential Equations*, 4th ed., Prentice-Hall, Englewood Cliffs, N.J., 2003.

Iooss, G., and D. D. Joseph, *Elementary Stability and Bifurcation Theory*, Springer, New York, 1980.

Joseph, D. D., *Stability of Fluid Motions*, 2 vols., Springer, Berlin, 1976.

Lin, C. C., and L. A. Segel, *Mathematics Applied to Deterministic Problems in the Natural Sciences*, Macmillan, New York, 1974.

Pinsky, M., *Introduction to Partial Differential Equations with Applications*, McGraw-Hill, New York, 1984.

Stakgold, I., *Green's Functions and Boundary Value Problems*, 2nd ed., Wiley-Interscience, New York, 1988.

Taylor, A. E., and W. R. Mann, *Advanced Calculus*, 3rd ed., Wiley, New York, 1983.

Tikhonov, A. N., and A. A. Samarskii, *Equations of Mathematical Physics*, Macmillan, New York, 1963.

Tolstoy, C. P., *Fourier Series*, Dover, New York, 1976.

Vladimirov, V. S., *Equations of Mathematical Physics*, Dekker, New York, 1971.

Weinberger, H. F., *A First Course in Partial Differential Equations*, Dover, New York, 1995.

Zachmanoglou, E. C., and D. W. Thoe, *Introduction to Partial Differential Equations with Applications*, Dover, New York, 1987.

CHAPTER 5

Bleistein, N., and R. Handelsman, *Asymptotic Expansions of Integrals*, Dover, New York, 1986.

Churchill, R. V., *Operational Mathematics*, 3rd ed., McGraw-Hill, New York, 1972.

Copson, E. T., *Asymptotic Expansions*, Cambridge University Press, Cambridge, 1965.

Davies, B., *Integral Transforms and Their Applications*, Springer, New York, 1984.

Duff, G. F. D., and D. Naylor, *Differential Equations of Applied Mathematics*, Wiley, New York, 1966.

Dym, H., and H. P. McKean, Jr., *Fourier Series and Integrals*, Academic, New York, 1972.

Erdelyi, A., et al., *Tables of Integral Transforms*, McGraw-Hill, New York, 1953.

Olver, F. W. J., *Asymptotics and Special Functions*, Academic, New York, 1974.

Pearson, C. E., Ed., *Handbook of Applied Mathematics*, 2nd ed., Van Nostrand, New York, 1990.

Sirovich, L., Asymptotic evaluation of multidimensional integrals, *J. Math. Phys.*, 11, 1970, 1365–1374.

Sirovich, L., *Techniques of Asymptotic Analysis*, Springer, New York, 1971.

Sneddon, I. N., *Fourier Transforms*, Dover, New York, 1999.

Sneddon, I. N., *The Use of Integral Transforms*, McGraw-Hill, New York, 1972.

Weinberger, H. F., *A First Course in Partial Differential Equations*, Dover, New York, 1995.

Wolf, K. B., *Integral Transforms in Science and Engineering*, Plenum, New York, 1979.

CHAPTER 6

Courant, R., *Methods of Mathematical Physics*, Vol. 2, Wiley-Interscience, New York, 1962.

Garabedian, P., *Partial Differential Equations*, 2nd ed., Chelsea, New York, 1986.

Hadamard, J., *Lectures on Cauchy's Problem*, Dover, New York, 1952.

Logan, J. D., *Applied Mathematics*, Wiley-Interscience, New York, 1987.

Tayler, A. B., *Mathematical Models in Applied Mechanics*, Oxford University Press, Oxford, 1985.

Tikhonov, A. N., and A. A. Samarskii, *Equations of Mathematical Physics*, Macmillan, New York, 1963.

Vladimirov, V. S., *Equations of Mathematical Physics*, Dekker, New York, 1971.

Webster, A. G., *Partial Differential Equations of Mathematical Physics*, Dover, New York, 1955.

