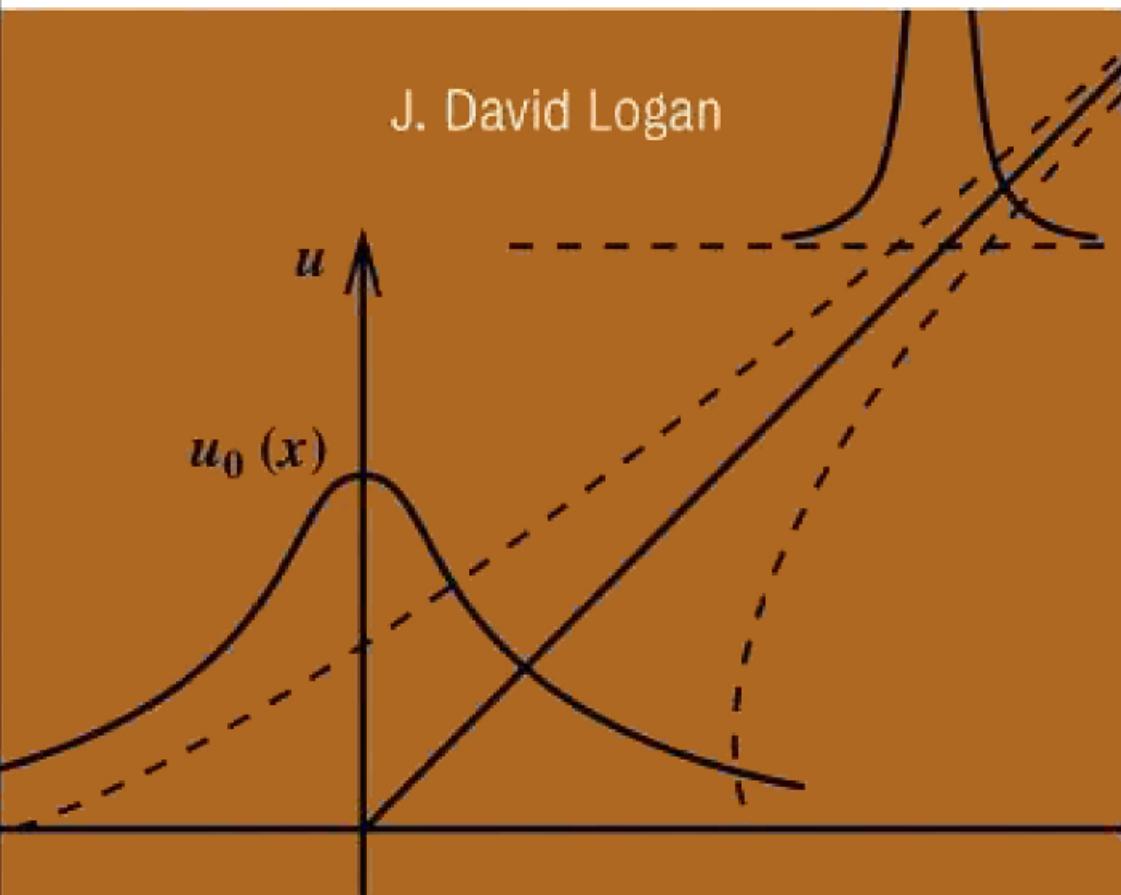


AN INTRODUCTION TO NONLINEAR PARTIAL DIFFERENTIAL EQUATIONS

SECOND EDITION

J. David Logan



An Introduction to Nonlinear Partial Differential Equations

PURE AND APPLIED MATHEMATICS

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

Founded by RICHARD COURANT

Editors Emeriti: MYRON B. ALLEN III, DAVID A. COX, PETER HILTON,
HARRY HOCHSTADT, PETER LAX, JOHN TOLAND

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

An Introduction to Nonlinear Partial Differential Equations

Second Edition

J. David Logan

Willa Cather Professor of Mathematics
University of Nebraska, Lincoln
Department of Mathematics
Lincoln, NE



A JOHN WILEY & SONS, INC., PUBLICATION

Copyright © 2008 by John Wiley & Sons, Inc. All rights reserved.

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

Published simultaneously in Canada.

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning, or otherwise, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 750-4470, or on the web at www.copyright.com. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at <http://www.wiley.com/go/permission>.

Limit of Liability/Disclaimer of Warranty: While the publisher and author have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives or written sales materials. The advice and strategies contained herein may not be suitable for your situation. You should consult with a professional where appropriate. Neither the publisher nor author shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

For general information on our other products and services or for technical support, please contact our Customer Care Department within the United States at (800) 762-2974, outside the United States at (317) 572-3993 or fax (317) 572-4002.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic format. For information about Wiley products, visit our web site at www.wiley.com.

Library of Congress Cataloging-in-Publication Data:

Logan, J. David (John David)

An introduction to nonlinear partial differential equations / J. David Logan. — 2nd ed.

p. cm.

Includes bibliographical references and index.

ISBN 978-0-470-22595-0 (cloth : acid-free paper)

1. Differential equations, Nonlinear. 2. Differential equations, Partial. I. Title.

QA377.L58 2008

515'.353—dc22

2007047514

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

To Tess, for all her affection and support

Contents

Preface	xi
1. Introduction to Partial Differential Equations	1
1.1 Partial Differential Equations	2
1.1.1 Equations and Solutions	2
1.1.2 Classification	5
1.1.3 Linear versus Nonlinear	8
1.1.4 Linear Equations	11
1.2 Conservation Laws	20
1.2.1 One Dimension	20
1.2.2 Higher Dimensions	23
1.3 Constitutive Relations	25
1.4 Initial and Boundary Value Problems	35
1.5 Waves	45
1.5.1 Traveling Waves	45
1.5.2 Plane Waves	50
1.5.3 Plane Waves and Transforms	52
1.5.4 Nonlinear Dispersion	54
2. First-Order Equations and Characteristics	61
2.1 Linear First-Order Equations	62
2.1.1 Advection Equation	62
2.1.2 Variable Coefficients	64
2.2 Nonlinear Equations	68
2.3 Quasilinear Equations	72
2.3.1 The General Solution	76

2.4	Propagation of Singularities	81
2.5	General First-Order Equation	86
2.5.1	Complete Integral	91
2.6	A Uniqueness Result	94
2.7	Models in Biology	96
2.7.1	Age Structure	96
2.7.2	Structured Predator–Prey Model	101
2.7.3	Chemotherapy	103
2.7.4	Mass Structure	105
2.7.5	Size-Dependent Predation	106
3.	Weak Solutions to Hyperbolic Equations	113
3.1	Discontinuous Solutions	114
3.2	Jump Conditions	116
3.2.1	Rarefaction Waves	118
3.2.2	Shock Propagation	119
3.3	Shock Formation	125
3.4	Applications	131
3.4.1	Traffic Flow	132
3.4.2	Plug Flow Chemical Reactors	136
3.5	Weak Solutions: A Formal Approach	140
3.6	Asymptotic Behavior of Shocks	148
3.6.1	Equal-Area Principle	148
3.6.2	Shock Fitting	152
3.6.3	Asymptotic Behavior	154
4.	Hyperbolic Systems	159
4.1	Shallow-Water Waves; Gas Dynamics	160
4.1.1	Shallow-Water Waves	160
4.1.2	Small-Amplitude Approximation	163
4.1.3	Gas Dynamics	164
4.2	Hyperbolic Systems and Characteristics	169
4.2.1	Classification	170
4.3	The Riemann Method	179
4.3.1	Jump Conditions for Systems	179
4.3.2	Breaking Dam Problem	181
4.3.3	Receding Wall Problem	183
4.3.4	Formation of a Bore	187
4.3.5	Gas Dynamics	190
4.4	Hodographs and Wavefronts	192
4.4.1	Hodograph Transformation	192
4.4.2	Wavefront Expansions	193

4.5	Weakly Nonlinear Approximations	201
4.5.1	Derivation of Burgers' Equation	202
5.	Diffusion Processes	209
5.1	Diffusion and Random Motion	210
5.2	Similarity Methods	217
5.3	Nonlinear Diffusion Models	224
5.4	Reaction–Diffusion; Fisher's Equation	234
5.4.1	Traveling Wave Solutions	235
5.4.2	Perturbation Solution	238
5.4.3	Stability of Traveling Waves	240
5.4.4	Nagumo's Equation	242
5.5	Advection–Diffusion; Burgers' Equation	245
5.5.1	Traveling Wave Solution	246
5.5.2	Initial Value Problem	247
5.6	Asymptotic Solution to Burgers' Equation	250
5.6.1	Evolution of a Point Source	252
	Appendix: Dynamical Systems	257
6.	Reaction–Diffusion Systems	267
6.1	Reaction–Diffusion Models	268
6.1.1	Predator–Prey Model	270
6.1.2	Combustion	271
6.1.3	Chemotaxis	274
6.2	Traveling Wave Solutions	277
6.2.1	Model for the Spread of a Disease	278
6.2.2	Contaminant Transport in Groundwater	284
6.3	Existence of Solutions	292
6.3.1	Fixed-Point Iteration	293
6.3.2	Semilinear Equations	297
6.3.3	Normed Linear Spaces	300
6.3.4	General Existence Theorem	303
6.4	Maximum Principles and Comparison Theorems	309
6.4.1	Maximum Principles	309
6.4.2	Comparison Theorems	314
6.5	Energy Estimates and Asymptotic Behavior	317
6.5.1	Calculus Inequalities	318
6.5.2	Energy Estimates	320
6.5.3	Invariant Sets	326
6.6	Pattern Formation	333

7. Equilibrium Models	345
7.1 Elliptic Models	346
7.2 Theoretical Results	352
7.2.1 Maximum Principle	353
7.2.2 Existence Theorem	355
7.3 Eigenvalue Problems	358
7.3.1 Linear Eigenvalue Problems	358
7.3.2 Nonlinear Eigenvalue Problems	361
7.4 Stability and Bifurcation	364
7.4.1 Ordinary Differential Equations	364
7.4.2 Partial Differential Equations	368
References	387
Index	395

Preface

Nonlinear partial differential equations (PDEs) is a vast area, and practitioners include applied mathematicians, analysts, and others in the pure and applied sciences. This introductory text on nonlinear partial differential equations evolved from a graduate course I have taught for many years at the University of Nebraska at Lincoln. It emerged as a pedagogical effort to introduce, at a fairly elementary level, nonlinear PDEs in a format and style that is accessible to students with diverse backgrounds and interests. The audience has been a mixture of graduate students from mathematics, physics, and engineering. The prerequisites include an elementary course in PDEs emphasizing Fourier series and separation of variables, and an elementary course in ordinary differential equations.

There is enough independence among the chapters to allow the instructor considerable flexibility in choosing topics for a course. The text may be used for a second course in partial differential equations, a first course in nonlinear PDEs, a course in PDEs in the biological sciences, or an advanced course in applied mathematics or mathematical modeling. The range of applications include biology, chemistry, gas dynamics, porous media, combustion, traffic flow, water waves, plug flow reactors, heat transfer, and other topics of interest in applied mathematics.

There are three major changes from the first edition, which appeared in 1994. Because the original chapter on chemically reacting fluids was highly specialized for an introductory text, it has been removed from the new edition. Additionally, because of the surge of interest in mathematical biology, considerable material on that topic has been added; this includes linear and nonlinear age structure, spatial effects, and pattern formation. Finally, the text has been reorganized with the chapters on hyperbolic equations separated from

the chapters on diffusion processes, rather than intermixing them.

The references have been updated and, as in the previous edition, are selected to suit the needs of an introductory text, pointing the reader to parallel treatments and resources for further study. Finally, many new exercises have been added. The exercises are intermediate-level and are designed to build the students' problem solving techniques beyond what is experienced in a beginning course.

Chapter 1 develops a perspective on how to understand problems involving PDEs and how the subject interrelates with physical phenomena. The subject is developed from the basic conservation law, which, when appended to constitutive relations, gives rise to the fundamental models of diffusion, advection, and reaction. There is emphasis on understanding that nonlinear hyperbolic and parabolic PDEs describe evolutionary processes; a solution is a signal that is propagated into a spacetime domain from the boundaries of that domain. Also, there is focus on the structure of the various equations and what the terms describe physically. Chapters 2–4 deal with wave propagation and hyperbolic problems. In *Chapter 2* we assume that the equations have smooth solutions and we develop algorithms to solve the equations analytically. In *Chapter 3* we study discontinuous solutions and shock formation, and we introduce the concept of a weak solution. In keeping with our strategy of thinking about initial waveforms evolving in time, we focus on the initial value problem rather than the general Cauchy problem. The idea of characteristics is central and forms the thread that weaves through these two chapters. Next, *Chapter 4* introduces the shallow-water equations as the prototype of a hyperbolic system, and those equations are taken to illustrate basic concepts associated with hyperbolic systems: characteristics, Riemann's method, the hodograph transformation, and asymptotic behavior. Also, the general classification of systems of first-order PDEs is developed, and weakly nonlinear methods of analysis are described; the latter are illustrated by a derivation of Burgers' equation.

Chapters 1–4 can form the basis of a one-semester course focusing on wave propagation, characteristics, and hyperbolic equations.

Chapter 5 introduces diffusion processes. After establishing a probabilistic basis for diffusion, we examine methods that are useful in studying the solution structure of diffusion problems, including phase plane analysis, similarity methods, and asymptotic expansions. The prototype equations for reaction–diffusion and advection–diffusion, Fisher's equation and Burgers' equation, respectively, are studied in detail with emphasis on traveling wave solutions, the stability of those solutions, and the asymptotic behavior of solutions. The Appendix to Chapter 5 reviews phase plane analysis. In *Chapter 6* we discuss systems of reaction–diffusion equations, emphasizing applications and model building, especially in the biological sciences. We expend some effort addressing theoret-

ical concepts such as existence, uniqueness, comparison and maximum principles, energy estimates, blowup, and invariant sets; a key application includes pattern formation. Finally, elliptic equations are introduced in *Chapter 7* as a asymptotic limit of reaction–diffusion equations; nonlinear eigenvalue problems, stability, and bifurcation phenomena form the core of this chapter.

Chapter 1, along with Chapters 5–8, can form the basis of a one-semester course in diffusion and reaction–diffusion processes, with emphasis on PDEs in mathematical biology.

I want to acknowledge many users of the first edition who suggested improvements, corrections, and new topics. Their excitement for a second edition, along with the unwavering encouragement of my editor Susanne Steitz-Filler at Wiley, provided the stimulus to actually complete it. My own interest in nonlinear PDEs was spawned over many years by collaboration with those with whom I have had the privilege of working: Kane Yee at Kansas State, John Bdzil at Los Alamos, Ash Kapila at Rensselaer Polytechnic Institute, and several of my colleagues at Nebraska (Professors Steve Cohn, Steve Dunbar, Tony Joern in biology, Glenn Ledder, Tom Shores, Vitaly Zlotnik in geology, and my former student Bill Wolessensky, now at the College of Saint Mary). Readers of this text will see the influence of the classic books of G. B. Whitham (*Linear and Nonlinear Waves*) and J. Smoller (*Shock Waves and Reaction-Diffusion Equations*), R. Courant and K. O. Friedrichs (*Supersonic Flow and Shock Waves*), and the text on mathematical biology by J. D. Murray (*Mathematical Biology*). Finally, I express my gratitude to the National Science Foundation and to the Department of Energy for supporting my research efforts over the last several years.

J. David Logan
Lincoln, Nebraska

1

Introduction to Partial Differential Equations

Partial differential equations (PDEs) is one of the basic areas of applied analysis, and it is difficult to imagine any area of applications where its impact is not felt. In recent decades there has been tremendous emphasis on understanding and modeling nonlinear processes; such processes are often governed by nonlinear PDEs, and the subject has become one of the most active areas in applied mathematics and central in modern-day mathematical research. Part of the impetus for this surge has been the advent of high-speed, powerful computers, where computational advances have been a major driving force.

This initial chapter focuses on developing a perspective on understanding problems involving PDEs and how the subject interrelates with physical phenomena. It also provides a transition from an elementary course, emphasizing eigenfunction expansions and linear problems, to a more sophisticated way of thinking about problems that is suggestive of and consistent with the methods in nonlinear analysis.

Section 1.1 summarizes some of the basic terminology of elementary PDEs, including ideas of classification. In Section 1.2 we begin the study of the origins of PDEs in physical problems. This interdependence is developed from the basic, one-dimensional conservation law. In Section 1.3 we show how constitutive relations can be appended to the conservation law to obtain equations that model the fundamental processes of diffusion, advection or transport, and reaction. Some of the common equations, such as the diffusion equation, Burgers' equation, Fisher's equation, and the porous media equation, are obtained

as models of these processes. In Section 1.4 we introduce initial and boundary value problems to see how auxiliary data specialize the problems. Finally, in Section 1.5 we discuss wave propagation in order to fix the notion of how evolution equations carry boundary and initial signals into the domain of interest. We also introduce some common techniques for determining solutions of a certain form (e.g., traveling wave solutions). The ideas presented in this chapter are intended to build an understanding of evolutionary processes so that the fundamental concepts of hyperbolic problems and characteristics, as well as diffusion problems, can be examined in later chapters with a firmer base.

1.1 Partial Differential Equations

1.1.1 Equations and Solutions

A *partial differential equation* is an equation involving an unknown function of several variables and its partial derivatives. To fix the notion, a *second-order PDE in two independent variables* is an equation of the form

$$G(x, t, u, u_x, u_t, u_{xx}, u_{tt}, u_{xt}) = 0, \quad (x, t) \in D, \quad (1.1.1)$$

where, as indicated, the independent variables x and t lie in some given domain D in \mathbb{R}^2 . By a *solution* to (1.1.1) we mean a twice continuously differentiable function $u = u(x, t)$ defined on D that, when substituted into (1.1.1), reduces it to an identity on D . The function $u(x, t)$ is assumed to be twice continuously differentiable, so that it makes sense to calculate its first and second derivatives and substitute them into the equation; a smooth solution like this is called a *classical solution* or *genuine solution*. Later we extend the notion of solution to include functions that may have discontinuities, or discontinuities in their derivatives; such functions are called *weak solutions*. The xt domain D where the problem is defined is referred to as a *spacetime domain*, and PDEs that include time t as one of the independent variables are called *evolution equations*. When the two independent variables are both spatial variables, say, x and y rather than x and t , the PDE is an *equilibrium* or *steady-state* equation. Evolution equations govern time-dependent processes, and equilibrium equations often govern physical processes after the transients caused by initial or boundary conditions die away.

Graphically, a solution $u = u(x, t)$ of (1.1.1) is a smooth surface in three-dimensional xtu space lying over the domain D in the xt plane, as shown in Figure 1.1. An alternative representation is a plot in the xu -plane of the function $u = u(x, t_0)$ for some fixed time $t = t_0$ (see Figures 1.1 and 1.2). Such

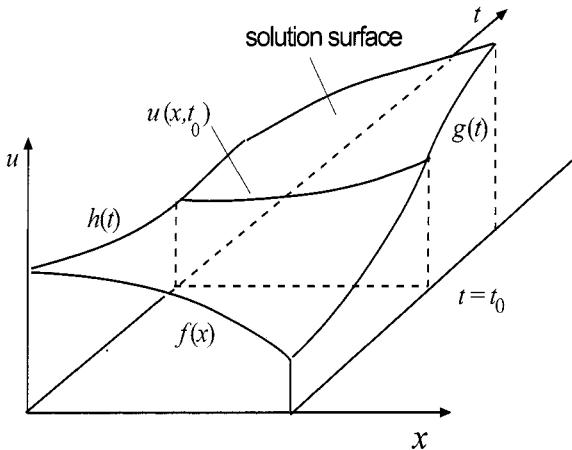


Figure 1.1 Solution surface $u = u(x, t)$ in xtu space, also showing a time snapshot or wave profile $u(x, t_0)$ at time t_0 . The functions f , g , and h represent values of u on the boundary of the domain, which are often prescribed as initial and boundary conditions.

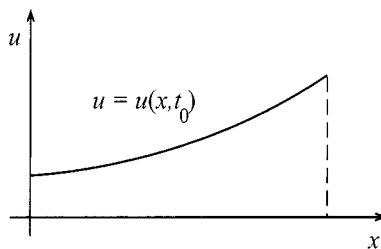


Figure 1.2 Time snapshot $u(x, t_0)$ at $t = t_0$ graphed in xu space. Often several snapshots for different times t are graphed on the same set of xu coordinates to indicate how the wave profiles are evolving in time.

representations are called *time snapshots* or *wave profiles* of the solution; time snapshots are profiles in space of the solution $u = u(x, t)$ frozen at a fixed time t_0 , or, stated differently, slices of the solution surface at a fixed time t_0 . Occasionally, several time snapshots are plotted simultaneously on the same set of xu axes to indicate how profiles change. It is also helpful on occasion to think of a solution in abstract terms. For example, suppose that $u = u(x, t)$ is a solution of a PDE for $x \in \mathbb{R}$ and $0 \leq t \leq T$. Then for each t , $u(x, t)$ is a function of x (a profile), and it generally belongs to some space of functions \mathbf{X} . To fix the idea, suppose that \mathbf{X} is the set of all twice continuously differentiable

functions on \mathbb{R} that approach zero at infinity. Then the solution can be regarded as a mapping from the time interval $[0, T]$ into the function space \mathbf{X} ; that is, to each t in $[0, T]$ we associate a function $u(\cdot, t)$, which is the wave profile at time t .

A PDE has infinitely many solutions, depending on arbitrary functions. For example, the *wave equation*

$$u_{tt} - c^2 u_{xx} = 0 \quad (1.1.2)$$

has a general solution that is the superposition (sum) of a right traveling wave $F(x - ct)$ of speed c and a left traveling wave $G(x + ct)$ of speed c ; that is,

$$u(x, t) = F(x - ct) + G(x + ct) \quad (1.1.3)$$

for any twice continuously differentiable functions F and G . (See the Exercises at the end of this section.) We contrast the situation in ordinary differential equations, where solutions depend on arbitrary constants; there, initial or boundary conditions fix the arbitrary constants and select a unique solution. For PDEs this occurs as well; initial and boundary conditions are usually imposed and select one of the infinitude of solutions. These auxiliary or subsidiary conditions are suggested by the underlying physical problem from which the PDE arises, or by the type of PDE. A condition on u or its derivatives given at $t = 0$ along some segment of the x axis is called an *initial condition*, while a condition along any other curve in the xt plane is called a *boundary condition*. PDEs with auxiliary conditions are called *initial value problems*, *boundary value problems*, or *initial-boundary value problems*, depending on the type of subsidiary conditions that are specified.

Example. The initial value problem for the wave equation is

$$u_{tt} - c^2 u_{xx} = 0, \quad x \in \mathbb{R}, \quad t > 0, \quad (1.1.4)$$

$$u(x, 0) = f(x), \quad u_t(x, 0) = g(x), \quad x \in \mathbb{R}, \quad (1.1.5)$$

where f and g are given twice continuously differentiable functions on \mathbb{R} . The unique solution is given by (see Exercise 2)

$$u(x, t) = \frac{1}{2}[f(x - ct) + f(x + ct)] + \frac{1}{2c} \int_{x-ct}^{x+ct} g(s) ds, \quad (1.1.6)$$

which is *D'Alembert's formula*. So, in this example we think of the auxiliary data (1.1.5) as selecting one of the infinitude of solutions given by (1.1.3). Note that the solution at (x, t) depends only on the initial data (1.1.5) in the interval $[x - ct, x + ct]$. \square

Statements regarding the single second-order PDE (1.1.1) can be generalized in various directions. Higher-order equations (as well as first-order equations), several independent variables, and several unknown functions (governed by systems of PDEs) are all possibilities.

1.1.2 Classification

PDEs are classified into different types, depending on either the type of physical phenomena from which they arise or a mathematical basis. As the reader has learned from previous experience, there are three fundamental types of equations: those that govern diffusion processes, those that govern wave propagation, and those that govern equilibrium phenomena. Equations of mixed type also occur. We consider a single, second order PDE of the form

$$a(x, t)u_{xx} + 2b(x, t)u_{xt} + c(x, t)u_{tt} = d(x, t, u, u_x, u_t), \quad (x, t) \in D, \quad (1.1.7)$$

where a , b , and c are continuous functions on D , and not all of a , b , and c vanish simultaneously at some point of D . The function d on the right side is assumed to be continuous as well. Classification is based on the combination of the second-order derivatives in the equation. If we define the *discriminant* Δ by $\Delta = b^2 - ac$, then (1.1.7) is *hyperbolic* if $\Delta > 0$, *parabolic* if $\Delta = 0$, and *elliptic* if $\Delta < 0$.

Hyperbolic and parabolic equations are evolution equations that govern wave propagation and diffusion processes, respectively, and elliptic equations are associated with equilibrium or steady-state processes. In the latter case, we use x and y as independent variables rather than x and t . There is also a close relationship between the classification and the kinds of initial and boundary conditions that may be imposed on a PDE to obtain a well-posed mathematical problem, or one that is physically relevant. Because classification is based on the highest-order derivatives in (1.1.7), or the *principal part* of the equation, and because Δ depends on x and t , equations may change type as x and t vary throughout the domain.

Now we demonstrate that equation (1.1.7) can be transformed to certain simpler, or *canonical*, forms, depending on the classification, by a change of independent variables

$$\xi = \xi(x, t), \quad \eta = \eta(x, t). \quad (1.1.8)$$

We now perform this calculation, with the view of actually trying to determine (1.1.8) such that (1.1.7) reduces to a simpler form in the $\xi\eta$ coordinate system. The transformation (1.1.8) is assumed to be invertible, which requires that the Jacobian $J = \xi_x\eta_t - \xi_t\eta_x$ be nonzero in any region where the transformation is applied. A straightforward application of the chain rule, which the reader

can verify, shows that the left side of (1.1.7) becomes, under the change of independent variables (1.1.8)

$$au_{xx} + 2bu_{xt} + cu_{tt} + \dots = Au_{\xi\xi} + 2Bu_{\xi\eta} + Cu_{\eta\eta} + \dots, \quad (1.1.9)$$

where the three dots denote terms with lower-order derivatives, and where

$$\begin{aligned} A &= a\xi_x^2 + 2b\xi_x\xi_t + c\xi_t^2, \\ B &= a\xi_x\eta_x + b(\xi_x\eta_t + \xi_t\eta_x) + c\xi_t\eta_t, \\ C &= a\eta_x^2 + 2b\eta_x\eta_t + c\eta_t^2. \end{aligned}$$

Notice that the expressions for A and C have the same form, namely

$$a\phi_x^2 + 2b\phi_x\phi_t + c\phi_t^2,$$

and are independent.

In the *hyperbolic case* we can choose ξ and η such that $A = C = 0$. To this end, set

$$a\phi_x^2 + 2b\phi_x\phi_t + c\phi_t^2 = 0. \quad (1.1.10)$$

Because the discriminant Δ is positive, we can write (1.1.10) as (assume that a is not zero)

$$\frac{\phi_x}{\phi_t} = -\frac{b \pm \sqrt{b^2 - ac}}{a}.$$

To determine ϕ , we regard it as defining loci (curves) in the xt plane via the equation $\phi(x, t) = \text{const}$. The differentials dx and dt along one of these curves satisfy the relation $\phi_x dx + \phi_t dt = 0$ or $dt/dx = -\phi_x/\phi_t$. Therefore

$$\frac{dt}{dx} = \frac{b \pm \sqrt{b^2 - ac}}{a} \quad (1.1.11)$$

is a differential equation whose solutions determine the curves $\phi(x, t) = \text{const}$. On choosing the $+$ and $-$ signs in (1.1.11), respectively, we obtain $\xi(x, t)$ and $\eta(x, t)$ as integral curves of (1.1.11), making $A = C = 0$. Consequently, if (1.1.7) is hyperbolic, it can be reduced to the *canonical hyperbolic form*

$$u_{\xi\eta} + \dots = 0,$$

where the three dots denote terms involving lower-order derivatives (we leave it as an exercise to show that B is nonzero in this case).

The differential equations (1.1.11) are called the *characteristic equations* associated with (1.1.7), and the two sets of solution curves $\xi(x, t) = \text{const}$ and $\eta(x, t) = \text{const}$ are called the *characteristic curves*, or just the *characteristics*; ξ and η are called *characteristic coordinates*. In summary, in the hyperbolic case there are two real families of characteristics that provide a coordinate system

where the equation reduces to a simpler form. Characteristics are the fundamental concept in the analysis of hyperbolic problems because characteristic coordinates form a natural curvilinear coordinate system in which to examine these problems. In some cases, PDEs simplify to ODEs along the characteristic curves.

In the *parabolic case* ($b^2 - ac = 0$) there is just one family of characteristic curves, defined by

$$\frac{dt}{dx} = \frac{b}{a}.$$

Thus we may choose $\xi = \xi(x, t)$ as an integral curve of this equation to make $A = 0$. Then, if $\eta = \eta(x, t)$ is chosen as any smooth function independent of ξ (i.e., so that the Jacobian is nonzero), one can easily determine that $B = 0$ automatically, giving the *parabolic canonical form*

$$u_{\xi\xi} + \dots = 0.$$

Characteristics rarely play a role in parabolic problems.

In the *elliptic case* ($b^2 - ac < 0$) there are no real characteristics and, as in the parabolic case, characteristics play no role in elliptic problems. However, it is still possible to eliminate the mixed derivative term in (1.1.7) to obtain an elliptic canonical form. The procedure is to determine complex characteristics by solving (1.1.11), and then take real and imaginary parts to determine a transformation (1.1.8) that makes $A = C$ and $B = 0$ in (1.1.9). We leave it as an exercise to show that the transformation is given by

$$\alpha = \frac{1}{2}(\xi + \eta), \quad \beta = \frac{1}{2i}(\xi - \eta).$$

Then the *elliptic canonical form* is

$$u_{\alpha\alpha} + u_{\beta\beta} + \dots = 0,$$

where the Laplacian operator becomes the principal part.

Example. It is easy to see that the characteristic curves for the wave equation (1.1.2), which is hyperbolic, are the straight lines $x - ct = \text{const}$ and $x + ct = \text{const}$. These are shown in Figure 1.3. In this case the characteristic coordinates are given by $\xi = x - ct$ and $\eta = x + ct$. In these coordinates the wave equation transforms to $u_{\xi\eta} = 0$. We regard characteristics as curves in spacetime moving with speeds c and $-c$, and from the general solution (1.1.3) we observe that signals are propagated along these curves. In hyperbolic problems, in general, the characteristics are curves in spacetime along which signals are transmitted. \square

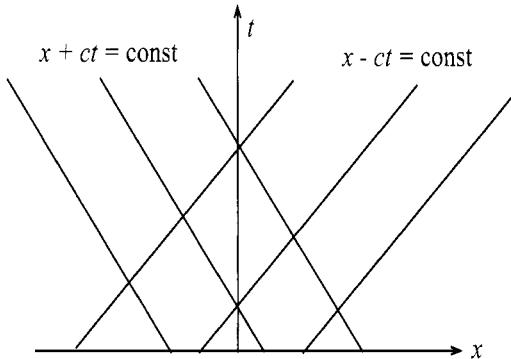


Figure 1.3 Characteristic diagram for the wave equation showing the forward and backward characteristics $x - ct = \text{const}$ and $x + ct = \text{const}$.

If the coefficients a , b , and c of the second-order derivatives in equation (1.1.7) depend on x , t , and u , then (1.1.7) is called a *quasilinear* equation. In this case we make the same classification as above, depending on the sign of the discriminant Δ ; now the type of the equation depends not only on the spacetime domain but also on the solution u itself. The canonical forms listed above are no longer valid in this case, and the characteristics defined by (1.1.11) cannot be determined a priori since a , b , and c depend on u , the unknown solution itself. Therefore, there is a significant increase in difficulty when the principal part of the equation is nonlinear.

There are other ways to approach the classification problem. In the preceding discussion the focus was on determining transformations under which a simplification occurs. In Section 6.1 we take a different perspective and ask whether it is possible to determine the solution u near a curve where the values of u and its first derivatives are known. That discussion is accessible to the reader at the present juncture, if desired. Yet another view of classification is presented in Chapter 4, where hyperbolic systems are discussed. Finally, from a physical perspective, we observe later in this chapter that hyperbolic problems are associated with wave propagation; parabolic problems, with diffusion; and elliptic problems, with equilibria.

1.1.3 Linear versus Nonlinear

The most important classification criterion is to distinguish PDEs as *linear* or *nonlinear*. Roughly, a homogeneous PDE is linear if the sum of two solutions is a solution, and a constant multiple of a solution is a solution. Otherwise, it is

nonlinear. The division of PDEs into these two categories is a significant one. The mathematical methods devised to deal with these two classes of equations are often entirely different, and the behavior of solutions differs substantially. One underlying cause is the fact that the solution space to a linear, homogeneous PDE is a vector space, and the linear structure of that space can be used with advantage in constructing solutions with desired properties that can meet diverse boundary and initial conditions. Such is not the case for nonlinear equations.

It is easy to find examples where nonlinear PDEs exhibit behavior with no linear counterpart. One is the breakdown of solutions and the formation of singularities, such as shock waves. A second is the existence of solitons, which are solutions to nonlinear dispersion equations. These solitary wave solutions maintain their shapes through collisions, in much the same way as linear equations do, even though the interactions are not linear. Nonlinear equations have come to the forefront because, basically, the world is nonlinear!

More formally, linearity and nonlinearity are usually defined in terms of the properties of the operator that defines the PDE itself. Let us assume that the PDE (1.1.1) can be written in the form

$$Lu = F, \quad (1.1.12)$$

where $F = F(x, t)$ and L is an operator that contains all the operations (differentiation, multiplication, composition, etc.) that act on $u = u(x, t)$. For example, the wave equation $u_{tt} - u_{xx} = 0$ can be written $Lu = 0$, where L is the partial differential operator $\partial_t^2 - \partial_x^2$. In (1.1.12) we reiterate that all terms involving the unknown function u are on the left side of the equation and are contained in the expression Lu ; the right side of (1.1.12) contains in F only expressions involving the independent variables x and t . If $F = 0$, then (1.1.12) is said to be *homogeneous*; otherwise, it is *nonhomogeneous*. We say that an operator L is *linear* if it is additive and if constants factor out of the operator, that is, (1) $L(u + v) = Lu + Lv$, and (2) $L(cu) = cLu$, where u and v are functions (in the domain of the operator) and c is any constant. The PDE (1.1.12) is *linear* if L is a linear operator; otherwise, the PDE is *nonlinear*.

Example. The equation $Lu = u_t + uu_x = 0$ is nonlinear because, for example, $L(cu) = cu_t + c^2uu_x$, which does not equal $cLu = c(u_t + uu_x)$. \square

Conditions (1) and (2) stated above imply that a linear homogeneous equation $Lu = 0$ has the property that if u_1, u_2, \dots, u_n are n solutions, the linear combination

$$u = c_1u_1 + c_2u_2 + \cdots + c_nu_n$$

is also a solution for any choice of the constants c_1, c_2, \dots, c_n . This fact is called the *superposition principle* for linear equations. For nonlinear equations we cannot superimpose solutions in this manner. The superposition principle can often be extended to infinite sums for linear problems, provided that convergence requirements are met. Superposition for linear equations allows one to construct, from a given set of solutions, another solution that meets initial or boundary requirements by choosing the constants c_1, c_2, \dots judiciously. This observation is the basis for the Fourier method, or eigenfunction expansion method, for linear, homogeneous boundary value problems, and we review this procedure at the end of the section. Moreover, superposition can often be extended to a family of solutions depending on a continuum of values of a parameter. More precisely, if $u = u(x, t; k)$ is a family of solutions of a linear homogeneous PDE for all values of k in some interval of real numbers I , one can superimpose these solutions formally using integration by defining

$$u(x, t) = \int_I c(k)u(x, t; k) dk,$$

where $c = c(k)$ is a function of the parameter k . Under certain conditions that must be established, the superposition $u(x, t)$ may again be a solution. As in the finite case, there is flexibility in selecting $c(k)$ to meet boundary or initial conditions. In fact, this procedure is the vehicle for transform methods for solving linear PDEs (Laplace transforms, Fourier transforms, etc.). We review this technique below. Finally, for a homogeneous, linear PDE the real and imaginary parts of a complex solution are both solutions. This is easily seen from the calculation

$$L(v + iw) = Lv + iLw = 0 + 0 = 0,$$

where the real-valued functions v and w satisfy $Lv = 0$ and $Lw = 0$. None of these methods based on superposition are applicable to nonlinear problems, and other methods must be sought. In summary, there is a profound difference between properties and solution methods for linear and nonlinear problems.

If most solution methods for linear problems are inapplicable to nonlinear equations, what methods can be developed? We mention a few.

1. *Perturbation Methods.* Perturbation methods are applicable to problems where a small or large parameter can be identified. In this case an approximate solution is sought as a series expansion in the parameter.
2. *Similarity Methods.* The similarity method is based on the PDE and its auxiliary conditions being invariant under a family of transformations depending on a small parameter. The invariance transformation allows one to identify a canonical change of variables that reduces the PDE to an ordinary differential equation (ODE), or reduces the order of the PDE.

3. *Characteristic Methods.* Nonlinear hyperbolic equations, which are associated with wave propagation, can be analyzed with success in characteristic coordinates (i.e., coordinates in spacetime along which the waves or signals propagate).
4. *Transformations.* Sometimes it is possible to identify transformations that change a given nonlinear equation into a simpler equation that can be solved.
5. *Numerical Methods.* Fast, large-scale computers have given tremendous impetus to the development and analysis of numerical algorithms to solve nonlinear problems and, in fact, have been a stimulus to the analysis of nonlinear equations.
6. *Traveling Wave Solutions.* Seeking solutions with special properties is a key technique. For example, traveling waves are solutions to evolution problems that represent fixed waveforms moving in time. The assumption of a traveling wave profile to a PDE sometimes reduces it to an ODE, often facilitating the analysis and solution. Traveling wave solutions form one type of similarity solution.
7. *Steady State Solutions and Their Stability.* Many PDEs have steady-state, or time-independent, solutions. Studying these equilibrium solutions and their stability is an important activity in many areas of application.
8. *Ad Hoc Methods.* The mathematical and applied science literature is replete with articles illustrating special methods that analyze a certain type of nonlinear PDE, or restricted classes of nonlinear PDEs.

These methods are primarily solution methods, which represent one aspect of the subject of nonlinear PDEs. Other basic issues are questions of existence and uniqueness of solutions, the regularity (smoothness) of solutions, and the investigation of stability properties of solutions. These and other theoretical questions have spawned investigations based on modern topological and algebraic concepts, and the subject of nonlinear PDEs has evolved into one of the most diverse, active areas of applied analysis.

1.1.4 Linear Equations

In this subsection we review, through examples, two techniques from elementary PDEs that illustrate the use of the superposition principles mentioned above. These calculations arise later in analyzing the local stability of equilibrium solutions to nonlinear problems.

Example. (*Separation of Variables*) Consider the following problem for $u = u(x, t)$ on the bounded interval $I : 0 \leq x \leq 1$ with $t > 0$, that is

$$u_t = Au, \quad 0 < x < 1, \quad t > 0, \quad (1.1.13)$$

$$u(0, t) = u(1, t) = 0, \quad t > 0, \quad (1.1.14)$$

$$u(x, 0) = f(x), \quad 0 \leq x \leq 1, \quad (1.1.15)$$

where A is a linear, spatial differential operator of the form

$$Au = -(pu_x)_x + qu.$$

The functions $p = p(x)$ and $q = q(x)$ are given, with p of one sign on I , and p , p' , and q continuous on I . Problems of this type are solved by Fourier's method, or the method of eigenfunction expansions. The idea is to construct infinitely many solutions that satisfy the PDE and the boundary conditions, equations (1.1.13) and (1.1.14), and then superimpose them, rigging up the constants so that the initial condition (1.1.15) is satisfied. This technique is called *separation of variables*, based on an assumption that the solution has the form $u(x, t) = g(t)y(x)$, where g and y are to be determined. When we substitute this form into the PDE and rearrange terms we obtain

$$\frac{g'}{g} = \frac{Ay}{y},$$

where the left side depends only on t and the right side depends only on x . A function of t can equal a function of x for all x and t only if both are equal to a constant, say, $-\lambda$, called the *separation constant*. Therefore

$$\frac{g'}{g} = \frac{Ay}{y} = -\lambda,$$

and we obtain two ODEs, one for g and one for y :

$$g' = -\lambda g, \quad -Ay = \lambda y.$$

We say that the equation separates. If we substitute the assumed form of u into the boundary conditions (1.1.14), then we obtain

$$y(0) = y(1) = 0.$$

The temporal equation is easily solved to get $g(t) = ce^{-\lambda t}$, where c is an arbitrary constant. The spatial equation along with its homogeneous (zero) boundary conditions give a boundary value problem (BVP) for y :

$$-Ay = \lambda y, \quad 0 < x < 1, \quad (1.1.16)$$

$$y(0) = y(1) = 0. \quad (1.1.17)$$

This BVP for y , which is differential eigenvalue problem called a *Sturm–Liouville problem*, has the property there are infinitely many real, discrete values of the separation constant λ , say, $\lambda = \lambda_n$, $n = 1, 2, \dots$, for which there are corresponding solutions $y = y_n(x)$, $n = 1, 2, \dots$. The λ_n are called the *eigenvalues* for the problem and the corresponding solutions $y = y_n(x)$ are called the *eigenfunctions*. The eigenvalues have the property that they are ordered and $|\lambda_n| \rightarrow \infty$ as $n \rightarrow \infty$. Therefore we have obtained a countably infinite number of solutions to the PDE that satisfy the boundary conditions:

$$u_n(x, t) = c_n e^{-\lambda_n t} y_n(x), \quad n = 1, 2, \dots$$

Now, here is where superposition is used. We add up these solutions and pick the constants c_n so that the initial condition (1.1.15) is satisfied, thus obtaining the solution to the problem; that is, we form

$$u(x, t) = \sum_{n=1}^{\infty} c_n e^{-\lambda_n t} y_n(x).$$

Formally applying the initial condition gives

$$u(x, 0) = f(x) = \sum_{n=1}^{\infty} c_n y_n(x). \quad (1.1.18)$$

The right side is an expansion of the initial condition f in terms of the eigenfunctions y_n , and we can use it to determine the coefficients c_n . This calculation is enabled by a very important property of the eigenfunctions, namely, orthogonality. If we define the inner product of two functions ϕ and ψ by

$$(\phi, \psi) = \int_0^1 \phi(x) \psi(x) dx,$$

then we say ϕ and ψ are *orthogonal* if $(\phi, \psi) = 0$. The set of eigenfunctions y_n of the Sturm–Liouville problem (1.1.16)–(1.1.17) are mutually orthogonal, or

$$(y_n, y_m) \equiv \int_0^1 y_n(x) y_m(x) dx = 0, \quad n \neq m.$$

Therefore, if we multiply (1.1.18) by a fixed but arbitrary y_m and formally integrate over the interval I , we then obtain

$$(f, y_m) = \sum_{n=1}^{\infty} c_n (y_n, y_m).$$

Because of orthogonality, the infinite series on the right side collapses to the single term $c_m (y_m, y_m)$. Therefore the coefficient c_m is given by

$$c_m = \frac{(f, y_m)}{(y_m, y_m)}.$$

This relation is true for any m , and so the coefficients c_n are

$$c_n = \frac{(f, y_n)}{(y_n, y_n)}, \quad n = 1, 2, \dots \quad (1.1.19)$$

Therefore, we have obtained the solution of (1.1.13)–(1.1.15) in the form of a series representation, or eigenfunction expansion,

$$u(x, t) = \sum_{n=1}^{\infty} \frac{(f, y_n)}{(y_n, y_n)} e^{-\lambda_n t} y_n(x).$$

The preceding calculation took a lot for granted, but it can be shown rigorously that the steps are valid. \square

An expansion of a function $f(x)$ in terms of the eigenfunctions $y_n(x)$, as in (1.1.18), is called the generalized *Fourier series* for f , and the coefficients c_n , given by (1.1.19), are the *Fourier coefficients*. It can be shown that that the series converges in the mean-square sense:

$$\int_0^1 \left(f(x) - \sum_{n=1}^N c_n y_n(x) \right)^2 dx \rightarrow 0 \quad \text{as } N \rightarrow \infty.$$

Pointwise and uniform convergence theorems require suitable smoothness conditions on the function f .