CHAPTER 7

Carrier, G. F., and C. E. Pearson, *Partial Differential Equations*, 2nd ed., Academic, New York, 1988.

Courant, R., *Methods of Mathematical Physics*, Vol. 2, Wiley-Interscience, New York, 1962.

Courant, R., and D. Hilbert, *Methods of Mathematical Physics*, Vol. 1, Interscience, New York, 1953.

Duff, G. F. D., and D. Naylor, *Differential Equations of Applied Mathematics*, Wiley, New York, 1966.

Greenberg, M. D., *Applications of Green's Functions in Science and Engineering*, Prentice-Hall, Englewood Cliffs, N.J., 1971.

Jones, D. S., *The Theory of Generalised Functions*, 2nd ed., Cambridge University Press, Cambridge, 1984.

Keller, J. B., and J. Papadakis, Eds., *Wave Propagation and Underwater Acoustics*, Springer, New York, 1977.

Lighthill, M. J., *Introduction to Fourier Analysis and Generalised Functions*, Cambridge University Press, Cambridge, 1960.

Morse, P. M., and H. Feshbach, *Methods of Theoretical Physics*, 2 vols., McGraw-Hill, New York, 1953.

Roach, G. F., *Green's Functions*, 2nd ed., Cambridge University Press, Cambridge, 1982.

Stakgold, I., *Green's Functions and Boundary Value Problems*, 2nd ed., Wiley-Interscience, New York, 1988.

Vladimirov, V. S., *Equations of Mathematical Physics*, Dekker, New York, 1971.

Weinberger, H. F., *A First Course in Partial Differential Equations*, Dover, New York, 1995.

CHAPTER 8

Ames, W. F., *Numerical Methods for Partial Differential Equations*, 2nd ed., Academic, New York, 1977.

Courant, R., *Methods of Mathematical Physics*, Vol. 2, Wiley-Interscience, New York, 1962.

Courant, R., and D. Hilbert, *Methods of Mathematical Physics*, Vol. 1, Interscience, New York, 1953.

Duff, G. F. D., and D. Naylor, *Differential Equations of Applied Mathematics*, Wiley, New York, 1966.

Finlayson, B. A., *The Method of Weighted Residuals and Variational Principles*, Academic, New York, 1972.

Garabedian, P., *Partial Differential Equations*, 2nd ed., Chelsea, New York, 1986.

Ockendon, H., and A. B. Tayler, *Inviscid Fluid Flows*, Springer, New York, 1983.

Protter, M. H., and H. F. Weinberger, *Maximum Principles in Differential Equations*, Springer, New York, 1984.

Tikhonov, A. N., and A. A. Samarskii, *Equations of Mathematical Physics*, Macmillan, New York, 1963.

Weinberger, H. F., *A First Course in Partial Differential Equations*, Dover, New York, 1995.

Widder, D. V., *The Heat Equation*, Academic, New York, 1975.

CHAPTER 9

Bender, C., and S. Orszag, *Advanced Mathematical Methods for Scientists and Engineers*, McGraw-Hill, New York, 1978.

Carrier, G. F., and C. E. Pearson, *Partial Differential Equations*, 2nd ed., Academic, New York, 1988.

Eckhaus, W., *Asymptotic Analysis of Singular Perturbations*, North-Holland, Amsterdam, 1979.

Kevorkian, J., and J. D. Cole, *Perturbation Methods in Applied Mathematics*, Springer, New York, 1981.

Nayfeh, A. H., *Perturbation Methods*, Wiley-Interscience, New York, 1973.

Van Dyke, M., *Perturbation Methods in Fluid Mechanics*, Parabolic Press, Palo Alto, Calif., 1975.

Whitham, G. B., *Linear and Nonlinear Waves*, Wiley-Interscience, New York, 1974.

CHAPTER 10

Akhmanov, S. A., A. P. Sukhorukov, and R. V Khoklov, Self-focusing and self-trapping of intense light beams in a nonlinear medium, *Sov. Phys. Usp.*, 10, 1968, 609–636.

Babič, V. M., and N. Y. Kirpičnikova, *The Boundary Layer Method in Diffraction Problems*, Springer, Berlin, 1979.

Bender, C., and S. Orszag, *Advanced Mathematical Methods for Engineers*, McGraw-Hill, New York, 1978.

Bleistein, N., *Mathematical Methods in Wave Propagation*, Academic, New York, 1984.

Buchal, R. N., and J. B. Keller, Boundary layer problems in diffraction theory, *Commun. Pure Appl. Math.*, 13, 1960, 85–114.

Courant, R., *Methods of Mathematical Physics*, Vol. 2, Wiley-Interscience, New York, 1962.

Haus, H. A., *Waves and Fields in Optoelectronics*, Prentice-Hall, Englewood Cliffs, N.J., 1984.

Jeffrey, A., and T. Kawahara, *Asymptotic Methods in Nonlinear Wave Theory*, Pitman, Boston, 1982.

Jones, D. S., *Methods in Electromagnetic Wave Propagation*, 2 vols., Oxford University Press, Oxford, 1987.

Keller, J. B., The geometrical theory of diffraction, *J. Opt. Soc. Am.*, 12, 1962, 116–132.

Keller, J. B., R. M. Lewis, and B. Seckler, Asymptotic solution of some diffraction problems, *Comm. Pure Appl. Math.*, 9, 1956, 207–251.

Kevorkian, J., and J. D. Cole, *Perturbation Methods in Applied Mathematics*, Springer, New York, 1981.

Newell, A. C., *Solitons in Mathematics and Physics*, SIAM, Philadelphia, Pa., 1985.

Newell, A. C., and J. V. Moloney, *Nonlinear Optics*, Addison-Wesley, Redwood City, Calif., 1992.

Taniuti, T., and K. Nishihara, *Nonlinear Waves*, Pitman, Boston, 1983.

Zauderer, E., Asymptotic decomposition of partial differential equations, *SIAM J. Appl. Math.*, 35, 1978, 575–596.

CHAPTER 11

Ames, W. F., *Numerical Methods for Partial Differential Equations*, 2nd ed., Academic, New York, 1977.

Dahlquist, G., and A. Björk, *Numerical Methods*, Prentice-Hall, Englewood Cliffs, N.J., 1974.

DuChateau, P., and D. Zachmann, *Applied Partial Differential Equations*, Dover, New York, 2003.

Gerald, C. F., and P. O. Wheatley, *Applied Numerical Analysis*, 5th ed., Addison-Wesley, Reading, Mass., 1994.

Gustafsson, B., H.-O. Kreiss, and J. Oliger, *Time Dependent Problems and Difference Methods*, Wiley-Interscience, New York, 1995.

Isaacson, E., and H. B. Keller, *Analysis of Numerical Methods*, Dover, New York, 1994.

Lapidus, L., and G. F. Pinder, *Numerical Solution of Partial Differential Equations in Science and Engineering*, Wiley-Interscience, New York, 1982.

Larsson, S., and V. Thomée, *Partial Differential Equations with Numerical Methods*, Springer, Berlin, 2003.

Richtmyer, R. D., and K. W. Morton, *Difference Methods for Initial-Value Problems*, 2nd ed., Wiley-Interscience, New York, 1967.

Smith, G. D., *Numerical Solution of Partial Differential Equations: Finite Difference Methods*, 3rd ed., Oxford University Press, New York, 1986.

Strang, G., *Introduction to Applied Mathematics*, Wellesley-Cambridge, Cambridge, 1986.

Strickwerda, J. C., *Finite Differences and Partial Differential Equations*, SIAM, Philadelphia, Pa., 2004.

CHAPTER 12

Gerald, C. F., and P. O. Wheatley, *Applied Numerical Analysis*, 5th ed., Addison-Wesley, Reading, Mass., 1994.