The method of separation of variables is successful under general boundary conditions of the form

$$\alpha u(0, t) + \beta u_x(0, t) = 0, \quad \gamma u(1, t) + \delta u_x(1, t) = 0,$$

where α , β , γ , and δ are given constants. Of course, the interval over which the problem is defined may be any *bounded* interval $a \leq x \leq b$; we chose $a = 0$ and $b = 1$ for simplicity of illustration. The method may be extended to problems over higher-dimensional, bounded, spatial domains, as well as to nonhomogeneous problems. For example, if the PDE in (1.1.13)–(1.1.15) is replaced by the nonhomogeneous equation

$$u_t = Au + F(x, t), \quad 0 < x < 1, \quad t > 0,$$

we can expand the nonhomogeneous term F as a Fourier series of the eigenfunctions for the homogenous problem, or

$$F(x, t) = \sum_{n=1}^{\infty} \gamma_n(t) y_n(x),$$

where the $\gamma_n(t)$ are the known Fourier coefficients (t is a parameter in the expansion) that can be computed from orthogonality property of the eigenfunctions. Then we assume the solution takes the form

$$u(x, t) = \sum_{n=1}^{\infty} c_n(t) y_n(x).$$

Substituting these forms into the PDE and the initial condition determines the $c_n(t)$ and therefore the solution to the nonhomogeneous problem.

Problems that are defined over infinite spatial domains require different techniques based on transform methods.

Example. (Transform Method) Consider the following problem on an infinite spatial domain:

$$\begin{aligned} u_t &= u_{xx}, \quad x > 0, \quad t > 0, \\ u(0, t) &= 0, \quad t > 0, \\ u(x, 0) &= f(x), \quad x > 0. \end{aligned}$$

Because there are two derivatives with respect to x , we expect to impose another boundary condition at infinity. Therefore, we demand that u be bounded as $x \rightarrow \infty$. Further, we assume that f is piecewise continuous and absolutely integrable over $x > 0$. We can proceed as in the preceding example and try a solution of the form $u(x, t) = g(t)y(x)$, where g and y are to be determined. Substituting into the differential equation leads to

$$g' = -\lambda g, \quad -y'' = \lambda y,$$

where λ is the separation constant. As before, $g(t) = ce^{-\lambda t}$, where c is an arbitrary constant. The boundary conditions imply that y is bounded and $y(0) = 0$. Thus we have the boundary value problem

$$\begin{aligned} -y'' &= \lambda y, \quad x > 0, \\ y(0) &= 0, \quad y \text{ bounded}. \end{aligned}$$

This is a boundary value problem on the semi-infinite domain $x > 0$. If $\lambda \leq 0$ there are no nontrivial, bounded solutions (check this), and therefore $\lambda > 0$. Let us write $\lambda = k^2$; then the general solution to the boundary value problem is

$$y(x) = a \sin kx,$$

where a is an arbitrary constant and $k > 0$. Consequently we have found a family of solutions depending upon a parameter k :

$$u(x, t, k) = ae^{-k^2 t} \sin kx.$$

In contrast to the last example, on a bounded interval, where the eigenvalues were discrete, the eigenvalues in the present case form a continuum. We superimpose these solutions over all k and write

$$u(x, t) = \int_0^\infty a(k) e^{-k^2 t} \sin kx \, dk,$$

where $a(k)$ is a function of k . We can determine $a(k)$ from the initial condition. Putting $t = 0$ in the last equation gives

$$f(x) = \int_0^\infty a(k) \sin kx \, dk. \quad (1.1.20)$$

This is an integral equation from which we can recover $a(k)$ using a special case of the Fourier integral theorem: If f is piecewise continuous and absolutely integrable on $x > 0$, and if

$$f(x) = \int_0^\infty a(k) \sin kx \, dk,$$

then

$$a(k) = \frac{2}{\pi} \int_0^\infty f(\xi) \sin k\xi \, d\xi.$$

The function a is the Fourier sine transform of f , and f is the inverse sine transform of a . Putting everything together gives an integral representation of the solution to the problem, namely

$$\begin{aligned} u(x, t) &= \frac{2}{\pi} \int_0^\infty \left(\int_0^\infty f(\xi) \sin k\xi \, d\xi \right) e^{-k^2 t} \sin kx \, dk \\ &= \frac{2}{\pi} \int_0^\infty f(\xi) \left(\int_0^\infty e^{-k^2 t} \sin k\xi \sin kx \, dk \right) d\xi, \end{aligned}$$

where we have changed the order of integration in the last step. Actually, the interior integral can be calculated analytically, or looked up in a table, which we leave as an exercise. \square

The reader should notice the great similarity of the solution forms in these two examples—finite domain versus infinite domain, discrete eigenvalues versus a continuum of eigenvalues, and expansions in terms of sums versus integrals. The role of superposition is critical in linear problems, but it does not carry over to nonlinear problems.

Terminology. We introduce some notation and terminology for some function spaces that commonly occur in analysis. Let D be an open domain (a set that does not contain any of its boundary) in either one or several dimensions. Because D is an open domain we do not have to deal with the question of

existence of derivatives at boundary points. Using an overbar, \overline{D} , we denote the closure of D , which consists of D and its boundary ∂D ; that is, $\overline{D} = D \cup \partial D$. By $C^n(D)$ we denote the set of all continuous functions on D that have n continuous derivatives in D (partial derivatives if D is of dimension greater than 1). The space of continuous functions on D is denoted by $C(D)$, and $C^\infty(D)$ denotes the space of continuous functions on D that have derivatives of all orders. When a function u belongs to one of these sets, e.g., $C^n(D)$, we sometimes say that u is of class C^n on D .

EXERCISES

1. Show that the wave equation $u_{tt} - c^2 u_{xx} = 0$ reduces to the canonical form $u_{\xi\eta} = 0$ under the change of variables $\xi = x - ct$, $\eta = x + ct$, and use this information to show that the general solution of the wave equation is $u(x, t) = F(x-ct) + G(x+ct)$, where F and G are arbitrary $C^2(\mathbb{R})$ functions.
2. Derive D'Alembert's formula for the initial value problem for the wave equation using Exercise 1 and determining F and G from the initial conditions.
3. Let u be of class C^3 . Show that $u = u(x, t)$ is a solution of the wave equation

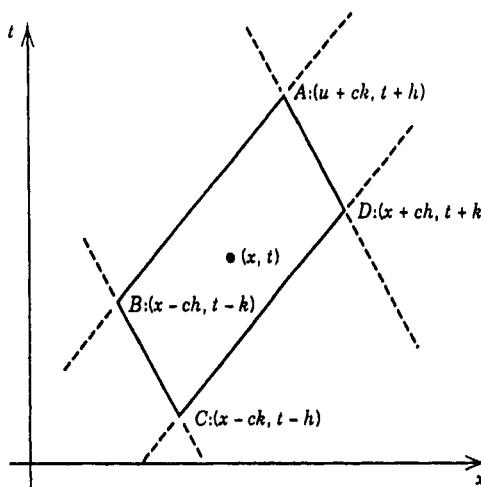


Figure 1.4 Characteristic parallelogram $ABCD$ whose sides are characteristic straight lines $x - ct = \text{const}$ and $x + ct = \text{const}$. The numbers h and k are positive constants that define the size of $ABCD$.

$u_{tt} - c^2 u_{xx} = 0$ if, and only if, u satisfies the difference equation

$$\begin{aligned} & u(x - ck, t - h) + u(x + ck, t + h) \\ &= u(x - ch, t - k) + u(x + ch, t + k) \end{aligned}$$

for all constants $h, k > 0$. Interpret this result geometrically in the xt -plane by observing that the difference equation relates the value of u at the vertices of a characteristic parallelogram whose sides are the characteristic straight lines $x + ct = \text{const}$ and $x - ct = \text{const}$ (see Figure 1.4).

4. Assuming that the initial data f and g for the initial value problem for the wave equation in Exercise 3 have compact support (i.e., f and g vanish for $|x|$ sufficiently large), prove that the solution $u(x, t)$ has compact support in x for each fixed time t .

5. Find the general solution of the PDE

$$x^2 u_{xx} + 2xtu_{xt} + t^2 u_{tt} = 0$$

by transforming the equation to canonical form using characteristic coordinates $\xi = t/x$, $\eta = t$.

6. Let $u = u(x, t)$ be a solution of the nonlinear equation

$$a(u_x, u_t)u_{xx} + 2b(u_x, u_t)u_{xt} + c(u_x, u_t)u_{tt} = 0.$$

Introduce new independent variables via $\xi = \xi(x, t)$ and $\eta = \eta(x, t)$, and a new function $\phi = \phi(\xi, \eta)$ defined by $\phi = xu_x + tu_t - u$. Prove that $\phi_\xi = x$, $\phi_\eta = t$, and ϕ satisfies the *linear* PDE

$$a(\xi, \eta)\phi_{\eta\eta} - 2b(\xi, \eta)\phi_{\xi\eta} + c(\xi, \eta)\phi_{\xi\xi} = 0.$$

(This transformation, known as a *hodograph*, or a *Legendre*, transformation, transforms a nonlinear equation of the given form to a linear equation by reversing the roles of the dependent and independent variables.)

7. Find a formula for the solution of the initial-boundary value problem

$$\begin{aligned} u_{tt} &= c^2 u_{xx}, \quad x > 0, \quad t > 0, \\ u(x, 0) &= f(x), \quad u_t(x, 0) = g(x), \quad x > 0, \\ u(0, t) &= h(t), \quad t > 0. \end{aligned}$$

Hint: Use D'Alembert's formula for $x > ct$ and the difference equation in Exercise 3 for $0 < x < ct$. Assume sufficient differentiability.

8. Solve the outgoing signaling problem

$$\begin{aligned} u_{tt} &= c^2 u_{xx}, \quad x > 0, \quad t \in \mathbb{R}, \\ u(0, t) &= s(t), \quad t \in \mathbb{R}. \end{aligned}$$

9. Consider the PDE

$$4u_{xx} + 5u_{xt} + u_{tt} = 2 - u_t - u_x.$$

Find the characteristic coordinates and graph the characteristic curves in the xt plane. Reduce the equation to canonical form and find the general solution.

10. Classify the PDE

$$xu_{xx} - 4u_{tt} = 0.$$

In the case $x > 0$ find the characteristic coordinates and sketch the characteristics in an appropriate region in the xt plane.

11. Use the separation of variables method to find a series representation of the solution to the following problems:

(a)

$$\begin{aligned} u_t &= Du_{xx}, \quad 0 < x < \pi, \quad t > 0, \\ u(0, t) &= 0, \quad u(\pi, t) = 0, \quad t > 0, \\ u(x, 0) &= f(t), \quad 0 < x < \pi. \end{aligned}$$

(b)

$$\begin{aligned} u_{tt} &= c^2 u_{xx}, \quad 0 < x < 1, \quad t > 0, \\ u(0, t) &= 0, \quad u(1, t) = 0, \quad t > 0, \\ u(x, 0) &= 0, \quad u_t(x, 0) = g(t), \quad 0 < x < \pi. \end{aligned}$$

(c)

$$\begin{aligned} u_{xx} + u_{yy} &= 0, \quad 0 < x < 1, \quad 0 < y < \pi, \\ u(x, 0) &= 0, \quad u_y(x, \pi) = 0, \quad 0 < x < 1, \\ u(0, y) &= 0, \quad u(1, y) = g(y), \quad 0 < y < \pi. \end{aligned}$$

12. Find the general solution of the Euler–Darboux equation

$$u_{x,y} = \frac{m}{x-y}(u_x - u_y)$$

in the case $m = 1$. Hint: Look at $((x-y)u)_{xy}$.

1.2 Conservation Laws

Many of the fundamental equations in the natural and physical sciences are obtained from *conservation laws*, which are balance laws, equations expressing the fact that some quantity is balanced throughout a process. In thermodynamics, for example, the first law states that the change in internal energy in a system is equal to, or is balanced by, the total heat added to the system plus the work done on the system. Thus the first law of thermodynamics is really an energy balance law, or conservation law. As another example, consider a fluid flowing in some region of space that consists of chemical species undergoing chemical reaction. For a given chemical species, the time rate of change of the total amount of that species in the region must equal the rate at which the species flows into the region, minus the rate at which the species flows out, plus the rate at which the species is created, or consumed, by the chemical reactions. This is a verbal statement of a conservation law for the amount of the given chemical species. Similar balance or conservation laws occur in all branches of science. In the biosciences, for example, the rate of change of an animal population in a fixed region must equal the birth rate, minus the death rate, plus the migration rate (emigration or immigration) into or out of the region.

1.2.1 One Dimension

Mathematically, conservation laws translate into integral or differential equations, which are then regarded as the *governing equations* or *equations of motion* of the process. These equations dictate how the process evolves in time. Here we are interested in processes governed by partial differential equations. We now formulate the basic one-dimensional conservation law, out of which will evolve some of the basic models and concepts in nonlinear PDEs.

Let us consider a quantity $u = u(x, t)$ that depends on a single spatial vari-

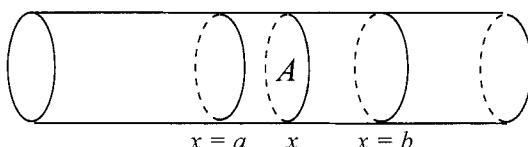


Figure 1.5 Cylindrical tube of cross-sectional area A showing a cross section at x and a finite section I : $a \leq x \leq b$. There is no variation of any quantity in a fixed cross section.

able $x \in \mathbb{R}$ and time $t > 0$. We assume that u is a density or concentration measured in an amount per unit volume, where the amount may refer to population, mass, energy, or any quantity. By definition, u varies in only one spatial direction, the direction denoted by x . We imagine further that the quantity is distributed in a tube of cross-sectional area A (see Figure 1.5). Again, by assumption, u is constant in any cross section of the tube, and the variation is only in the x direction. Now consider an arbitrary segment of the tube denoted by the interval $I = [a, b]$. The total amount of the quantity u inside I at time t is

$$\text{Total amount of quantity in } I = \int_a^b u(x, t) A dx.$$

Now assume that there is motion of the quantity in the tube in the axial direction. We define the *flux* $\phi(x, t)$ of u at x at time t as the amount of the quantity u flowing through the cross section at x at time t , per unit area, per unit time. Thus the dimensions of ϕ are $[\phi] = \text{amount}/(\text{area} \cdot \text{time})$, where the bracket notation denotes *dimensions of*. By convention, we take ϕ to be positive if the flow at x is in the positive x direction, and to be negative at x if the flow is in the negative x direction. Therefore, at time t the net rate that the quantity is flowing into the interval I is the rate that it is flowing in at $x = a$ minus the rate that it is flowing out at $x = b$:

$$\text{Net rate that the quantity flows into } I = A\phi(a, t) - A\phi(b, t).$$

Finally, the quantity u may be created or destroyed inside I by an external or internal source (e.g., by a chemical reaction if u were a species concentration, or by birth or death if u were a population density). We denote this *source function*, which is a local function acting at each x , by $f(x, t, u)$ and its dimensions are given by $[f] = \text{amount}/(\text{volume} \cdot \text{time})$. (The source f could also depend on derivatives of u .) Consequently, f is the rate that u is created (or destroyed) at x at time t , per unit volume. Note that the source function f may depend on u itself, as well as space and time. If f is positive, we say that it is a *source*, and if f is negative, we say that it is a *sink*. Now, given f , we may calculate the total rate that u is created in I by integration. We have

$$\text{Rate that quantity is produced in } I \text{ by sources} = \int_a^b f(x, t, u(x, t)) A dx.$$

The fundamental conservation law may now be formulated for the quantity u : For any interval I , we have

$$\begin{aligned} &\text{Time rate of change of the total amount in } I \\ &= \text{net rate that the quantity flows into } I \\ &\quad + \text{rate that the quantity is produced in } I. \end{aligned}$$

In terms of the mathematical symbols and expressions that we introduced above, after canceling the constant cross-sectional area A , we have

$$\frac{d}{dt} \int_a^b u(x, t) dx = \phi(a, t) - \phi(b, t) + \int_a^b f(x, t, u) dx. \quad (1.2.1)$$

In summary, (1.2.1) states that the rate that u changes in the interval I must equal the net rate at which u flows into I plus the rate that u is produced in I by sources. Equation (1.2.1) is called a *conservation law in integral form*, and it holds even if u , ϕ , or f is not a smooth (continuously differentiable) function. The latter remark is important when we consider physical processes giving rise to shock waves, or discontinuous solutions, in subsequent chapters.

If conditions are placed on the triad u , ϕ , and f , then (1.2.1) may be transformed into a single PDE. Two results from elementary integration theory are required to make this transformation: (1) the fundamental theorem of calculus, and (2) the result on differentiating an integral with respect to a parameter in the integrand. Precisely:

1. $\int_a^b \phi_x(x, t) dx = \phi(b, t) - \phi(a, t).$
2. $d/dt \int_a^b u(x, t) dx = \int_a^b u_t(x, t) dx.$

These results are valid if ϕ and u are continuously differentiable functions on \mathbb{R}^2 . Of course, (1) and (2) remain correct under less stringent conditions, but the assumption of smoothness is all that is required in the subsequent discussion. Therefore, assuming smoothness of u and ϕ , as well as continuity of f , equations (1) and (2) imply that the conservation law (1.2.1) may be written

$$\int_a^b [u_t(x, t) + \phi_x(x, t) - f(x, t, u)] dx = 0 \quad \text{for all intervals } I = [a, b]. \quad (1.2.2)$$

Because the integrand is a continuous function of x , and because (1.2.2) holds for all intervals of integration I , it follows that the integrand must vanish identically:

$$u_t + \phi_x = f(x, t, u), \quad x \in \mathbb{R}, \quad t > 0. \quad (1.2.3)$$

Equation (1.2.3) is a PDE relating the density $u = u(x, t)$ and the flux $\phi = \phi(x, t)$. Both are regarded as unknowns, whereas the form of the source function f is assumed to be given. Equation (1.2.3) is called a *local conservation law*, in contrast to the integral form (1.2.1). The ϕ_x term is called the *flux term* because it arises from the movement, or transport, of u through the cross section at x . The source term f is called a *reaction term* (especially in chemical contexts) or a *growth* or *interaction* term (in biological contexts). Finally, we have defined the flux ϕ as a function of x and t ; this dependence on space and time may occur through dependence on u or its derivatives. For example, a physical assumption

may require us to posit $\phi(x, t) = \phi(x, t, u(x, t))$, where the flux is dependent on u itself.

It is important to observe that (1.2.3) was derived under assumptions of smoothness. If smoothness of the density or flux is not guaranteed, as occurs in the study of discontinuous solutions, then (1.2.3) must be abandoned in favor of the integral form (1.2.1) of the conservation law, which is always valid. This issue becomes the focus of study in Chapter 3.

Finally, we make a general observation about deriving conservation laws. In the preceding discussion the balance law was applied to an entire interval I . Assuming smoothness, we can derive the conservation law directly using a small interval $[x, x + \Delta x]$ and then take the limit as $\Delta x \rightarrow 0$. Applying the balance law to the small box $[x, x + \Delta x]$, we obtain

$$\frac{\partial}{\partial t} (u(\xi, t) A \Delta x) = A\phi(x, t) - A\phi(x + \Delta x, t) + f(t, \eta, u(\eta, t)) A \Delta x,$$

where ξ and η are points in $[x, x + \Delta x]$, guaranteed by the mean value theorem. Now, dividing by $A \Delta x$ gives

$$\frac{\partial}{\partial t} u(\xi, t) = \frac{\phi(x, t) - \phi(x + \Delta x, t)}{\Delta x} + f(t, \eta, u(\eta, t)).$$

Taking the limit as $\Delta x \rightarrow 0$ gives the conservation law (1.2.3). These two methods for obtaining the conservation law are called, appropriately, the *large-box method* and the *small-box method*.

1.2.2 Higher Dimensions

It is straightforward to formulate conservation laws in higher dimensions. In this section we limit the discussion to three-dimensional Euclidean space \mathbb{R}^3 . Let $x = (x_1, x_2, x_3)$ denote a point in \mathbb{R}^3 , and assume that $u = u(x, t)$ is a scalar density function representing the amount per unit volume of some quantity of interest distributed throughout a domain $D \subset \mathbb{R}^3$. In this domain let V be an arbitrary region with a smooth boundary denoted by ∂V . By an argument similar to the one-dimensional case, the total amount of the quantity in V is given by the volume integral

$$\text{Total amount in } V = \int_V u(x, t) dx,$$

where $dx = dx_1 dx_2 dx_3$ represents a volume element in \mathbb{R}^3 . We prefer to write the volume integral over V with a single integral sign rather than the usual triple integral. Now, we know that the time rate of change of the total amount in V must be balanced by the rate that the quantity is produced in V by

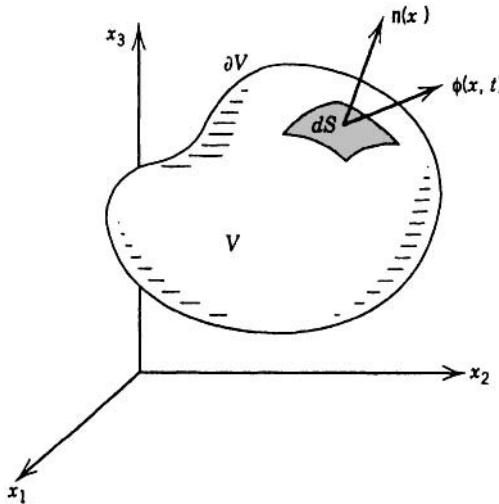


Figure 1.6 Volume V with boundary ∂V showing a surface element dS with outward normal \mathbf{n} and flux vector ϕ . The outward normal determines the orientation of the surface element.

sources, plus the net rate that the quantity flows through the boundary of V . We let $f(x, t, u)$ denote the source term, so that the rate that the quantity is produced in V is given by

$$\text{Rate that } u \text{ is produced by sources} = \int_V f(x, t, u) dx.$$

In three dimensions the flow can be in any direction, and therefore the flux is given by a vector $\phi(x, t)$. If $\mathbf{n}(x)$ denotes the outward unit normal vector to the region V (see Figure 1.6), then the net outward flux of the quantity u through the boundary ∂V is given by the surface integral

$$\text{Net outward flux through } \partial V = \int_{\partial V} \phi(x, t) \cdot \mathbf{n}(x) dS,$$

where dS denotes a surface element on ∂V . Hence, the balance law for u is given by

$$\frac{d}{dt} \int_V u dx = - \int_{\partial V} \phi \cdot \mathbf{n} dS + \int_V f dx. \quad (1.2.4)$$

The minus sign on the flux term occurs because outward flux decreases the rate that u changes in V .

The integral form of the conservation law (1.2.4) can be reformulated as a local condition, that is, a PDE, provided that u and ϕ are sufficiently smooth functions. In this case the surface integral can be written as a volume integral

over V using the *divergence theorem* (the divergence theorem is the fundamental theorem of calculus in three dimensions)

$$\int_V \operatorname{div} \phi \, dx = \int_{\partial V} \phi \cdot \mathbf{n} \, dS, \quad (1.2.5)$$

where div is the divergence operator. Using (1.2.5) and bringing the derivative under the integral on the left side of (1.2.4) yields

$$\int_V u_t \, dx = - \int_V \operatorname{div} \phi \, dx + \int_V f \, dx.$$

The arbitrariness of V then implies the differential form of the balance law:

$$u_t + \operatorname{div} \phi = f(x, t, u), \quad x \in D, \quad t > 0. \quad (1.2.6)$$

Equation (1.2.6) is the three-dimensional version of equation (1.2.3).

EXERCISES

1. The derivation of the fundamental conservation law (1.2.3) assumed the cross-sectional area A of the tube to be constant. Derive integral and differential forms of the conservation law assuming that the area is a slowly varying function of x , that is, $A = A(x)$ and $A'(x)$ is small. [Note that $A(x)$ cannot change significantly over small changes in x ; otherwise, the one-dimensional assumption of the state functions u and ϕ being constant in any cross section would be violated.]
2. Assuming that there are no sources and $\phi = \phi(u)$, show that the conservation law (1.2.1) is equivalent to

$$\begin{aligned} \int_a^b u(x, t_2) \, dx &= \int_a^b u(x, t_1) \, dx + \\ &\int_{t_1}^{t_2} \phi(u(a, t)) \, dt - \int_{t_1}^{t_2} \phi(u(b, t)) \, dt \end{aligned}$$

for all t_1 and t_2 .

1.3 Constitutive Relations

Because equation (1.2.3), or (1.2.6), is a single PDE for two unknown quantities (the density u and the flux ϕ), intuition indicates that another equation is required to have a well-determined system. This additional equation is a relation that is usually based on an assumption about the physical properties of the

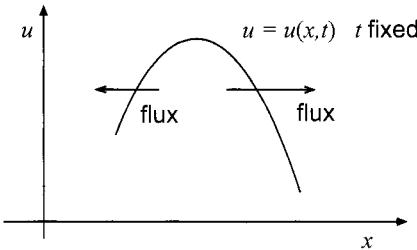


Figure 1.7 Time snapshot of the density distribution $u(x, t)$ illustrating Fick's law. The arrows indicate the direction of the flow, from higher concentrations to lower concentrations. The flow is said to be *down the gradient*.

medium, or the processes involved, which in turn is based on empirical reasoning. Equations expressing these assumptions are called *constitutive relations* or *equations of state*. Thus, a constitutive equation is on a different level from the basic conservation law; the latter is a fundamental law of nature connecting the density u to the flux ϕ , whereas a constitutive relation is an approximate equation whose origin is in empirics. We present several key examples.

Example. (Diffusion Equation) At the outset assume that no sources are present ($f = 0$) and the process is governed by the basic conservation law in one dimension

$$u_t + \phi_x = 0, \quad x \in \mathbb{R}, \quad t > 0. \quad (1.3.1)$$

In many physical processes it is observed that the amount of the substance, represented by its density u , flowing through a cross section at x at time t is proportional to the density gradient u_x ; that is, $\phi(x, t) \propto u_x(x, t)$. If $u_x > 0$, then $\phi < 0$ and the substance flows to the left, and if $u_x < 0$, then $\phi > 0$ and the substance flows to the right. Figure 1.7 illustrates the situation. We say that u flows *down the gradient*. For example, the second law of thermodynamics states that heat behaves in this manner; heat flows from hotter regions to colder regions, and the steeper the temperature distribution curve, the more rapid the flow of heat. As another example, if u represents a concentration of insects, one may observe that insects move from high concentrations to low concentrations with a rate proportional to the concentration gradient. Therefore, we assume the basic constitutive law

$$\phi(x, t) = -Du_x(x, t), \quad (1.3.2)$$

which is known as *Fick's law*. Processes described by this law are called *linear diffusion processes*. The positive proportionality constant D is called the *diffusion constant*, and it has dimensions $[D] = \text{length}^2/\text{time}$. Fick's law accurately

describes the behavior of many physical and biological systems, and in Chapter 5 we give a supporting argument for it on the basis of a probability model and random walk.

Equations (1.3.1) and (1.3.2) give a pair of PDEs for the two unknowns u and ϕ . They combine easily to form a single second-order linear PDE for the unknown density $u = u(x, t)$ given by

$$u_t - Du_{xx} = 0. \quad (1.3.3)$$

Equation (1.3.3), called the *diffusion equation*, governs conservative processes when the flux is specified by Fick's law. It may not be clear at this time why (1.3.3) should be termed the diffusion equation; suffice it to note for the moment that Fick's law implies that a substance moves into adjacent regions because of concentration gradients. We refer to the Du_{xx} term in (1.3.3) as the *diffusion term*. \square

The diffusion constant D defines a crude characteristic time (or time scale) T for the process. If L is a length scale (e.g., the length of the container), the quantity

$$T = \frac{L^2}{D} \quad (1.3.4)$$

is the only constant in the process with dimensions of time, and T gives a measure of the time required for discernible changes in concentration to occur over the length L .

Example. (Heat Equation) If $u = u(x, t)$ is the thermal energy density (energy per volume) in a heat-conducting medium, then $u = \rho CT$, where ρ is the mass density (mass per unit volume), C is the specific heat (energy per unit mass per degree), and $T = T(x, t)$ is the temperature (degrees). In the absence of sources the conservation law is

$$(\rho CT)_t + \phi_x = 0,$$

where ϕ is the energy flux. The constitutive law, which is a heat analog of Fick's law, is

$$\phi = -KT_x(x, t),$$

where K is the thermal conductivity of the region. In heat transfer this is called *Fourier's law of heat conduction*. Thus, heat energy moves down the temperature gradient. The conservation law becomes

$$T_t - kT_{xx} = 0,$$

where $k = K/\rho C$ is the diffusivity. This is the *heat equation*, which is just the diffusion equation in the context of heat flow. \square

Example. (*Reaction–Diffusion Equation*) If sources are present ($f \neq 0$), the conservation law

$$u_t + \phi_x = f(x, t, u) \quad (1.3.5)$$

and Fick's law (1.3.2) combine to give

$$u_t - Du_{xx} = f(x, t, u), \quad (1.3.6)$$

which is called a *reaction–diffusion equation*. Reaction–diffusion equations are nonlinear if the reaction term f is nonlinear in u . These equations are of great interest in nonlinear analysis and applications, particularly in combustion processes and in biological systematics. \square

Example. (*Fisher's Equation*) In studies of elementary population dynamics, one proposal is that populations are governed by the logistic law, which states that the rate of change of the total population $u = u(t)$ in a fixed spatial domain is given by

$$\frac{du}{dt} = ru \left(1 - \frac{u}{K}\right), \quad (1.3.7)$$

where $r > 0$ is the *growth rate* and $K > 0$ is the *carrying capacity*. Initially, if u is small, the linear growth term ru in (1.3.7) dominates and rapid population growth results; as u becomes large, the quadratic competition term $-ru^2/K$ kicks in to inhibit the growth. For large times t , the population equilibrates toward the asymptotically stable state $u = K$, the carrying capacity. Now we add spatial effects. Suppose that the population u is a population density and depends on a spatial variable x as well as time t ; that is, $u = u(x, t)$. Then, as in the preceding discussion, a conservation law may be formulated as

$$u_t + \phi_x = ru \left(1 - \frac{u}{K}\right), \quad (1.3.8)$$

where $f = f(u) = ru(1 - u/K)$ is the assumed local source term given by the logistics growth law, and ϕ is the population flux. Assuming Fick's law for the flux, we have

$$u_t - Du_{xx} = ru \left(1 - \frac{u}{K}\right). \quad (1.3.9)$$

The reaction–diffusion equation (1.3.9) is *Fisher's equation*, after R. A. Fisher, who studied the equation in the context of investigating the distribution of an advantageous gene as it diffuses through a given population. Discussion of this equation, which is one of the fundamental equations in mathematical biology, is given in subsequent chapters. \square

Example. (*Burgers' Equation*) To the basic conservation law with no sources we now append the constitutive relation

$$\phi = -Du_x + Q(u). \quad (1.3.10)$$

The density u then satisfies

$$u_t - Du_{xx} + Q(u)_x = 0. \quad (1.3.11)$$

Now there are two terms contributing to the flux, a Fick's law type of term $-Du_x$, which introduces a diffusion effect, and a flux term $Q(u)$, depending only on u itself, that leads to what is interpreted later as advection. In the special case that $Q(u) = u^2/2$, equation (1.3.11) can be written

$$u_t + uu_x = Du_{xx}, \quad (1.3.12)$$

which is *Burgers' equation*. This is one of the fundamental model equations in fluid mechanics and illustrates the coupling between advection and diffusion. Later we derive (1.3.12) using a weakly nonlinear approximation of the equations of gas dynamics. When $D = 0$ (no diffusion), then

$$u_t + uu_x = 0, \quad (1.3.13)$$

which is called the *inviscid Burgers' equation* (also, Riemann's equation). It is the prototype equation for nonlinear advection and arises in gas dynamics, traffic flow, chromatography, and flood waves in rivers. Equation (1.3.13) is hyperbolic and describes wave propagation, whereas (1.3.12) is parabolic and models diffusion. \square

Example. (*Advection Equation*) The simplest flux term occurs when the material forming the density is carried along by the medium having a fixed velocity, as in the case of particulates carried by, for example, wind or water. In these cases the flux is given by the simple linear relationship

$$\phi = cu, \quad (1.3.14)$$

where c is a positive constant having the dimensions of speed. The basic conservation law is

$$u_t + cu_x = 0, \quad (1.3.15)$$

which is the *advection equation*. The term *advection* refers to the horizontal movement of a physical property (e.g., a density wave); other equivalent terms are *convection* and *transport*, which have the same meaning. For example, biologists use the term advection while many engineers use the term convection. We use these terms interchangeably. This linear first-order PDE (1.3.15) is the simplest wave equation. Figure 1.8 compares how a signal at time $t = 0$ propagates under advection, diffusion, and advection-diffusion processes. \square

Example. (*Diffusion in \mathbb{R}^3*) In Section 1.2 we obtained the basic conservation law

$$u_t + \operatorname{div} \phi = f(x, t, u).$$

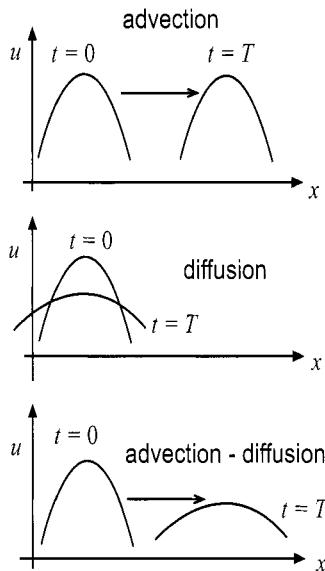


Figure 1.8 Comparison of advection, diffusion, and advection–diffusion processes.

In \mathbb{R}^3 Fick's law takes the form $\phi = -D \operatorname{grad} u$; that is, the flux is in the direction of the negative gradient of u . Recall the direction of maximum increase is in the direction of the gradient. Using the vector identity $\operatorname{div}(\operatorname{grad} u) = \Delta u$, where

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}$$

is the *Laplacian* operator, and assuming that the diffusion coefficient D is constant, the conservation law becomes

$$u_t - D\Delta u = f(x, t, u),$$

which is a reaction–diffusion equation in \mathbb{R}^3 . If there are no sources ($f = 0$), we obtain the three-dimensional diffusion equation $u_t - D\Delta u = 0$. \square

In summary, we introduced a few of the fundamental PDEs that have been examined extensively by practitioners of the subject. An understanding of these equations and a development of intuition regarding the fundamental processes that they, and other equations, describe is one of the goals of this treatment. We end this section with the problem of modeling flow through a porous medium.

Example. (Porous Media) Consider a fluid (e.g., water) seeping downward through the soil, and let $\rho = \rho(x, t)$ be the density of the fluid, with positive

x measured downward. In a given volume of soil only a fraction of the space is available to the fluid; the remaining space is reserved for the soil itself. In this sense, the soil is a porous medium through which the fluid flows. If the fraction of the volume that is available to the fluid is denoted by κ , called the *porosity*, the fluid mass balance law in a given section of unit cross-sectional area between $x = a$ and $x = b$ is given by

$$\frac{d}{dt} \int_a^b \kappa \rho(x, t) dx = \phi(a, t) - \phi(b, t), \quad (1.3.16)$$

where ϕ is the mass flux. We assume that the mass flux is $\phi = \rho v$, where v is the volumetric flow rate. Assuming requisite smoothness, the integral balance law (1.3.16) can be written

$$\kappa \rho_t + (\rho v)_x = 0. \quad (1.3.17)$$

Here we assumed that the porosity is constant, but in general it could depend on x or even ρ . The conservation law (1.3.17) contains two unknowns, the density and the flow rate, and therefore we need a constitutive equation that relates the two. For reasonably slow flows it is observed experimentally that the volumetric flow rate is given by

$$v = -\frac{\mu}{\nu}(p_x + g\rho), \quad (1.3.18)$$

where $p = p(x, t)$ is the pressure, and the positive constants g , μ , and ν are the acceleration due to gravity, the permeability, and the viscosity of the fluid, respectively. Equation (1.3.18) is *Darcy's law*, and it is a basic assumption in many groundwater problems. It is physically plausible because it confirms our intuition that the flow rate should depend on the pressure gradient as well as gravity. Darcy's law is a statement replacing a momentum balance law. When (1.3.18) is substituted into (1.3.17), we obtain the *groundwater equation*:

$$\kappa \rho_t - \frac{\mu}{\nu}(\rho p_x + g\rho^2)_x = 0. \quad (1.3.19)$$

Equation (1.3.19) contains two unknowns, ρ and p , because the constitutive relation (1.3.18) introduced yet another unknown, the pressure. Consequently, we require another equation to obtain a determined system.

If the fluid is a gas, it is common to neglect gravity and assume an equation of state for the gas of the form $p = p(\rho)$. In particular, we assume a γ -law gas having equation of state

$$\frac{p}{p_0} = \left(\frac{\rho}{\rho_0} \right)^\gamma, \quad (1.3.20)$$

where p_0 and ρ_0 are positive constants and $\gamma > 1$. Substituting (1.3.20) into (1.3.19) and setting $g = 0$ yields a single PDE for the density ρ having the form

$$\rho_t - \frac{\alpha}{\kappa}(\rho^\gamma \rho_x)_x = 0, \quad (1.3.21)$$

where $\alpha = \gamma \mu p_0 / \nu \rho_0^\gamma$. Equation (1.3.21) is a nonlinear diffusion equation called the *porous medium equation*, and it governs flows subject to three laws: mass conservation, Darcy's law, and the gas equation of state. See Aronson (1986) for a survey on the porous medium equation. \square

Finally, we point out that it is often crucial to nondimensionalize a problem. Differential equations, when formulated, involve dimensioned dependent and independent variables such as time, length, and temperature. The Buckingham Pi theorem guarantees that a dimensionally consistent physical law can always be transformed to one where the variables, as well as the parameters, are dimensionless. A valid comparison of the relative magnitudes of the terms in an equation can be made only when a problem is nondimensionalized. Further, the dimensionless problem often offers an economy over the dimensioned version in that there is a reduction in the number of parameters. The process of nondimensionalization is sometimes called *scaling*. Lin & Segel (1974) and Logan (2006a) thoroughly discuss scaling and dimensional analysis.

Example. (Scaling) In Fisher's equation,

$$u_t = Du_{xx} + ru \left(1 - \frac{u}{K}\right),$$

the variables t , x , and u have dimensions of time, length, and animals per area, respectively. The parameters are the growth rate r with units of 1/time, the carrying capacity K with units of animals per area, and the diffusion constant D with dimensions length-squared per time. Using the parameters we can build dimensionless variables by defining

$$\tau = \frac{t}{r^{-1}}, \quad \xi = \frac{x}{\sqrt{D/r}}, \quad v = \frac{u}{K}.$$

Note that each has the form of a dimensioned variable divided by a constant with the same dimension. We refer to these constants in the denominator as *scales*. For example, the population density is scaled by K , which means that the population is being measured relative to the carrying capacity; r^{-1} is the time scale, meaning that time is being measured relative to the (inverse) growth rate, and so on. There are usually several ways to determine the scales. By the chain rule we can transform the PDE into the dimensionless variables. Observe that derivatives in the PDE transform via

$$\frac{\partial u}{\partial t} = rK \frac{\partial v}{\partial \tau}, \quad \frac{\partial^2 u}{\partial x^2} = \frac{K}{D/r} \frac{\partial^2 v}{\partial \xi^2},$$

and therefore the PDE becomes

$$rK \frac{\partial v}{\partial \tau} = D \frac{K}{D/r} \frac{\partial^2 v}{\partial \xi^2} + rKv(1-v),$$

or

$$v_\tau = v_{\xi\xi} + v(1-v).$$

Therefore, in dimensionless variables Fisher's equation reduces to a model equation without any constants at all. This simpler equation may be analyzed, and, if required, a return to interpretations in terms of the original dimensioned quantities can be made. \square

EXERCISES

1. Write the PDE

$$u_t + uu_x + u_{xxx} = 0$$

in the form of a conservation law, identifying the flux ϕ . Given u as a solution for which u , u_x , and u_{xx} approach zero as $|x| \rightarrow \infty$, show that

$$\int_{-\infty}^{\infty} u(x, t) dx = \text{const}, \quad \int_{-\infty}^{\infty} u^2(x, t) dx = \text{const},$$

for all $t > 0$.

2. In three dimensions assume show that advection should be modeled by the equation

$$u_t + \operatorname{div} \mathbf{c} u = 0,$$

where $\mathbf{c} = \mathbf{c}(\mathbf{x})$ is the velocity of the medium. Given \mathbf{c} as a constant vector, show that $u = f(\mathbf{x} - \mathbf{c}t)$ is a solution for any real-valued differentiable function f .

3. Show that Burgers' equation (1.3.12) can be transformed into the linear diffusion equation (1.3.3) by the *Cole-Hopf transformation*

$$u = -\frac{2Dv_x}{v}.$$

4. Show that the PDE

$$u_t + kuu_x + q(t)u = 0$$

can be reduced to the inviscid Burgers' equation

$$v_s + vv_x = 0$$

using the transformation

$$v = u \exp \int q(t) dt, \quad s = \int k \exp \left(- \int q(y) dy \right) dt.$$

Show that the same transformation transforms

$$u_t + kuu_x + q(t)u - Du_{xx} = 0$$

into

$$v_s + vv_x - g^{-1}(s)v_{xx} = 0,$$

where $g = (k/D) \exp(-\int q(t) dt)$.

5. By rescaling, show that the porous media equation can be written in the form

$$u_\tau = (u^m)_{\xi\xi} \quad (m > 2)$$

for appropriately chosen dimensionless variables τ , ξ , and u .

6. Show that the nonlinear growth-diffusion equation

$$S_t = k(S^3)_{xx} + aS$$

can be reduced to the porous medium equation by the transformations

$$S = \rho(x, t)e^{at}, \quad \tau = \frac{1}{2a}e^{2at}.$$

7. Consider a porous medium where the fluid is water, and assume that the density ρ is constant. What equation must the pressure p satisfy? Describe the pressure distribution.
8. Nondimensionalize the growth-advection-diffusion equation

$$u_t = Du_{xx} - cu_x + ru.$$

9. The population density $u(x, t)$ of zooplankton in a deep lake varies as a function of depth $x > 0$ and time t ($x = 0$ is the surface). Zooplankton diffuse with diffusion constant D , and buoyancy effects cause them to migrate toward the surface with an advection speed of αg , where g is the acceleration due to gravity. Ignore birth and death rates.
- (a) Find a PDE model for the population density of zooplankton in the lake, along with the appropriate boundary conditions at $x = 0$ and $x = +\infty$.
- (b) Find the steady-state population density $u = U(x)$ for zooplankton as a function of depth, and sketch its graph.

10. Let $u = u(x, y, t)$ satisfy the following PDE and boundary conditions:

$$\begin{aligned}\varepsilon^2 u_{xx} + u_{yy} &= 0, \quad 0 < y < 1, x \in \mathbb{R}, t > 0, \\ u_y(x, 0, t) &= 0, \quad u_y(x, 1, t) + \varepsilon^2 u_{tt}(x, 1, t) = 0, \\ x \in \mathbb{R}, t > 0,\end{aligned}$$

where ε is small. Assuming a perturbation expansion

$$u = u_0(x, t) + u_1(x, y, t)\varepsilon^2 + \dots,$$

show that u_0 satisfies the wave equation.

1.4 Initial and Boundary Value Problems

So far we encountered several types of PDEs governing different types of physical processes:

$$\begin{array}{ll}u_t = Du_{xx} & \text{(diffusion)} \\u_t = Du_{xx} + f(x, t, u) & \text{(reaction-diffusion)} \\u_t + cu_x = 0 & \text{(advection)} \\u_t + cu_x = Du_{xx} & \text{(advection-diffusion)} \\u_t + uu_x = Du_{xx} & \text{(nonlinear advection-diffusion)}\end{array}$$

As we noted earlier, we are seldom interested in the general solution to a PDE, which contains arbitrary functions. Rather, we are interested in solving the PDE subject to auxiliary conditions such as initial conditions, boundary conditions, or both.

One of the fundamental problems in PDEs is the *pure initial value problem* (or *Cauchy problem*) on \mathbb{R} having the form

$$u_t + F(x, t, u, u_x, u_{xx}) = 0, \quad x \in \mathbb{R}, \quad t > 0, \quad (1.4.1)$$

$$u(x, 0) = u_0(x), \quad x \in \mathbb{R}, \quad (1.4.2)$$

where $u_0(x)$ is a given function. Interpreted physically, the function $u_0(x)$ represents a *signal* at time $t = 0$, and the PDE is the equation that propagates the signal in time. Figure 1.9 depicts this interpretation. In wave propagation problems, the signal is usually called a *wave* or *wave profile*. There are several fundamental questions associated with the pure initial value problem (1.4.1)–(1.4.2).