Larsson, S., and V. Thomée, *Partial Differential Equations with Numerical Methods*, Springer, Berlin, 2003.

Strang, G., *Introduction to Applied Mathematics*, Wellesley-Cambridge, Cambridge, 1986.

Strang, G and G. J. Fix, *An Analysis of the Finite Element Method*, Prentice-Hall, Englewood Cliffs, N.J., 1973.

INDEX

C

D

E

F

G

H

I

J

K

L

M

N

O

P

R

S

T

U

V

W

Y

PURE AND APPLIED MATHEMATICS

A Wiley-Interscience Series of Texts, Monographs, and Tracts

Consulting Editor: DAVID A. COX
Founded by RICHARD COURANT
Editors Emeriti: MYRON B. ALLEN III, DAVID A. COX, PETER HILTON, HARRY HOCHSTADT, PETER LAX, JOHN TOLAND

ADÁMEK, HERRLICH, and STRECKER—Abstract and Concrete Catetories
ADAMOWICZ and ZBIERSKI—Logic of Mathematics
AINSWORTH and ODEN—A Posteriori Error Estimation in Finite Element Analysis
AKIVIS and GOLDBERG—Conformal Differential Geometry and Its Generalizations
ALLEN and ISAACSON—Numerical Analysis for Applied Science
*ARTIN—Geometric Algebra
AUBIN—Applied Functional Analysis, Second Edition
AZIZOV and IOKHVIDOV—Linear Operators in Spaces with an Indefinite Metric
BERG—The Fourier-Analytic Proof of Quadratic Reciprocity
BERMAN, NEUMANN, and STERN—Nonnegative Matrices in Dynamic Systems
BERKOVITZ—Convexity and Optimization in $\mathbb{R}^n$
BOYARINTSEV—Methods of Solving Singular Systems of Ordinary Differential Equations
BURK—Lebesgue Measure and Integration: An Introduction
*CARTER—Finite Groups of Lie Type
CASTILLO, COBO, JUBETE, and PRUNEDA—Orthogonal Sets and Polar Methods in Linear Algebra: Applications to Matrix Calculations, Systems of Equations, Inequalities, and Linear Programming
CASTILLO, CONEJO, PEDREGAL, GARCIÁ, and ALGUACIL—Building and Solving Mathematical Programming Models in Engineering and Science
CHATELIN—Eigenvalues of Matrices
CLARK—Mathematical Bioeconomics: The Optimal Management of Renewable Resources, Second Edition
COX—Galois Theory
†COX—Primes of the Form $x^2 + ny^2$: Fermat, Class Field Theory, and Complex Multiplication
*CURTIS and REINER—Representation Theory of Finite Groups and Associative Algebras
*CURTIS and REINER—Methods of Representation Theory: With Applications to Finite Groups and Orders, Volume I
CURTIS and REINER—Methods of Representation Theory: With Applications to Finite Groups and Orders, Volume II
DINCULEANU—Vector Integration and Stochastic Integration in Banach Spaces
*DUNFORD and SCHWARTZ—Linear Operators
Part 1—General Theory
Part 2—Spectral Theory, Self Adjoint Operators in Hilbert Space
Part 3—Spectral Operators
FARINA and RINALDI—Positive Linear Systems: Theory and Applications
FATICONI—The Mathematics of Infinity: A Guide to Great Ideas
FOLLAND—Real Analysis: Modern Techniques and Their Applications

*Now available in a lower priced paperback edition in the Wiley Classics Library.
†Now available in paperback.