1. *Existence of Solutions.* Given an initial signal $u_0(x)$ satisfying specified regularity conditions (e.g., continuous, bounded, integrable, or whatever), does a solution $u = u(x, t)$ exist for all $x \in \mathbb{R}$ and $t > 0$? If a solution exists for all $t > 0$, it is called a *global solution*. Sometimes solutions are only *local*; that is, they exist for only up to finite times. For nonlinear hyperbolic problems, for example, a signal can propagate up to a finite time and blowup occurs; that is, the signal experiences a gradient catastrophe where u_x becomes infinite and the solution ceases to be smooth. In other problems, for example in some reaction-diffusion equations, the solution u itself may blow up.
2. *Uniqueness.* If a solution of (1.4.1)–(1.4.2) exists, is the solution unique? For a properly posed physical problem we expect an affirmative answer, and therefore we expect the governing initial value problem, which is regarded as a mathematical model for the physical system, to mirror the properties of the system when considering uniqueness and existence questions.
3. *Continuous Dependence on Data.* Another requirement of a physical problem is that of stability; that is, if the initial condition is changed by only a small amount, the system should behave in nearly the same way. Mathematically, this is translated into the statement that the solution should depend continuously on the initial data. In PDEs, if the initial value problem has a unique solution that depends continuously on the initial conditions, we say that the problem is *well-posed*. A similar statement can be made for boundary value problems. A basic question in PDEs is the problem of well-posedness.
4. *Regularity of Solutions.* If a solution exists, how regular is it? In other words, is it continuous, continuously differentiable, or piecewise smooth?

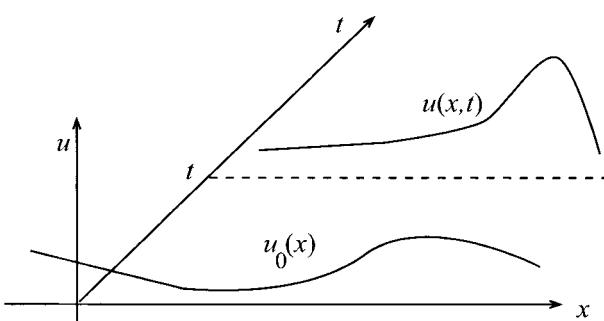


Figure 1.9 Schematic indicating the time evolution, or propagation, of an initial signal or waveform $u_0(x)$.

5. *Asymptotic Behavior.* If an initial signal can be propagated for all times $t > 0$, we may inquire about its asymptotic behavior, or the form of the signal for long times. If the signal decays, for example, what is the decay rate? Does the signal disperse, or does it remain coherent for long times? Does it keep the same shape?

These are a few of the issues in the study of PDEs. The primary issue, however, from the point of view of the applied scientist, may be methods of solution. If a physical problem leads to an initial value problem as a mathematical model, what methods are available or can be developed to obtain a solution, either exact or approximate? Or, if no solution can be obtained (say, other than numerical), what properties can be inferred from the governing PDEs themselves? For example, what is the speed of propagation? Are solutions wave-like, diffusion-like, or dispersive? These questions are addressed in subsequent chapters.

Several examples illustrate the diversity of solutions.

Example. (Diffusion Equation) The solution to initial value problem for the diffusion equation

$$\begin{aligned} u_t &= Du_{xx}, \quad x \in \mathbb{R}, \quad t > 0, \\ u(x, 0) &= u_0(x), \quad x \in \mathbb{R}, \end{aligned}$$

is, as one may verify (e.g., using Fourier transforms), as follows:

$$u(x, t) = \frac{1}{\sqrt{4\pi Dt}} \int_{-\infty}^{\infty} u_0(\xi) e^{(x-\xi)^2/4Dt} d\xi.$$

The solution is valid for all $t > 0$ and $x \in \mathbb{R}$ under rather mild restrictions on the initial signal $u_0(x)$, and the solution has a high degree of smoothness even if the initial data u_0 are discontinuous. Succinctly stated, diffusion smooths out signals. \square

Example. (Advection Equation) Consider the linear initial value problem for the advection equation

$$\begin{aligned} u_t + cu_x &= 0, \quad x \in \mathbb{R}, \quad t > 0, \\ u(x, 0) &= u_0(x), \quad x \in \mathbb{R}, \end{aligned}$$

where c is a positive constant. It is easy to check that $u(x, t) = f(x - ct)$ is a solution of the PDE for any differentiable function f . We can apply the initial condition to determine f by writing $u(x, 0) = f(x) = u_0(x)$. Therefore the global solution to the initial value problem is

$$u(x, t) = u_0(x - ct) \quad x \in \mathbb{R}, \quad t > 0.$$

Graphically, the solution is the initial signal $u_0(x)$ shifted to the right by the amount ct , as shown in Figures 1.10 and 1.11. Therefore, the initial signal moves forward undistorted in spacetime at speed c . Regarding regularity, even if u_0 is discontinuous, it appears that the solution holds, provided we can make sense of derivatives of discontinuous functions. \square

Example. (*Inviscid Burgers' Equation*) A more complicated example is the nonlinear Cauchy problem

$$\begin{aligned} u_t + uu_x &= 0, \quad x \in \mathbb{R}, \quad t > 0, \\ u(x, 0) &= \frac{1}{1+x^2}, \quad x \in \mathbb{R}. \end{aligned}$$

In contrast to the two preceding examples, the solution does not exist for all $t > 0$. The initial waveform, in the form of a bell-shaped curve, distorts during propagation and a gradient catastrophe occurs in finite time. The argument we present to show this nonexistence of a global solution is typical of the types of general arguments that are developed later to study nonlinear hyperbolic problems. Assume that this problem has a solution $u = u(x, t)$, and consider the family of curves in xt space defined by the differential equation

$$\frac{dx}{dt} = u(x, t).$$

Denote a curve C in this family by $x = x(t)$. Along this curve we have $du/dt = u_t(x(t), t) + u_x(x(t), t)dx/dt = 0$, and therefore $u = \text{constant}$ on C . The curve C must be a straight line because $d^2x/dt^2 = du/dt = 0$. The curve C , which is called a *characteristic curve*, is shown in Figure 1.12 emanating from a point

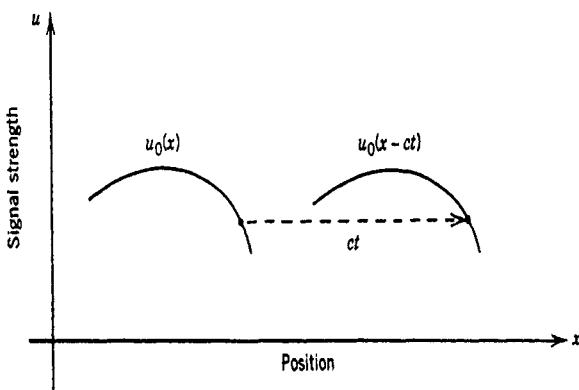


Figure 1.10 Right traveling wave, which represents the solution to the advection equation, shown in xu space.

$(\xi, 0)$ on the initial timeline (x axis) to an arbitrary point (x, t) in spacetime. The equation of C is given by

$$x - \xi = u(\xi, 0)t,$$

where its speed (the reciprocal of its slope), $u(\xi, 0) = (1 + \xi^2)^{-1}$, is determined by the initial condition and the fact that u is constant on C . Now let us determine how the gradient u_x of u evolves along curve C . For simplicity denote $g(t) = u_x(x(t), t)$. Then

$$g'(t) = u_{xx} \frac{dx}{dt} + u_{xt} = (u_t + uu_x)_x - u_x^2 = -u_x^2 = -g(t)^2.$$

The general solution of $g' = -g^2$ is

$$g(t) = \frac{1}{t + c}, \quad c \text{ const.}$$

But $g(0) = 1/c = -2\xi/(1 + \xi^2)^2$ is the initial gradient, and therefore g is given by

$$g(t) = \frac{1}{t - (1 + \xi^2)^2/2\xi}.$$

Note that ξ may be chosen positive. Therefore, along the straight line C the gradient u_x becomes infinite at the finite time $t = (1 + \xi^2)^2/2\xi$. Therefore a smooth solution cannot exist for all $t > 0$. \square

The nonexistence of a global solution to the initial value problem is a typically nonlinear phenomenon. Because physical processes are often governed by nonlinear equations, we may well ask what happens after the gradient catastrophe. Actually, the distortion of the wave profile and development of an infinite gradient is the witnessing of the formation of a shock wave (i.e., a discontinuous solution that propagates thereafter). However, the idea of a nonsmooth

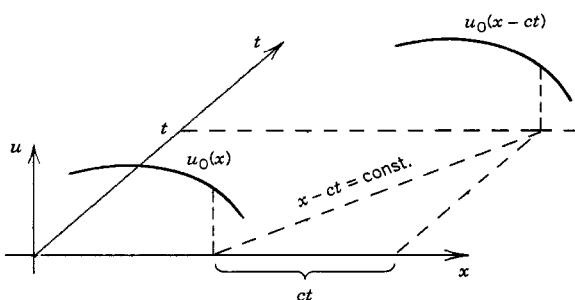


Figure 1.11 Right traveling wave shown in Figure 1.10 represented in xtu -space.

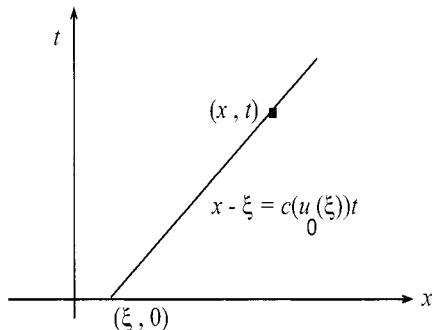


Figure 1.12 Characteristic C emanating from the x axis.

solution to a PDE is a concept that must be formulated carefully, and we carry out this program in Chapter 3.

Another type of problem associated with PDEs is the *signaling problem*. In this case the domain of the problem is the first quadrant $x > 0, t > 0$ in spacetime, and initial data are given along the positive x axis; data are also prescribed along the positive t axis as boundary conditions, or signaling data (see Figure 1.13). The form of a *signaling problem* is

$$u_t + F(x, t, u, u_x, u_{xx}) = 0, \quad x > 0, \quad t > 0, \quad (1.4.3)$$

$$u(x, 0) = u_0(x), \quad x > 0, \quad (1.4.4)$$

$$u(0, t) = u_1(t), \quad t > 0, \quad (1.4.5)$$

where $u_0(x)$ is the given initial state and $u_1(t)$ is a specified signal imposed

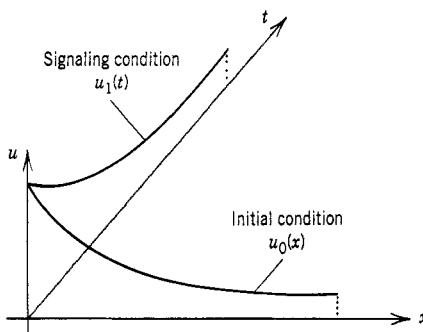


Figure 1.13 Schematic representing a signaling problem where signaling data are prescribed at $x = 0$ along the time axis and initial data are prescribed at $t = 0$ along the spatial axis.

at $x = 0$ for all times $t > 0$. As in the case of the initial value problem, the signaling problem may or may not have a solution that exists for all times t .

In lieu of the condition (1.4.5) given at $x = 0$, one may impose a condition on the derivative u of the form

$$u_x(0, t) = u_2(t), \quad t > 0, \quad (1.4.6)$$

or some combination of u and u_x ,

$$u(0, t) + \alpha u_x(0, t) = u_2(t), \quad t > 0.$$

If the PDE is a conservation law and Fick's law $\phi(x, t) = -Du_x(x, t)$ holds, condition (1.4.6) translates into a condition on the flux ϕ . The condition that the flux be zero at $x = 0$ is the physical condition that the amount of u that passes through $x = 0$ is zero; in heat flow problems this condition is called the *insulated boundary condition*. A boundary condition of the type (1.4.5) is called a *Dirichlet condition*, and one of the type (1.4.6) is called a *Neumann condition*. Mixed conditions are called *Robin conditions*.

If the spatial domain is finite, that is, $a \leq x \leq b$, one may expect to impose boundary data along both $x = a$ and $x = b$, and therefore we consider the *initial-boundary value problem*

$$u_t + F(x, t, u, u_x, u_{xx}) = 0, \quad a < x < b, \quad t > 0, \quad (1.4.7)$$

$$u(x, 0) = u_0(x), \quad a < x < b, \quad (1.4.8)$$

$$u(a, t) = u_1(t), \quad u(b, 0) = u_2(t), \quad t > 0, \quad (1.4.9)$$

where u_0 , u_1 , and u_2 are given functions. If (1.4.7) is the diffusion equation, this problem has a solution under mild restrictions on the data. However, if we consider the advection equation, the problem seldom has a solution for arbitrary boundary data, as the following example shows.

Example. Consider the initial-boundary value problem

$$u_t + cu_x = 0, \quad 0 < x < 1, \quad t > 0,$$

with initial and boundary conditions given by (1.4.8) and (1.4.9) with $a = 0$ and $b = 1$. We have noted already that the general solution of the advection equation is $u(x, t) = f(x - ct)$, for an arbitrary function f . Consequently, u must be constant on the straight lines $x - ct = \text{constant}$ (see Figure 1.14). Clearly, therefore, data cannot be independently specified along the boundary $x = b$. In this case, the initial data along $0 < x < 1$ are carried along the straight lines to the segment A on $x = 1$; the boundary data along $x = 0$ is carried to the segment B on $x = 1$. Thus $u(1, t)$ cannot be specified arbitrarily.

□

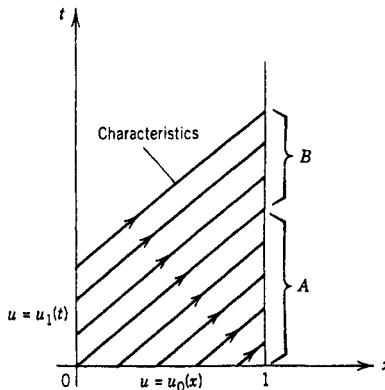


Figure 1.14 Right-moving characteristics carrying left boundary and initial data into the region of interest. Arbitrary data may not be prescribed along the right boundary $x = 1$ labeled A and B .

The preceding example shows that we must be careful in defining a well-posed problem. Mathematically correct initial conditions and boundary conditions are associated with types of PDEs (hyperbolic, parabolic, elliptic), as well as their order. Conditions that ensure well-posedness are often suggested by the underlying physical problem.

Example. (Wave Equation) The wave equation

$$u_{tt} - c^2 u_{xx} = 0, \quad x \in \mathbb{R}, \quad t > 0, \quad (1.4.10)$$

is second-order in t , and therefore it does not fit into the category of equations defined by (1.4.1). Because it is second order in t , we are guided by our experiences with ordinary differential equations to impose two conditions at $t = 0$, a condition on u and a condition on u_t . Therefore, the pure initial value problem for the wave equation consists of (1.4.10) subject to the initial conditions

$$u(x, 0) = u_0(x), \quad u_t(x, 0) = u_1(x), \quad x \in \mathbb{R}, \quad (1.4.11)$$

where u_0 and u_1 are given functions. If $u_0 \in C^2(\mathbb{R})$ and $u_1 \in C^1(\mathbb{R})$, the unique, global solution to (1.4.10)–(1.4.11) is given by D'Alembert's formula (1.1.6) with $f = u_0$ and $g = u_1$. \square

Example. The initial value problem

$$u_{tt} + u_{xx} = 0, \quad x \in \mathbb{R}, \quad t > 0, \quad (1.4.12)$$

$$u(x, 0) = u_0(x), \quad u_t(x, 0) = u_1(x), \quad x \in \mathbb{R} \quad (1.4.13)$$

is not well-posed because small changes, or perturbations, in the initial data can lead to arbitrarily large changes in the solution (see Exercise 5). Equation (1.4.12) is Laplace's equation (with variables x and t , rather than the usual x and y), which is elliptic. In general, initial conditions are not correct for elliptic equations, which are naturally associated with equilibrium phenomena and boundary data. \square

To summarize, the auxiliary conditions imposed on a PDE must be considered carefully. In the sequel, as the subject is developed, the reader should become aware of which conditions go with which equations in order to ensure, in the end, a well-formulated problem.

EXERCISES

1. Solve the signaling problem

$$\begin{aligned} u_t + cu_x &= 0, \quad x > 0, \quad t > 0, \\ u(x, 0) &= 1, \quad x > 0, \\ u(0, t) &= \frac{1+t^2}{1+2t^2}, \quad t > 0, \end{aligned}$$

using the fact that u must be constant on the curves $x = ct + \xi$, where ξ is constant. *Hint:* Treat the regions $x > ct$ and $x < ct$ separately.

2. Obtain the solution to the initial value problem

$$\begin{aligned} u_t &= u_{xx}, \quad x \in \mathbb{R}, \quad t > 0, \\ u(x, 0) &= u_0 \quad \text{if } |x| < L \quad \text{and} \quad u(x, 0) = 0 \quad \text{if } |x| > L, \end{aligned}$$

where u_0 is a constant, in the form

$$u(x, t) = -\frac{u_0}{2} \left[\operatorname{erf}\left(\frac{x-L}{\sqrt{4t}}\right) - \operatorname{erf}\left(\frac{x+L}{\sqrt{4t}}\right) \right],$$

where erf is the *error function*

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-s^2} ds.$$

Show that for x fixed and for large t , we obtain

$$u(x, t) \sim \frac{u_0 L}{\sqrt{\pi t}}.$$

3. Find a formula for the solution to the initial-boundary value problem

$$\begin{aligned} u_t - u_x &= 0, \quad 0 < x < 1, \quad t > 0, \\ u(x, 0) &= 2, \quad 0 < x < 1, \\ u(1, t) &= \frac{2}{1+t^2}, \quad t > 0. \end{aligned}$$

4. Let $u \in C^1$ be a solution to the Cauchy problem

$$\begin{aligned} u_t + Q(u)_x &= 0, \quad x \in \mathbb{R}, \quad t > 0, \\ u(x, 0) &= u_0(x), \quad x \in \mathbb{R}, \end{aligned}$$

where $Q(0) = 0$, $Q''(u) > 0$, and u_0 is integrable on \mathbb{R} with $u_0(x) = 0$ for $x < x_0$ (for some x_0), and $u_0(x) > 0$ otherwise. Define

$$U(x, t) = \int_{-\infty}^x u(s, t) ds$$

- (a) Show that U satisfies the equation $U_t + Q(U_x) = 0$.
 - (b) Prove that $Q(u) \geq Q(v) + c(v)(u - v)$, where $c(u) = Q'(u)$.
 - (c) Prove that $U_t + c(v)U_x \leq c(v)v - Q(v)$.
 - (d) Show that $U(x, t) \leq U(\xi, 0) + t[c(v)v - Q(v)]$, where ξ is the point where the line defined by $dx/dt = c(v)$ through the point (x, t) intersects the x axis.
 - (e) In the inequalities above, show that equality holds if $v = u$.
 - (f) Give a geometric-physical interpretation of the results above.
5. By considering solutions $u_n(x, t) = n^{-1} \cos nx \cosh nt$, show that the initial value problem (1.4.12)-(1.4.13) for Laplace's equation is not well-posed.
6. Consider the initial value problem for the *backward* diffusion equation:

$$\begin{aligned} u_t + u_{xx} &= 0, \quad x \in \mathbb{R}, \quad t > 0, \\ u(x, 0) &= 1, \quad x \in \mathbb{R}. \end{aligned}$$

Show that the solution does not depend continuously on the initial condition by considering the functions

$$u_n(x, t) = 1 + \frac{1}{n} \exp(n^2 t) \sin nx.$$

1.5 Waves

One of the cornerstones of PDEs is wave propagation. A *wave* is a recognizable signal that is transferred from one part of the medium to another part with a recognizable speed of propagation. Energy is often transferred as a wave propagates, but matter may not be. There is hardly any area of science or engineering where wave phenomena are not a critical part of the subject. We mention a few areas where wave propagation is of fundamental importance.

Fluid mechanics (water waves, aerodynamics, meteorology, traffic flow)

Acoustics (sound waves in air and liquids)

Elasticity (stress waves, earthquakes)

Physics (optics, electromagnetic waves, quantum mechanics)

Biology (spread of diseases, population dispersal, nerve signal transmission)

Porous media (groundwater dynamics, contaminant migration)

Chemistry (combustion and detonation waves)

1.5.1 Traveling Waves

The simplest form of a mathematical wave is a function of the form

$$u(x, t) = f(x - ct). \quad (1.5.1)$$

We interpret the density u as the strength of the signal. At $t = 0$ the wave has the form $f(x)$, which is the initial wave profile. Then $f(x - ct)$ represents the profile at time t , which is just the initial profile translated to the right ct spatial units. The constant c represents the speed of the wave. Evidently, (1.5.1) represents a right traveling wave of speed c . Similarly, $u(x, t) = f(x + ct)$ represents a left traveling wave of speed c . These types of waves propagate undistorted along the straight lines $x - ct = \text{const.}$ (or $x + ct = \text{const.}$) in spacetime.

A key question is whether a given PDE can propagate such a traveling wave or, in different words, whether a traveling wave solution (TWS) exists for a given PDE. This question is generally posed without regard to initial conditions (time is usually regarded as varying from $-\infty$ to $+\infty$), so that the wave is assumed to have existed for all times. However, boundary conditions of the form

$$u(-\infty, t) = \text{const.}, \quad u(+\infty, t) = \text{const.} \quad (1.5.2)$$

may be imposed. A *wavefront solution* is a TWS of the form $u(x, t) = f(x - ct)$ (or $f(x + ct)$) subject to the conditions (1.5.2) of constancy at plus and minus infinity (not necessarily the same constant); at present, the function f

is assumed to have the requisite degree of smoothness defined by the PDE ($C^1(\mathbb{R})$, $C^2(\mathbb{R})$, ...). Figure 1.15 shows a typical wavefront. If u approaches the same constant at both plus and minus infinity, the wavefront solution is called a *pulse*.

The key computational device is to substitute the form $u(x, t) = f(x - ct)$ into the PDE and observe whether it transforms the PDE into an ODE for the unknown wave profile f , which is a function of a single variable $z = x - ct$, interpreted as a moving coordinate. The speed c of the wave is not known. To carry out the substitution it is necessary to calculate how the derivatives of u transform. By the chain rule, we have

$$\begin{aligned} u_t &= f'(z)z_t = -cf'(z), \\ u_x &= f'(z)z_x = f'(z). \end{aligned}$$

We easily find $u_{tt} = c^2 f''(z)$, $u_{xx} = f''(z)$, and so on for higher derivatives. Therefore, the general equation

$$u_t = G(u, u_x, u_{xx})$$

transforms into

$$-cf' = G(f, f', f''), \quad z \in \mathbb{R},$$

which is an ODE for f and the unknown wave speed c . If boundary conditions

$$f(-\infty) = f_0, \quad f(+\infty) = f_1$$

are imposed, we interpret this problem as a nonlinear eigenvalue problem for f , with c as an eigenvalue.

Example. (Advection Equation) We have already observed that

$$u(x, t) = f(x - ct),$$

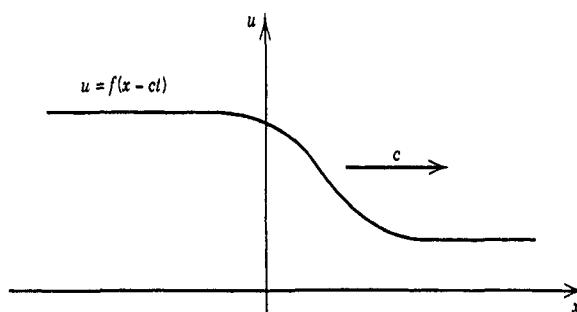


Figure 1.15 A right traveling wavefront with speed c . A wavefront solution is a TWS with constant states at $\pm\infty$.

where f is a differentiable function, is a solution to the PDE

$$u_t + cu_x = 0. \quad (1.5.3)$$

Therefore, the advection equation (1.5.3) admits wavefront solutions, and it is the simplest example of a wave equation. \square

Example. (*Wave Equation*) The wave equation

$$u_{tt} = c^2 u_{xx}$$

has solutions that are the superposition of right and left traveling waves (see Section 1.1). \square

Example. (*Diffusion Equation*) The diffusion equation cannot propagate non-constant wavefronts. To verify this fact, we substitute $u = f(z)$, where $z = x - ct$, into the diffusion equation $u_t = Du_{xx}$ to obtain the following ordinary differential equation for the wave profile f :

$$-cf'(z) = Df''(z), \quad z \in \mathbb{R}$$

This linear differential equation has the general solution

$$f(z) = a + be^{-cz/D},$$

where a and b are arbitrary constants. The only possibility for f to be constant at both plus and minus infinity is to require $b = 0$. Thus, there are traveling wave solutions, but no nonconstant wavefront solutions to the diffusion equation. \square

Example. (*Korteweg–de Vries Equation*) The KdV equation is

$$u_t + uu_x + u_{xxx} = 0. \quad (1.5.4)$$

This nonlinear equation admits traveling wave solutions of different types. One particular type is the soliton, or solitary wave, which is now derived. Assume that $u = f(z)$, where $z = x - ct$. Substituting into (1.5.4) gives

$$-cf' + ff' + f''' = 0, \quad (1.5.5)$$

where the prime denotes d/dz . Integrating (1.5.5) 2 times yields, after rearrangement

$$\frac{df}{dz} = \frac{1}{\sqrt{3}}(-f^3 + 3cf^2 + 6af + 6b)^{1/2}, \quad (1.5.6)$$

where a and b are constants of integration. Here we took the plus sign on the square root; later we observe that this can be done without loss of generality.

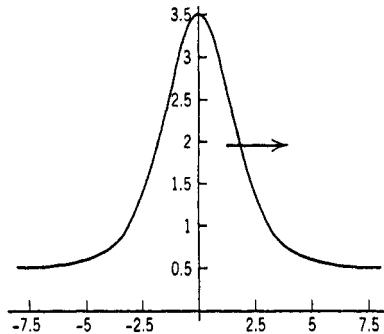


Figure 1.16 Plot of the soliton (1.5.8) at $t = 0$ with $\gamma = 0.5$ and $\alpha - \gamma = 3.0$. A soliton is an example of a pulse.

The first-order differential equation (1.5.6) for the form of the wave f is separable. However, the expression under the radical on the right side is a cubic in f , and therefore different cases must be considered depending on the number of real roots, double roots, and so on. We examine only one case, namely, when the cubic has three real roots, where one is a double root. So let us assume that

$$-f^3 + 3cf^2 + 6af + 6b = (f - \gamma)^2(\alpha - f), \quad 0 < \gamma < \alpha.$$

Then the differential equation (1.5.6) becomes

$$(f - \gamma)^{-1}(\alpha - f)^{-1/2} df = \frac{1}{\sqrt{3}} dz. \quad (1.5.7)$$

The substitution $f = \gamma + (\alpha - \gamma)\operatorname{sech}^2 w$ easily reduces (1.5.7) to

$$-2(\alpha - \gamma)^{-1/2} dw = \frac{1}{\sqrt{3}} dz,$$

which integrates to $w = \sqrt{(\alpha - \gamma)/12}z$. Consequently, traveling waves in the case we are considering have the form

$$u(x, t) = \gamma + (\alpha - \gamma)\operatorname{sech}^2[\kappa(x - ct)], \quad \kappa = \left(\frac{\alpha - \gamma}{12}\right)^{1/2}. \quad (1.5.8)$$

A graph of the waveform, which is a pulse, is shown in Figure 1.16. It is instructive to write the roots α and γ in terms of the original parameters. To this end we have

$$\begin{aligned} -f^3 + 3cf^2 + 6af + 6b &= (f - \gamma)^2(\alpha - f) \\ &= -f^3 + (\alpha + 2\gamma)f^2 - (2\gamma\alpha + \gamma^2)f + \gamma^2\alpha. \end{aligned}$$

Therefore, the wave speed c is given by

$$c = \frac{\alpha + 2\gamma}{3} = \frac{\alpha - \gamma}{3} + \gamma. \quad (1.5.9)$$

The speed of the wave relative to the state γ ahead of the wave is proportional to the amplitude $A = \alpha - \gamma$. Both the amplitude and the width of the wave depend on the wave speed c . Thus, the taller and wider the wave, the faster it moves. For linear waves, say, governed by the wave equation, the speed of propagation is independent of the amplitude of the wave. Such a waveform (1.5.8) is known as a *soliton*, or *solitary wave*, and many of the important equations of mathematical physics exhibit soliton-type solutions (e.g., the Boussinesq equation, the Sine–Gordon equation, the Born–Infeld equation, and nonlinear Schrödinger equations). In applications, the value of such solutions is that if a pulse or signal travels as a soliton, the information contained in the pulse can be carried over long distances with no distortion or loss of intensity. Solitons occur in fluid mechanics, nonlinear optics, and other nonlinear phenomena. \square

There is an interesting history of solitary waves. In 1836 John Scott Russell, a Scottish engineer, observed such a wave moving along a canal. He followed it on horseback, noting that it propagated a long distance without changing form. Many at the time doubted his observation and thought that such waves could not exist. However, Boussinesq, in 1872, showed that in a special limit of the equations of fluid flow such water waves can exist; Korteweg and deVries obtained (1.5.4) in the 1890. It wasn't until the 1960s that physicists Kruskal and Zabusky obtained the equation as a continuum limit of a model of nonlinear spring mass chains, and they coined the term *soliton*.

The interaction between two solitons is both interesting and unexpected. If two solitons are moving to the right and the one behind represents a stronger signal (therefore moving faster) than the one ahead, the large one overtakes the slower, smaller wave and a complicated nonlinear interaction occurs, after which both waves return to their original shape. The only change is a phase shift in the two waves. The initial value problem for the KdV equation, with $u(x, 0) = u_0(x)$, $x \in \mathbb{R}$, also has interesting behavior. If the initial profile u_0 approaches zero fast enough at $x = \pm\infty$, then, over time, the solution divides into a finite number of solitons moving to the right at their respective speeds (smaller to larger), plus a small dispersive disturbance moving to $-\infty$. \square

The problem of determining whether a given PDE admits wavefront solutions is fundamental, and it occupies much attention in subsequent chapters. The search for these types of solutions is one of the basic methods in the analysis of nonlinear problems.

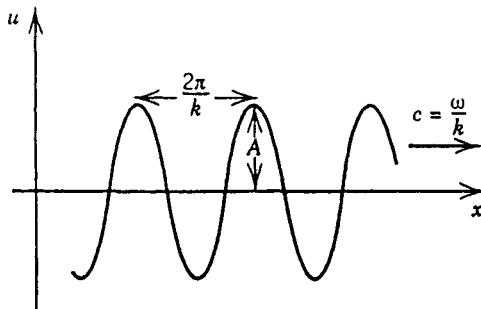


Figure 1.17 A plane wave of amplitude A , wave number k , and frequency ω .

1.5.2 Plane Waves

Another type of wave of interest is a *plane wave*, or *wave train*. These waves are traveling, periodic waves of the form

$$u(x, t) = A \cos(kx - \omega t), \quad (1.5.10)$$

where A is the amplitude of the wave, k the wave number, and ω the frequency. The *wave number* k is a measure of the number of spatial oscillations (per 2π units) observed at a fixed time, and the *frequency* ω is a measure of the number of oscillations in time (per 2π units) observed at a fixed spatial location. The number $\lambda = 2\pi/k$ is the *wavelength*, and $P = 2\pi/\omega$ is the *period*. The wavelength measures the distances between successive crests, and the period measures the time for an observer located at a fixed position x to see a repeat pattern (see Figure 1.17). Note that (1.5.10) may be written

$$u(x, t) = A \cos k \left(x - \frac{\omega}{k} t \right),$$

so (1.5.10) represents a traveling wave moving to the right with speed $c = \omega k$. This number is called the *phase velocity*, and it is the speed that one would have to move to remain on the crest of the wave. For calculations the complex form

$$u(x, t) = A e^{i(kx - \omega t)} \quad (1.5.11)$$

is preferred because differentiations with the exponential function are simpler. After the calculations are completed, one can use Euler's formula $\exp(i\theta) = \cos \theta + i \sin \theta$ and take real or imaginary parts to recover a real solution. The technique of searching for plane wave solutions is applicable to linear equations, but a modification applies to some nonlinear problems. Finding plane wave solutions is a powerful method for determining properties of linear problems.

Example. (*Diffusion Equation*) We seek solutions of the form of equation (1.5.11) to the diffusion equation

$$u_t = Du_{xx}. \quad (1.5.12)$$

Substituting (1.5.11) into (1.5.12) gives

$$\omega = -iDk^2. \quad (1.5.13)$$

Equation (1.5.13) is a condition, called a *dispersion relation*, between the frequency ω and the wave number k that must be satisfied for (1.5.12) to admit plane wave solutions. Thus we have determined a class of solutions

$$u(x, t; k) = Ae^{-Dk^2t}e^{ikx}, \quad k \in \mathbb{R} \quad (1.5.14)$$

for the diffusion equation that depends on an arbitrary parameter k , the wave number. The factor $\exp(ikx)$ represents a spatial oscillation, and the factor $A \exp(-Dk^2t)$ represents a decaying amplitude. Note that the rate of decay depends on the wave number k ; waves of shorter wavelength decay more rapidly than do waves of longer wavelength. \square

Example. (*Wave Equation*) Substituting (1.5.11) into the wave equation

$$u_{tt} = c^2 u_{xx} \quad (1.5.15)$$

forces

$$\omega = \pm ck. \quad (1.5.16)$$

Therefore, the wave equation admits solutions of the form

$$u(x, t) = Ae^{ik(x \pm ct)},$$

which are right and left sinusoidal traveling waves of speed c . \square

The technique of looking for solutions of the form (1.5.11) for linear, homogeneous PDEs of the form

$$Lu = 0, \quad (1.5.17)$$

where L is a linear constant-coefficient operator, always leads to a relation connecting the frequency ω and the wave number k of the form

$$G(k, \omega) = 0. \quad (1.5.18)$$

This condition is called the *dispersion relation* corresponding to the PDE (1.5.17), and it characterizes plane wave solutions entirely. Generally, linear PDEs (1.5.17) can be classified according to their dispersion relation in the following way. Assume that (1.5.18) may be solved for ω in the form

$$\omega = \omega(k). \quad (1.5.19)$$

We say that the PDE is *dispersive* if $\omega(k)$ is real and if $\omega''(k) \neq 0$. When $\omega(k)$ is complex, the PDE is *diffusive*. The diffusion equation, from the example above, has dispersion relation $\omega = -ik^2D$, which is complex; thus the diffusion equation is classified as diffusive. The classical wave equation (1.5.15) is neither diffusive nor dispersive under the foregoing classification. Even though its solutions are wave-like, it is not classified as dispersive since its dispersion relation (see 1.5.16) satisfies $\omega'' = 0$. For the wave equation, the speed of propagation is c and is independent of the wave number k . For dispersive equations, the speed of propagation, or the phase velocity ω/k , depends on the wave number (or equivalently, the wavelength). The wave equation is generally regarded as the prototype of a hyperbolic equation, and the term *dispersive* is reserved for equations where the phase velocity depends on k . We caution the reader that the term *dispersion* is often used in a diffusion context (e.g., a group of animals dispersing), and it is important to be aware of this.

Example. (Schrödinger Equation) In quantum mechanics, the Schrödinger equation for a free particle, under appropriate scalings, is

$$u_t = iu_{xx}.$$

It is easy to see that the dispersion relation is $\omega = k^2$, so that the Schrödinger equation is dispersive. The Schrödinger equation is neither parabolic nor hyperbolic. \square

1.5.3 Plane Waves and Transforms

Using a plane wave assumption, we already constructed the class of solutions $u(x, t; k) = Ae^{-Dk^2t}e^{ikx}$, $k \in \mathbb{R}$, to the diffusion equation

$$u_t = Du_{xx}, \quad x \in \mathbb{R}, \quad t > 0.$$

Formally, superimposing these solutions, we obtain

$$u(x, t) = \int_{\mathbb{R}} A(k)e^{-k^2Dt}e^{ikx} dk, \quad (1.5.20)$$

which can be verified to be a solution to the diffusion equation, provided that $A(k)$ is a well-behaved function (e.g., continuous, bounded, and integrable on \mathbb{R}). Having (1.5.20) as a solution opens up the possibility of selecting the function $A(k)$ so that other conditions (e.g., an initial condition) can be met. Therefore, let us impose the initial condition

$$u(x, 0) = u_0(x), \quad x \in \mathbb{R},$$

to the diffusion equation. From (1.5.20) it follows that

$$u_0(x) = \int_{\mathbb{R}} A(k) e^{ikx} dk. \quad (1.5.21)$$

We recognize $u_0(x)$ as the Fourier transform of $A(k)$, and therefore $A(k)$ must be the inverse Fourier transform of the function $u_0(x)$. We remind the reader of these facts by recalling the *Fourier integral theorem* [see, e.g., Stakgold 1998].

Theorem. Let f be a continuous, bounded, integrable function on \mathbb{R} , and let

$$\hat{f}(k) = \int_{\mathbb{R}} f(x) e^{ikx} dx$$

be the Fourier transform of f . Then, for all $x \in \mathbb{R}$, we have

$$f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{f}(k) e^{-ikx} dk.$$

Applying this theorem (the variables k and x have been interchanged) to 1.5.21) leads us to conclude that

$$A(k) = \frac{1}{2\pi} \int_{\mathbb{R}} u_0(x) e^{-ikx} dx.$$

Consequently, we have obtained the solution to the Cauchy problem for the diffusion equation

$$u_t - Du_{xx} = 0, \quad x \in \mathbb{R}, \quad t > 0, \quad (1.5.22)$$

$$u(x, 0) = u_0(x), \quad x \in \mathbb{R}, \quad (1.5.23)$$

in the form

$$u(x, t) = \frac{1}{2\pi} \int_{\mathbb{R}} \left(\int_{\mathbb{R}} u_0(\xi) e^{-ik\xi} d\xi \right) e^{-k^2 Dt} e^{ikx} dk. \quad (1.5.24)$$

Interchanging the order of integration allows us to formally write the solution as

$$u(x, t) = \frac{1}{2\pi} \int_{\mathbb{R}} u_0(\xi) \left(\int_{\mathbb{R}} e^{ik(x-\xi)-k^2 Dt} dk \right) d\xi. \quad (1.5.25)$$

The inner integral may be calculated outright by noting

$$\begin{aligned} \int_{\mathbb{R}} e^{ik(x-\xi)-k^2 Dt} dk &= 2 \int_0^\infty e^{-k^2 Dt} \cos(k(x-\xi)) dk \\ &= \sqrt{\frac{\pi}{Dt}} e^{-(x-\xi)^2/4Dt}. \end{aligned}$$

Therefore, the solution to the initial value problem (1.5.22)–(1.5.23) for the diffusion equation is given by

$$u(x, t) = \frac{1}{\sqrt{4\pi Dt}} \int_{\mathbb{R}} u_0(\xi) e^{-(x-\xi)^2/4Dt} d\xi. \quad (1.5.26)$$

We derived this solution formally, but we have not proved that it is indeed a solution. A proof would consist of a rigorous argument that (1.5.25) satisfies the diffusion equation (1.5.22) and the initial condition (1.5.23). To show well posedness, a uniqueness argument would have to be supplied as well as a proof that the solution is stable to small perturbations of the initial data. We shall not carry out these arguments here, but rather, refer the reader to the references for the details.

This method (finding plane wave solutions followed by superposition and use of the Fourier integral theorem) is equivalent to the classical Fourier transform method learned in elementary courses where one takes a Fourier transform of the PDE and initial condition to reduce the problem to an ordinary differential equation in the transform domain, which is then solved. Then the inverse Fourier transform is applied to return the solution in the original domain. This method is generally applicable to the pure initial value problems on the real line for linear equations with constant coefficients.

Fourier integral expressions of the type obtained above can be approximated for large times by the method of stationary phase [e.g., see Bhatnagar 1979].

1.5.4 Nonlinear Dispersion

For nonlinear equations we do not expect plane wave solutions of the form (1.5.16), and therefore a dispersion relation will not exist as it does for linear equations. Moreover, superposition for nonlinear equations is invalid. However, in some nonlinear problems, there may exist traveling periodic wave trains of the form

$$u(x, t) = U(\theta), \quad \theta = kx - \omega t, \quad (1.5.27)$$

where U is a periodic function. For example, consider the nonlinear PDE

$$u_{tt} - u_{xx} + f'(u) = 0, \quad (1.5.28)$$

where $f(u)$ is some function of u , yet to be specified. If (1.5.27) is substituted into (1.5.28), we obtain the ordinary differential equation

$$(\omega^2 - k^2)U_{\theta\theta} + f'(U) = 0.$$

Here we are using subscripts θ to denote ordinary derivatives of U with respect

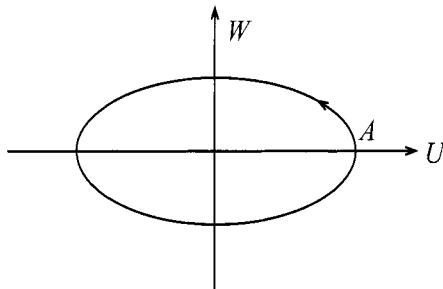


Figure 1.18 Phase space orbit representing the periodic solution (1.5.31) with amplitude A .

θ. Multiplication by U_θ and subsequent integration gives

$$\frac{1}{2}(\omega^2 - k^2)U_\theta^2 + f(U) = A, \quad (1.5.29)$$

where A is a constant of integration. The goal is to determine U as a periodic function of θ . Equation (1.5.29) has the same form as an energy conservation law, where f is a potential function, which suggests introducing the variable W defined by $W = U_\theta$. Then

$$W^2 = \frac{2}{\omega^2 - k^2}[A - f(U)], \quad (1.5.30)$$

which are the integral curves. Assuming that $\omega^2 > k^2$, we have

$$W = \pm \sqrt{\frac{2}{\omega^2 - k^2}} \sqrt{A - f(U)}, \quad (1.5.31)$$

which defines a locus of points in the UW plane. For example, let us choose $f(U) = U^4$. Then the locus (1.5.31) is a closed path, representing a periodic solution of (1.5.29), as shown in Figure 1.18. Notice that A , which has been taken positive, is the amplitude of the oscillation represented by the closed path. To find U as a function of θ in the case $f(U) = U^4$ we write (1.5.29) as

$$\theta = \pm \sqrt{\frac{\omega^2 - k^2}{2}} \int_0^U \frac{1}{\sqrt{A - s^4}} ds. \quad (1.5.32)$$

This formula defines the periodic function $U = U(\theta)$ implicitly; in this case $U(\theta)$ can be determined explicitly as an elliptic function, and we leave this as an exercise. The period of the oscillation can be determined by integrating over one-quarter period in the integral on the right side of (1.5.32), taking care to choose the appropriate sign. If P denotes the period, we obtain an equation of the form

$$P = P(\omega, k, A). \quad (1.5.33)$$

In other words, the frequency ω will depend on the amplitude A as well as the wave number k . Consequently, the wave speed $c = \omega/k$ will be amplitude-dependent. This amplitude dependence in the nonlinear dispersion relation (1.5.33) is one important distinguishing aspect of nonlinear phenomena.

These calculations can be carried out for nonlinear equations of the form (1.5.28) for various potential functions $f(u)$. Periodic solutions are obtained when U oscillates between two simple zeros of $A - f(U)$. A thorough discussion of nonlinear dispersion can be found in Whitham (1974).

EXERCISES

- Find the dispersion relation for the advection–diffusion equation

$$u_t + au_x = Du_{xx}, \quad (a, D > 0),$$

and show that it is diffusive. Use superposition and the Fourier integral theorem to find an integral representation of the solution of the initial value problem for this equation. *Hint:* You will need the Fourier transform of $e^{iax} f(x)$.