Partial Differential Equations of Appplied Mathematics, Third Edition. By Erich Zauderer

FRÖLICHER and KRIEGL—Linear Spaces and Differentiation Theory
GARDINER—Teichmüller Theory and Quadratic Differentials
GILBERT and NICHOLSON—Modern Algebra with Applications, Second Edition
*GRIFFITHS and HARRIS—Principles of Algebraic Geometry
GRILLET—Algebra
GROVE—Groups and Characters
GUSTAFSSON, KREISS and OLIGER—Time Dependent Problems and Difference Methods
HANNA and ROWLAND—Fourier Series, Transforms, and Boundary Value Problems, Second Edition
*HENRICI—Applied and Computational Complex Analysis
Volume 1, Power Series—Integration—Conformal Mapping—Location of Zeros
Volume 2, Special Functions—Integral Transforms—Asymptotics—Continued Fractions
Volume 3, Discrete Fourier Analysis, Cauchy Integrals, Construction of Conformal Maps, Univalent Functions
*HILTON and WU—A Course in Modern Algebra
*HOCHSTADT—Integral Equations
JOST—Two-Dimensional Geometric Variational Procedures
KHAMSI and KIRK—An Introduction to Metric Spaces and Fixed Point Theory
*KOBAYASHI and NOMIZU—Foundations of Differential Geometry, Volume I
*KOBAYASHI and NOMIZU—Foundations of Differential Geometry, Volume II
KOSHY—Fibonacci and Lucas Numbers with Applications
LAX—Functional Analysis
LAX—Linear Algebra
LOGAN—An Introduction to Nonlinear Partial Differential Equations
MARKLEY—Principles of Differential Equations
MORRISON—Functional Analysis: An Introduction to Banach Space Theory
NAYFEH—Perturbation Methods
NAYFEH and MOOK—Nonlinear Oscillations
PANDEY—The Hilbert Transform of Schwartz Distributions and Applications
PETKOV—Geometry of Reflecting Rays and Inverse Spectral Problems
*PRENTER—Splines and Variational Methods
RAO—Measure Theory and Integration
RASSIAS and SIMSA—Finite Sums Decompositions in Mathematical Analysis
RENELT—Elliptic Systems and Quasiconformal Mappings
RIVLIN—Chebyshev Polynomials: From Approximation Theory to Algebra and Number Theory, Second Edition
ROCKAFELLAR—Network Flows and Monotropic Optimization
ROITMAN—Introduction to Modern Set Theory
ROSSI—Theorems, Corollaries, Lemmas, and Methods of Proof
*RUDIN—Fourier Analysis on Groups
SENDOV—The Averaged Moduli of Smoothness: Applications in Numerical Methods and Approximations
SENDOV and POPOV—The Averaged Moduli of Smoothness
SEWELL—The Numerical Solution of Ordinary and Partial Differential Equations, Second Edition
SEWELL—Computational Methods of Linear Algebra, Second Edition
*SIEGEL—Topics in Complex Function Theory
Volume I—Elliptic Functions and Uniformization Theory
Volume 2—Automorphic Functions and Abelian Integrals
Volume 3—Abelian Functions and Modular Functions of Several Variables

*Now available in a lower priced paperback edition in the Wiley Classics Library.
†Now available in paperback.

SMITH and ROMANOWSKA—Post-Modern Algebra
ŠOLÍN–Partial Differential Equations and the Finite Element Method
STADE—Fourier Analysis
STAKGOLD—Green's Functions and Boundary Value Problems, Second Editon
STAHL—Introduction to Topology and Geometry
STANOYEVITCH—Introduction to Numerical Ordinary and Partial Differential Equations Using MATLAB®
*STOKER—Differential Geometry
*STOKER—Nonlinear Vibrations in Mechanical and Electrical Systems
*STOKER—Water Waves: The Mathematical Theory with Applications
WATKINS—Fundamentals of Matrix Computations, Second Edition
WESSELING—An Introduction to Multigrid Methods
†WHITHAM—Linear and Nonlinear Waves
ZAUDERER—Partial Differential Equations of Applied Mathematics, Third Edition

*Now available in a lower priced paperback edition in the Wiley Classics Library.
†Now available in paperback.