- Consider the KdV equation in the form $u_t - 6uu_x + u_{xxx} = 0$, $x \in \mathbb{R}$, $t > 0$. Let $u = u(x, t)$ be a solution that decays, along with its derivatives, very rapidly to zero as $|x| \rightarrow \infty$. Show that $\int_{\mathbb{R}} u dx$ and $\int_{\mathbb{R}} u^2 dx$ are both constant in time.
- Examine the form of traveling wave solutions of the KdV equation (1.5.4) in the case that the cubic expression on the right side of (1.5.6) has a triple real root.
- Determine and sketch traveling wave solutions of the equation

$$u_{tt} - u_{xx} = -\sin u$$

in the form

$$u(x, t) = 4 \arctan \left\{ \exp \pm \left[\frac{x - ct}{\sqrt{1 - c^2}} \right] \right\}, \quad 0 < c < 1.$$

- For the following PDEs find the dispersion relation and classify the equations as diffusive, dispersive, or neither:

- (a) $u_{tt} + a^2 u_{xxxx} = 0$ (beam equation)
- (b) $u_t + u_{xxx} = 0$ (dispersive wave equation)
- (c) $u_t + au_x + bu_{xxx} = 0$ (linearized KdV equation)
- (d) $u_{tt} - c^2 u_{xx} + bu = 0$ (Klein–Gordon equation)

6. Consider the heat equation in an infinite domain $x > 0$ where the boundary condition at $x = 0$ is a periodic function over all time:

$$\begin{aligned} u_t &= ku_{xx}, \quad x > 0, t \in \mathbb{R}, \\ u(0, t) &= T_0 + Ae^{i\omega t}, \quad t \in \mathbb{R}, \end{aligned}$$

where T_0 , A , and ω of constants, and where $u = u(x, t)$ is the temperature. This problem models temperatures in the ground subject to surface periodic temperatures. Find $u(x, t)$ and determine the amplitude and phase shift, relative to the values at the surface, at a depth x . Answer the same questions if instead the flux is imposed at the boundary:

$$-Ku_x(0, t) = Ae^{i\omega t}, \quad t \in \mathbb{R}.$$

7. By superimposing plane wave solutions to the dispersive wave equation $u_t + u_{xxx} = 0$, find an integral representation of the solution to the Cauchy problem and write your answer in terms of the Airy function

$$\text{Ai}(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \cos\left(\frac{z^3}{3} + xz\right) dz.$$

8. Find solutions of the outgoing signaling problem for the wave equation:

$$\begin{aligned} u_{tt} - c^2 u_{xx} &= 0, \quad x > 0, t \in \mathbb{R}. \\ u_x(0, t) &= s(t), \quad t \in \mathbb{R}. \end{aligned}$$

9. In equation (1.5.28) take $f(u) = u^2/2$ and determine periodic solutions of the form $u = U(\theta)$, where $\theta = kx - \omega t$. What is the period of oscillation? Does ω depend on the amplitude of oscillation in this case?

10. In equation (1.5.28) assume that the potential function $f(u)$ has the expansion

$$f(u) = \frac{u^2}{2} + \sigma u^4 + \dots,$$

where σ is a small known parameter. Assuming that the amplitude is small, show that periodic wave trains are given by

$$U(\theta) = a \cos \theta + \frac{1}{8}\sigma a^3 \cos 3\theta + \dots,$$

where the frequency and amplitude are

$$\omega^2 = 1 + k^2 + 3\sigma a^2 + \dots,$$

$$A = \frac{a^2}{2} + \frac{9}{8}\sigma a^4 + \dots.$$

11. The *nonlinear Schrödinger equation* occurs in the description of water waves, nonlinear optics, and plasma physics. It is given by

$$iu_t + u_{xx} + \gamma|u^2|u = 0, \quad (1.5.34)$$

where $\gamma > 0$ and $u = u(x, t)$ is complex-valued.

- (a) If $u = U(z)e^{i(kz-\omega t)}$, where $z = x - ct$, show that

$$\frac{d^2U}{dz^2} + i(2k - c)\frac{dU}{dz} + (kx - \omega + k^2)U + \gamma U^3 = 0.$$

- (b) If $c = 2k$, show that

$$\left(\frac{dU}{dz}\right)^2 = aU^2 - \frac{\gamma}{2}U^4 + C$$

for some appropriately chosen constant a , where C is an arbitrary constant.

- (c) Taking $C = 0$ and $a > 0$, show that

$$U(z) = \sqrt{\frac{2a}{\gamma}} \operatorname{sech}(\sqrt{a}(z - ct)),$$

and comment on the properties of this solution.

12. Let $F = F(u)$ be a smooth function and suppose

$$u_{tt} - u_{xx} = F'(u).$$

Assuming that u and its partial derivative u_x both go to zero as $|x| \rightarrow \infty$, show that

$$\int_{\mathbb{R}} \left(\frac{1}{2}u_t^2 + \frac{1}{2}u_x^2 + F(u) \right) dx = \text{const.}$$

13. Show that solutions of the nonlinear Schrödinger equation (1.5.34) with $\gamma = 1$ have the properties

$$\int_{\mathbb{R}} |u|^2 dx = \text{const}, \quad \int_{\mathbb{R}} (u_x^2 - \frac{1}{2}|u|^4) dx = \text{const},$$

provided u and its derivatives approach zero sufficiently fast as $|x| \rightarrow 0$.

Reference Notes. Partial differential equations have a long history and a correspondingly vast literature. The early developments in PDEs were in the post calculus years of the early 1700s and involved the geometry of surfaces. However, it soon became clear that PDEs were models of physical phenomena like fluid flow, vibrating strings, heat conduction, and so on. Euler, Bernoulli

D'Alembert, Laplace, Lagrange, Fourier, Cauchy, and others developed many of the basic ideas in the linear theory and its applications. It wasn't until the last half of the twentieth century, as part of the general interest in nonlinear science and advancement of computation, that nonlinear PDEs came to the forefront, particularly in their role in wave propagation and diffusion.

It is impractical to cite more than just a few key references. There are books at all levels and several research journals, in both mathematics and the pure and applied sciences, that can be consulted for an entry point to the literature. Here we reference only a few of the texts and cite only articles that are relevant to the particular topic under discussion.

Elementary, entry-level texts focus almost exclusively on linear problems, emphasizing Fourier series, integral transforms, and boundary value problems. The long-time standard has been Churchill (1969), which remains an excellent introduction to Fourier series and boundary value problems. There are many more recent introductory texts, too many to cite. We mention Strauss (1992), an outstanding treatment, and the author's text (Logan 2004), which is a brief introduction. More advanced texts include the classic by John (1982), as well as Evans (1998), McOwen (2003), Guenther & Lee (1996), Renardy & Rogers (2004), and Stakgold (1998). Strichartz (1994) gives an outstanding perspective on Fourier transforms and distributions.

More specialized, theoretical books include Friedman (1964) for parabolic equations, Gilbarg & Trudinger (1983) for elliptic equations, Protter & Weinberger (1967) for maximum principles, and Pao (1992) for both nonlinear parabolic and elliptic equations. The treatises by Courant & Hilbert (1953, 1962) provide a wealth of information on elliptic and hyperbolic equations, and Courant & Friedrichs (1948) is still a standard in shock waves and gas dynamics. Kevorkian & Cole (1981) and Zauderer (2006) discuss perturbation methods. For nonlinear equations one can consult Lax (1973), Smoller (1994) and Whitham (1974), all of which are key books. Another text on nonlinear PDEs that is very similar to the first edition of the present text (Logan 1994) is Debnath (1997). Bhatnagar (1979) is an excellent introduction to nonlinear dispersive waves, and solitons are discussed in Drazin & Johnson (1989).

Two excellent volumes that introduce several common and ad hoc methods for nonlinear PDEs are Ames (1965, 1972)

First-Order Equations and Characteristics

A single first-order PDE is a hyperbolic equation that is wave-like, that is, associated with the propagation of signals at finite speeds. The fundamental idea associated with hyperbolic equations is the notion of a characteristic, a curve in spacetime (a hypersurface in higher dimensions) along which signals propagate. There are several ways of looking at the concept of a characteristic; one definition, and this forms the basis of our definition of a characteristic, is that it is a curve along which the PDE can be reduced to a simpler form, for example, an ordinary differential equation. But ultimately, the characteristics are the curves along which information is carried.

The aim in this chapter is to build a solid base of understanding of characteristics by examining a first-order PDE in one spatial variable and time. We focus on the initial value problem and the signaling problem, and carry out the analysis under the assumption that a continuous smooth solution exists. In Chapter 3 the concept of a weak solution is discussed, and the smoothness assumptions are relaxed. In Sections 2.1 and 2.2 we study linear and nonlinear equations, respectively, and in Section 2.3 we examine the general quasilinear equation. Then in Section 2.4 we address the subject of wavefront expansions and show that discontinuities in the derivatives also propagate along the characteristics. In Section 2.5 we discuss the general nonlinear equation of first order.

2.1 Linear First-Order Equations

2.1.1 Advection Equation

We already showed that the initial value problem for the advection equation, namely

$$u_t + cu_x = 0, \quad x \in \mathbb{R}, \quad t > 0, \quad (2.1.1)$$

$$u(x, 0) = u_0(x), \quad x \in \mathbb{R}, \quad (2.1.2)$$

has a unique solution given by $u(x, t) = u_0(x - ct)$, which is a right traveling wave of speed c . We now present an alternative derivation of this solution that is in the spirit of the analysis of hyperbolic PDEs. In this entire section we assume that the initial signal u_0 is a smooth (continuously differentiable) function.

First we recall a basic fact from elementary calculus. If $u = u(x, t)$ is a function of two variables and $x = x(t)$ defines a smooth curve C in the xt plane, the total derivative of u along the curve C is given, according to the chain rule, by

$$\frac{d}{dt}u(x(t), t) = u_t(x(t), t) + u_x(x(t), t)\frac{dx}{dt}.$$

This expression defines how u changes along C . By observation, the left side of equation (2.1.1) is a total derivative of u along the curves defined by the equation $dx/dt = c$. We may therefore recast (2.1.1) into the statement

$$\frac{du}{dt} = 0 \text{ along the curves defined by } \frac{dx}{dt} = c,$$

or, equivalently

$$u = \text{constant} \quad \text{on} \quad x - ct = \xi,$$

where ξ is a constant. In these expressions we have suppressed the arguments of u for concise notation; all expressions involving u are assumed to be evaluated along the curve, namely, at $(x(t), t)$. Consequently, the PDE (2.1.1) reduced to an ordinary differential equation, which was subsequently integrated, along the family of curves $x - ct = \xi$, which are solutions to $dx/dt = c$. If we draw one of these curves in the xt plane that passes through an arbitrary point (x, t) , it intersects the x axis at $(\xi, 0)$; its speed is c and its slope is $1/c$ because we are graphing t versus x rather than x versus t . It is common to refer to the speed of a curve in spacetime, rather than its slope (see Figure 2.1). Now, because u is constant on this curve, we have

$$u(x, t) = u(\xi, 0) = u_0(\xi) = u_0(x - ct),$$

which is the solution to the initial value problem (2.1.1)–(2.1.2). The totality of all the curves $x - ct = \xi$, where ξ is constant, is called the set of *characteristic*

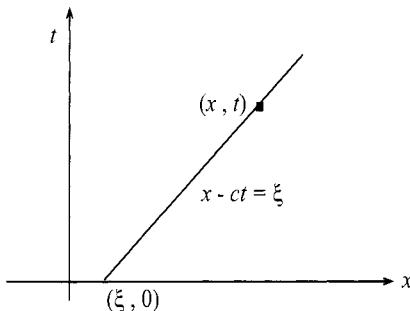


Figure 2.1 Characteristic curve with speed c passing through the arbitrary point (x, t) .

curves for this problem. The constant ξ acts as a parameter and distinguishes the curves. A graph of the set of characteristic curves on a spacetime diagram (i.e., in the xt plane) is called the *characteristic diagram* for the problem. Figure 2.2 shows the characteristic diagram for the problem (2.1.1)–(2.1.2). The characteristic curves, or just characteristics, are curves in spacetime along which signals propagate. In the present case, the signal is the *constancy of u* that is carried along the characteristics. Further—and this will turn out to be a defining feature—the PDE reduces to an ordinary differential equation along the characteristics.

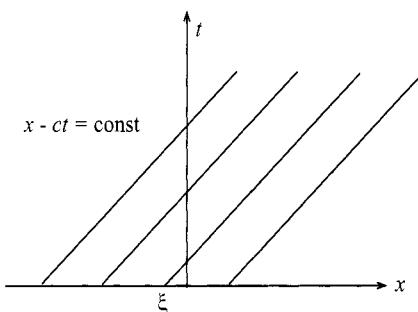


Figure 2.2 Characteristic diagram for the PDE (2.1.1) showing the family of characteristic curves $x - ct = \xi$. The curves have speed c and are parameterized by their intersection ξ with the x axis.

2.1.2 Variable Coefficients

We can extend the preceding notions to more complicated problems. First consider the linear initial value problem

$$u_t + c(x, t)u_x = 0, \quad x \in \mathbb{R}, \quad t > 0, \quad (2.1.3)$$

$$u(x, 0) = u_0(x), \quad x \in \mathbb{R}, \quad (2.1.4)$$

where $c = c(x, t)$ is a given continuous function. The left side of the PDE (2.1.3) is a total derivative along the curves in the xt plane defined by the differential equation

$$\frac{dx}{dt} = c(x, t). \quad (2.1.5)$$

Along these curves

$$\frac{du}{dt} = u_t + u_x \frac{dx}{dt} = u_t + c(x, t)u_x = 0,$$

or, in other words, u is constant. Therefore

$$u = \text{const} \quad \text{on} \quad \frac{dx}{dt} = c(x, t).$$

Again the PDE (2.1.3) reduces to an ordinary differential equation (which was integrated to a constant) along a special family of curves, the characteristics defined by (2.1.5). The function $c = c(x, t)$ gives the speed of these characteristic curves, which varies in spacetime.

Example. Consider the initial value problem

$$u_t - xt u_x = 0, \quad x \in \mathbb{R}, \quad t > 0, \quad (2.1.6)$$

$$u(x, 0) = u_0(x), \quad x \in \mathbb{R}. \quad (2.1.7)$$

The characteristics are defined by the equation

$$\frac{dx}{dt} = -xt,$$

which, on quadrature, gives

$$x = \xi e^{-t^2/2}, \quad (2.1.8)$$

where ξ is a constant. On these curves the PDE becomes

$$u_t - xt u_x = u_t - \frac{dx}{dt} u_x = \frac{du}{dt} = 0,$$

or

$$u = \text{const} \quad \text{on} \quad x = \xi e^{-t^2/2}.$$

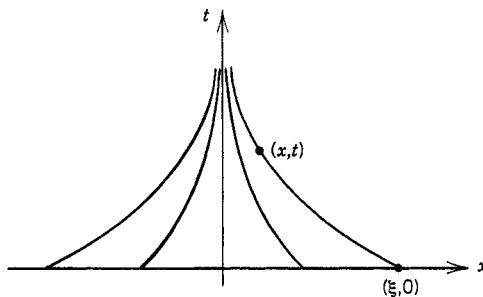


Figure 2.3 Characteristic diagram for (2.1.6) showing the family of characteristics given by (2.1.8).

The characteristic diagram is shown in Figure 2.3, with an arbitrary point (x, t) labeled on a given characteristic that emanates from a point $(\xi, 0)$ on the x axis. From the constancy of u along the characteristics, we have

$$u(x, t) = u(\xi, 0) = u_0(\xi) = u_0 \left(x e^{t^2/2} \right),$$

which is a solution to (2.1.6)–(2.1.7), valid for all $t > 0$. Figure 2.4 shows how an initial signal $u_0(x)$ is focused along the characteristics to a region near $x = 0$ as time increases. \square

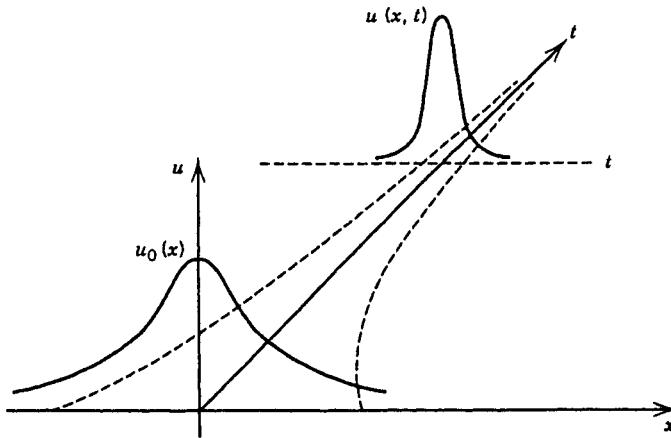


Figure 2.4 Diagram showing how an initial signal is propagated in spacetime by the PDE (2.1.6) along the characteristics (2.1.8).

This method, called the *method of characteristics*, can be extended to non-

homogeneous initial value problems of the form

$$u_t + c(x, t)u_x = f(x, t), \quad x \in \mathbb{R}, \quad t > 0, \quad (2.1.9)$$

$$u(x, 0) = u_0(x), \quad x \in \mathbb{R}, \quad (2.1.10)$$

where c and f are given continuous functions. Now the PDE (2.1.9) reduces to the ordinary differential equation

$$\frac{du}{dt} = f(x, t) \quad (2.1.11)$$

on the characteristic curves defined by

$$\frac{dx}{dt} = c(x, t). \quad (2.1.12)$$

The pair of differential equations (2.1.11)–(2.1.12) can, in theory, be solved subject to the initial conditions

$$x = \xi, \quad u = u_0(\xi) \quad \text{on } t = 0,$$

to find the solution. The constant ξ is again a parameter that distinguishes each characteristic by defining its intersection with the x axis. In practice it may be impossible to solve the characteristic system (2.1.11)–(2.1.12) in closed form, and one may have to adopt numerical quadrature methods. Existence and uniqueness results from ordinary differential equations imply that a unique solution of (2.1.11)–(2.1.12) exists in some neighborhood of each point $(x_0, 0)$; however, a global solution for all $t > 0$ may not exist.

Boundary value problems are handled in the same way. Boundary data along $x = 0$ (the t axis) of the form

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

can be parameterized by

$$t = \tau, \quad u = g(\tau) \quad \text{on } x = 0,$$

which give conditions on the characteristic system.

Example. (Boundary Value Problem) Consider the initial-boundary value problem

$$u_t + u_x = 0, \quad x > 0, \quad t > 0,$$

$$u(x, 0) = 0, \quad x > 0,$$

$$u(0, t) = te^{-t}, \quad t > 0.$$

The characteristic system is

$$\frac{du}{dt} = 0, \quad \frac{dx}{dt} = 1,$$

which has general solution

$$u = c_1, \quad x = t + c_2.$$

The characteristics are straight lines of slope 1 emanating from both the x and t axes. Because conditions on the axes are different, we separate the problem into two regions, $x > t$ and $x < t$ (ahead and behind the line $x = t$, which is the separating characteristic). It is clear that $u = 0$ in the entire region $x > t$ because u is constant on characteristics, and it is zero along the x axis. Along the t axis the data can be parameterized by

$$t = \tau, \quad u = \tau e^{-\tau}, \quad x = 0.$$

It follows that the arbitrary constants c_1 and c_2 are given by

$$c_1 = \tau e^{-\tau}, \quad c_2 = -\tau.$$

Therefore the solution is given in parametric form by

$$u = \tau e^{-\tau}, \quad x = t - \tau.$$

Eliminating the parameter τ gives

$$u(x, t) = (t - x)e^{x-t}, \quad x > t. \quad \square$$

EXERCISES

1. Solve the initial value problem

$$u_t + 2u_x = 0, \quad x \in \mathbb{R}, \quad t > 0,$$

$$u(x, 0) = \frac{1}{1+x^2}, \quad x \in \mathbb{R}.$$

Sketch the characteristics. Sketch wave profiles at $t = 0$, $t = 1$, and $t = 4$.

2. Solve the initial value problem

$$u_t + 2tu_x = 0, \quad x \in \mathbb{R}, \quad t > 0,$$

$$u(x, 0) = e^{-x^2}, \quad x \in \mathbb{R}.$$

3. Solve the initial value problem

$$u_t - x^2u_x = 0, \quad x \in \mathbb{R}, \quad t > 0,$$

$$u(x, 0) = x + 1, \quad x \in \mathbb{R}.$$

4. Solve the signaling problem

$$\begin{aligned} u_t - x^2 u_x &= 0, \quad x > 0, \quad t \in \mathbb{R}, \\ u(0, t) &= g(t), \quad t \in \mathbb{R}. \end{aligned}$$

5. Solve the initial-boundary value problem

$$\begin{aligned} u_t - x^2 u_x &= 0, \quad x > 0, \quad t > 0, \\ u(x, 0) &= e^{-x}, \quad x > 0, \\ u(0, t) &= 1, \quad t > 0. \end{aligned}$$

6. Insofar as possible, write down a formula for the solution to the initial value problem

$$\begin{aligned} u_t + c u_x &= f(t), \quad x \in \mathbb{R}, \quad t > 0, \\ u(x, 0) &= u_0(x), \quad x \in \mathbb{R}. \end{aligned}$$

2.2 Nonlinear Equations

When nonlinear terms are introduced into PDEs, the situation changes dramatically from the linear case discussed in Section 2.1. The method of characteristics, however, still provides the vehicle for obtaining information about how initial signals propagate. To begin, consider the simple nonlinear initial value problem

$$u_t + c(u)u_x = 0, \quad x \in \mathbb{R}, \quad t > 0, \tag{2.2.1}$$

$$u(x, 0) = u_0(x), \quad x \in \mathbb{R}, \tag{2.2.2}$$

where $c = c(u)$ is a given smooth function of u ; here and in this entire section, the initial signal u_0 is assumed to be smooth. We recognize (2.2.1) as the basic conservation law

$$u_t + \phi(u)_x = 0, \quad c(u) = \phi'(u),$$

where $\phi = \phi(u)$ is the flux. The nonlinearity in (2.2.1) occurs in the advection term $c(u)u_x$. Equation 2.2.1, often called a *kinematic wave equation*, arises in nonlinear wave phenomena when dissipative effects such as viscosity and diffusion are ignored. It is a special case of the more general quasilinear equation that is discussed in the next section.

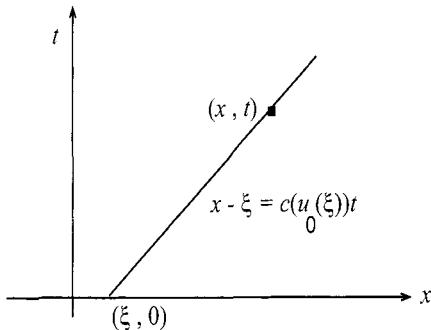


Figure 2.5 Characteristic given by $x - \xi = c(u_0(\xi))t$ having speed $c(u_0(\xi))$.

To analyze (2.2.1) let us assume for a moment that a C^1 solution $u = u(x, t)$ exists for $t > 0$. Motivated by the approach for linear equations in Section 2.1, we define characteristic curves by the differential equation

$$\frac{dx}{dt} = c(u), \quad (2.2.3)$$

where $u = u(x, t)$. Of course, contrary to the situation for linear equations, the right side of this equation is not known a priori because the solution is not yet determined. Thus the characteristics cannot be determined in advance. In any case, along the curves defined by (2.2.3) the PDE (2.2.1) becomes

$$u_t + c(u)u_x = u_t + u_x \frac{dx}{dt} = \frac{du}{dt} = 0,$$

or $u = \text{const.}$ Thus u is constant on the characteristic curves. It is easy to observe that the characteristic curves defined by (2.2.3) are straight lines, for

$$\frac{d^2x}{dt^2} = \frac{dc(u)}{dt} = c'(u)\frac{du}{dt} = 0.$$

Therefore, let us draw a characteristic back in time from an arbitrary point (x, t) in spacetime to a point $(\xi, 0)$ on the x axis (see Figure 2.5). The equation of this characteristic is given by

$$x - \xi = c(u_0(\xi))t, \quad (2.2.4)$$

where we have noted that its speed (reciprocal slope) is given by dx/dt or $c(u(0, \xi))$, because u is constant on the entire characteristic line. Further, it follows that

$$u(x, t) = u(\xi, 0) = u_0(\xi).$$

Therefore, if a solution to the initial value problem exists for $t > 0$, then necessarily it must be given by

$$u(x, t) = u_0(\xi), \quad (2.2.5)$$

where $\xi = \xi(x, t)$ is given implicitly by (2.2.4).

The solution of (2.2.1)–(2.2.2) determined implicitly by (2.2.4) and (2.2.5) was obtained under the assumption that a C^1 solution exists for all $t > 0$. Let us now verify that we indeed have a solution by substituting (2.2.5) into the PDE (2.2.1). We have

$$u_t = u'_0(\xi)\xi_t, \quad u_x = u'_0(\xi)\xi_x.$$

The partial derivatives ξ_t and ξ_x can be calculated by implicit differentiation of (2.2.4) to obtain

$$-\xi_t = c'(u_0(\xi))u'_0(\xi)\xi_t + c(u_0(\xi)), \quad 1 - \xi_x = c'(u_0(\xi))u'_0(\xi)\xi_x,$$

and therefore

$$u_t = -\frac{c(u_0(\xi))u'_0(\xi)}{D}, \quad u_x = \frac{u'_0(\xi)}{D}, \quad (2.2.6)$$

where $D = 1 + c'(u_0(\xi))u'_0(\xi)t$. Consequently, $u_t + c(u)u_x = 0$.

The preceding equations were obtained under the assumption that a unique solution $u = u(x, t)$ exists for all $t > 0$. Implicit, therefore, is the assumption that each point (x, t) in the xt plane uniquely determines a value of ξ , and the partial derivatives (2.2.6) exist and are finite for all x and t . It seems clear that the derivatives in (2.2.6), however, will not be finite unless some restrictions are placed on the initial signal u_0 and the function $c(u)$ in order to guarantee that the denominator D will not vanish. One such set of conditions is that the derivatives u' and c' both have the same sign; then the denominator D in (2.2.6) is always positive, and it can be verified that (2.2.4) and (2.2.5) provide a solution to (2.2.1)–(2.2.2). We record this result as a theorem.

Theorem. If the functions c and u_0 are $C^1(\mathbb{R})$ and if u_0 and c are either both nondecreasing or both nonincreasing on \mathbb{R} , the nonlinear initial value problem (2.2.1)–(2.2.2) has a unique solution defined implicitly by the parametric equations

$$\begin{aligned} u(x, t) &= u_0(\xi), \\ x - \xi &= c(u_0(\xi))t. \end{aligned}$$

Example. Consider the problem

$$u_t + uu_x = 0, \quad x \in \mathbb{R}, \quad t > 0, \quad (2.2.7)$$

$$u(x, 0) = 0 \quad \text{if } x \leq 0; \quad u(x, 0) = e^{-1/x} \quad \text{if } x > 0. \quad (2.2.8)$$

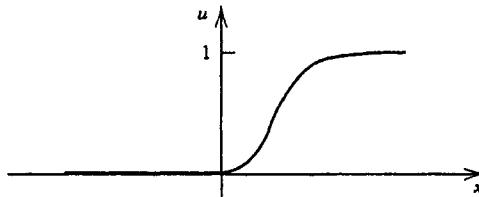


Figure 2.6 Graph of the initial waveform (2.2.8).

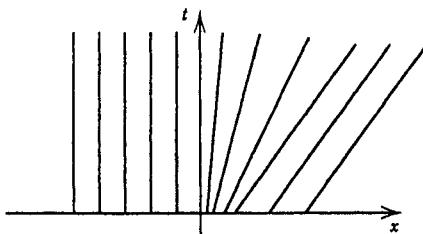


Figure 2.7 Characteristic diagram associated with (2.2.7)–(2.2.8).

Here $c(u) = u$ and $u_0(x)$ are smooth nondecreasing functions on \mathbb{R} . The initial signal $u_0(x)$ is shown in Figure 2.6. The characteristics emanating from a point $(\xi, 0)$ on the x axis have speed $c(u_0(\xi)) = u_0(\xi)$. Hence $c = 0$ if $\xi \leq 0$, and the characteristics are vertical lines; $c = \exp(-1/\xi)$ if $\xi > 0$ and the characteristics fan out and approach speed unity as ξ increases. Figure 2.7 shows the characteristic diagram. The wave in the region where the characteristics fan out is called a *rarefaction wave* or *release wave*. Figure 2.8 shows how the initial wave is propagated along the characteristics (recall that u is constant on the characteristics). The solution to the problem (2.2.7)–(2.2.8) is given implicitly by

$$u(x, t) = 0 \quad \text{for } x \leq 0,$$

$$u(x, t) = e^{-1/\xi}, \quad \text{where } x - \xi = te^{-1/\xi} \quad \text{for } x > 0. \quad \square$$

In Chapter 3 we address the important question of what happens when the hypotheses of the last theorem are not satisfied and the derivatives u_x and u_t in (2.2.6) blow up. In this case the wave develops a discontinuity known as a *shock*.

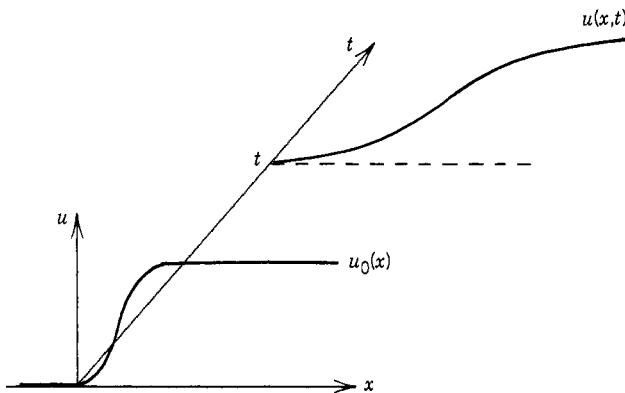


Figure 2.8 Rarefaction, or release, wave.

EXERCISES

1. Consider the initial value problem

$$u_t + uu_x = 0, \quad x \in \mathbb{R}, \quad t > 0, \\ u(x, 0) = -x, \quad x \in \mathbb{R}.$$

Sketch the characteristic diagram and find the solution.

2. Solve the nonlocal advection equation

$$u_t + \left(\int_0^1 F(u(\xi, t)) d\xi \right) u_x = 0, \quad 0 < x < 1, \quad t > 0, \\ u(x, 0) = f(x), \quad 0 < x < 1, \\ u(0, t) = 0, \quad t > 0.$$

Solve it also in the case the boundary condition is $u(0, t) = g(t)$ (see Logan 2003).

3. Solve the signaling problem

$$u_t = -u_x + u^2, \quad x > 0, \quad t \in \mathbb{R}, \\ u(0, t) = g(t), \quad t \in \mathbb{R}.$$

2.3 Quasilinear Equations

Now we examine the nonlinear PDE

$$u_t + c(x, t, u)u_x = f(x, t, u), \quad x \in \mathbb{R}, \quad t > 0, \quad (2.3.1)$$

where the coefficients c and f are continuous functions. Such equations are called *quasi-linear* because of the way the nonlinearity occurs—that is, the equation is linear in the derivatives and the nonlinearity occurs through multiplication by coefficients that depend on u . We consider the initial value problem and append to (2.3.1) the initial condition

$$u(x, 0) = u_0(x), \quad x \in \mathbb{R}, \quad (2.3.2)$$

where $u_0(x)$ is continuously differentiable on \mathbb{R} . Our approach in this section will be the same as in Section 2.2; we examine (2.3.1) and (2.3.2) under the assumption that a smooth solution exists, postponing a discussion of discontinuous solutions to Chapter 3. The goal at present is to build further on the method of characteristics and show how an algorithm can be constructed to solve (2.3.1)–(2.3.2).

Therefore, let $u = u(x, t)$ be a smooth solution to (2.3.1)–(2.3.2). From the prior discussion we observe that the PDE (2.3.1) reduces to the ordinary differential equation

$$\frac{du}{dt} = f(x, t, u) \quad (2.3.3)$$

along the family of curves (characteristics) defined by

$$\frac{dx}{dt} = c(x, t, u). \quad (2.3.4)$$

We may regard (2.3.3) and (2.3.4) as a system of two ordinary differential equations (called the *characteristic system*) for u and x , and we may solve them in principle, subject to the initial conditions

$$u = u_0(\xi), \quad x = \xi \quad \text{at} \quad t = 0 \quad (2.3.5)$$

to obtain the solution. Here, as before, we regard ξ as a number that parameterizes the characteristic curves, giving the intersection of the curves with the x axis.

Example. Consider the initial value problem

$$u_t + uu_x = -u, \quad x \in \mathbb{R}, \quad t > 0, \quad (2.3.6)$$

$$u(x, 0) = -\frac{x}{2}, \quad x \in \mathbb{R}. \quad (2.3.7)$$

The characteristic system is

$$\frac{du}{dt} = -u \quad \text{on} \quad \frac{dx}{dt} = u, \quad (2.3.8)$$

with initial data

$$u = -\frac{\xi}{2}, \quad x = \xi \quad \text{at} \quad t = 0. \quad (2.3.9)$$

It is straightforward to solve (2.3.8) to obtain

$$u = ae^{-t}, \quad x = b - ae^{-t}, \quad (2.3.10)$$

where a and b are constants of integration. Applying the initial conditions (2.3.9) gives a and b in terms of the parameter ξ according to

$$a = -\frac{\xi}{2}, \quad b = \frac{\xi}{2}.$$

Then the solution of the characteristic system (2.3.8) is, in parametric form

$$u = -\frac{\xi}{2}e^{-t}, \quad x = \frac{\xi}{2}(1 + e^{-t}), \quad \xi \in \mathbb{R}.$$

In the present case we may eliminate the parameter ξ from the second equation to obtain

$$\xi = \frac{2x}{1 + e^{-t}}.$$

Thus

$$u(x, t) = -\frac{xe^{-t}}{1 + e^{-t}},$$

which is a smooth solution valid for all $t > 0$ and $x \in \mathbb{R}$. \square

It is unlikely in the general case that the calculations can be performed as in the example and a smooth, closed-form solution derived. The signaling problem, as well as other boundary value problems, may be handled in a similar manner using the method of characteristics.

Example. Consider the initial-boundary value problem

$$u_t + uu_x = 0, \quad x > 0, \quad t > 0, \quad (2.3.11)$$

$$u(x, 0) = 1, \quad x > 0, \quad (2.3.12)$$

$$u(0, t) = \frac{1}{1 + t^2}, \quad t > 0. \quad (2.3.13)$$

The PDE reduces to the total derivative $du/dt = 0$ on the characteristic curves defined by $dx/dt = u$. Thus

$$u = \text{const} \quad \text{on} \quad \frac{dx}{dt} = u.$$

As we determined earlier, the characteristics are straight lines with speed u . Therefore, all the characteristics emanating from the x axis, where $u = 1$, have speed 1. Figure 2.9 shows the characteristic diagram for (2.3.11)–(2.3.13). This means that the constant solution $u = 1$ is carried into the region $x > t$ by these characteristics. For $x < t$ the characteristics fan out because the speed

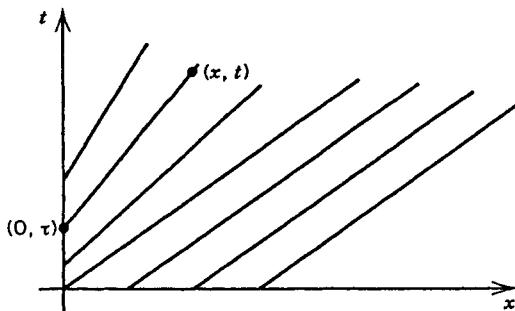


Figure 2.9 Characteristic diagram associated with (2.3.11)–(2.3.13)

u decreases along the t axis according to the boundary condition (2.3.13). Let us now select an arbitrary point (x, t) in the region $x < t$ where we want to determine the solution. Following this characteristic back to the boundary $x = 0$, we denote its intersection point by $(0, \tau)$. The equation of this characteristic is given by

$$x = u(0, \tau)(t - \tau),$$

or

$$x = \frac{1}{1 + \tau^2}(t - \tau). \quad (2.3.14)$$

By the constancy of u along the characteristics, we know that

$$u(x, t) = u(0, \tau) = \frac{1}{1 + \tau^2},$$

and $\tau = \tau(x, t)$ is given implicitly by formula (2.3.14). In the present case we may solve (2.3.14) for τ to obtain

$$\tau = \frac{-1 + \sqrt{1 + 4x(t - x)}}{2x}. \quad (2.3.15)$$

Note that (2.3.14) reduces to a quadratic equation in τ that can be solved by the quadratic formula, taking the plus sign on the square root in order to satisfy the boundary condition. Consequently, a solution to the problem (2.3.11)–(2.3.13) is

$$\begin{aligned} u(x, t) &= 1 \quad \text{for } x \geq t, \\ u(x, t) &= \frac{1}{1 + \tau^2} \quad \text{for } x < t, \end{aligned}$$

where τ is given by (2.3.15). \square

2.3.1 The General Solution

In some contexts the general solution of the quasilinear equation (2.3.1) is needed, in terms of a single, arbitrary function. To determine the general solution we make some definitions and observations. An expression $\psi(x, t, u)$ is called a *first integral* of the characteristic system (2.3.3)–(2.3.4) if $\psi(x, t, u) = k$ (constant), on solutions to (2.3.3)–(2.3.4). In other words, $\psi(X(t), t, U(t)) = k$ for all t in some interval I , if $x = X(t)$, $u = U(t)$, $t \in I$, is a solution to (2.3.3)–(2.3.4). Taking the total derivative of this last equation with respect to t and using the chain rule gives

$$\psi_x c + \psi_t + \psi_u f = 0, \quad t \in I, \quad (2.3.16)$$

where each of the terms in this equation are evaluated at $(X(t), t, U(t))$.

In addition, if $\psi(x, t, u)$ is a first integral of the characteristic system, then the equation $\psi(x, t, u) = k$ defines, implicitly, a surface $u = u(x, t)$, on some domain D in the xt plane, provided $\psi_u \neq 0$. Thus

$$\psi(x, t, u(x, t)) = k, \quad (x, t) \in D.$$

Using the chain rule we calculate the partial derivatives of u as

$$\psi_t + \psi_u u_t = 0, \quad \psi_x + \psi_u u_x = 0,$$

or

$$u_t = -\frac{\psi_t}{\psi_u}, \quad u_x = \frac{\psi_x}{\psi_u}, \quad (x, t) \in D.$$

Each term on the right is evaluated at $(x, t, u(x, t))$. Observe that the curve $(X(t), t, U(t))$ lies on this surface because $\psi(X(t), t, U(t)) = k$.

We claim that this surface $u = u(x, t)$ is a solution to the partial differential equation (2.3.1). To prove this, fix an arbitrary point (ξ, τ) in D . Then

$$u_t(\xi, \tau) + cu_x(\xi, \tau) = -\frac{\psi_t}{\psi_u} - c\frac{\psi_x}{\psi_u}, \quad (2.3.17)$$

where c and all the terms on the right side are evaluated at (ξ, τ, ω) , where $\omega = u(\xi, \tau)$. But a solution curve $(X(t), t, U(t))$ lies on the surface and passes through the point (ξ, τ, ω) at $t = \tau$. That is, $(\xi, \tau, \omega) = (X(\tau), \tau, U(\tau))$. Therefore, the right side of (2.3.17) may be evaluated at $(X(\tau), \tau, U(\tau))$.

In general, we expect to find two independent first integrals to the characteristic system, $\psi(t, x, u)$ and $\chi(t, x, u)$. The *general solution* of (2.3.1) is then

$$G(\psi(t, x, u), \chi(t, x, u)) = 0, \quad (2.3.18)$$

where G is an arbitrary function. We can formally solve for one of the variables and write this as $\psi(t, x, u) = g(\chi(t, x, u))$, where g is an arbitrary function.

We leave it as an exercise (using the chain rule) to show that if $\psi(t, x, u)$ and $\chi(t, x, u)$ are first integrals of the characteristic system, then (2.3.18) defines, implicitly, a solution $u = u(x, t)$ to (2.3.1) for any function G , provided $G_\psi \psi_u + G_\chi \chi_u \neq 0$.

Example. Consider the equation

$$u_t + 2tu_x = \frac{x}{2u+1}.$$

The characteristic system is

$$\frac{dx}{dt} = 2t, \quad \frac{du}{dt} = \frac{x}{2u+1}.$$

The first equation gives first integral $\psi = x - t^2 = c_1$. Then the second equation becomes

$$\frac{du}{dt} = \frac{t^2 + c_1}{2u+1}.$$

Separating variables and integrating gives another first integral $\chi = u^2 + u + \frac{2}{3}t^3 - xt = c_2$. Therefore, the general solution is

$$G(x - t^2, u^2 + u + \frac{2}{3}t^3 - xt) = 0,$$

where G is an arbitrary function. This expression defines implicit solutions. It can also be written

$$u^2 + u + \frac{2}{3}t^3 - xt = g(x - t^2),$$

where g is an arbitrary function. This quadratic in u can be solved to determine explicit solutions. \square

Example. Consider the linear problem

$$\begin{aligned} u_t + \frac{t}{x}u_x &= \frac{u}{x}, \quad x, t > 0, \\ u(x, 0) &= x^3, \quad x > 0, \\ \lim_{x \rightarrow 0} u(x, t) &= t^3, \quad t > 0. \end{aligned}$$

The boundary condition should be interpreted in a limiting sense. The characteristic system is

$$\frac{dx}{dt} = \frac{t}{x}, \quad \frac{du}{dt} = \frac{u}{x}.$$

The first equation integrates immediately to give the first integral

$$x^2 - t^2 = C_1,$$

which defines the characteristics, a family of hyperbolas, in the xt plane. To solve the second characteristic equation we can substitute for x and write

$$\frac{du}{dt} = \frac{u}{\pm\sqrt{t^2 + C_1}}.$$

But this integration is rather involved, and sometimes it is easier to find a second first integral by manipulating the characteristic equations directly. If we divide the two characteristic equations we get

$$\frac{du}{dx} = \frac{u}{t}.$$

Therefore we have

$$\frac{dx}{du} = \frac{t}{u}, \quad \frac{dt}{du} = \frac{x}{u}.$$

Adding these two equations gives

$$\frac{d(t+x)}{du} = \frac{1}{u}(t+x).$$

Now we can separate variables and integrate to get

$$\frac{u}{x+t} = C_2.$$

We can immediately write the general solution as

$$F\left(\frac{u}{x+t}, x^2 - t^2\right) = 0,$$

or

$$u = (x+t)G(x^2 - t^2).$$

Now we determine the arbitrary function G on each side of the separating characteristic $x = t$. For $x > t$ we use the initial data to get

$$u(x, 0) = xG(x^2) = x^3,$$

which implies the form of G is $G(z) = z$. Therefore we have

$$u = (x+t)(x^2 - t^2), \quad x > t.$$

For $x < t$ we use the boundary data to get

$$u(0, t) = tG(-t^2) = t^3,$$

which implies the form of G is $G(z) = -z$. Therefore we have

$$u = (x+t)(t^2 - x^2), \quad x < t.$$

Note that as $x \rightarrow 0$, the boundary condition is satisfied.

It is clear from the solution that u is continuous across the characteristic $x = t$. However, u_x is not continuous across $x = t$ because for $x > t$ we have $u_x = (x+t)(2x) + x^2 - t^2 \rightarrow 4t^2$ as $x \rightarrow t^+$ and for $x < t$ we have $u_x = (x+t)(-2x) + t^2 - x^2 \rightarrow -4t^2$ as $x \rightarrow t^-$. The fact that discontinuities are propagated along characteristics is a feature of nonlinear PDEs. \square

EXERCISES

1. Solve the initial value problem

$$\begin{aligned} u_t + uu_x &= -ku^2, \quad x \in \mathbb{R}, \quad t > 0, \\ u(x, 0) &= 1, \quad x \in \mathbb{R}, \end{aligned}$$

where k is a positive constant. Sketch the characteristics.

2. Consider the initial value problem

$$\begin{aligned} u_t + uu_x &= -ku, \quad x \in \mathbb{R}, \quad t > 0, \\ u(x, 0) &= u_0(x), \quad x \in \mathbb{R}, \end{aligned}$$

where k is a positive constant. Determine a condition on u_0 so that a smooth solution exists for all $t > 0$ and $x \in \mathbb{R}$, and determine the solution in parametric form.

3. Solve the initial value problem

$$\begin{aligned} u_t + cu_x &= xu, \quad x \in \mathbb{R}, \quad t > 0, \\ u(x, 0) &= u_0(x), \quad x \in \mathbb{R}, \end{aligned}$$

where c is a positive constant. Sketch the characteristics.

4. Solve the initial value problem

$$\begin{aligned} u_t - xt u_x &= x, \quad x \in \mathbb{R}, \quad t > 0, \\ u(x, 0) &= u_0(x), \quad x \in \mathbb{R}. \end{aligned}$$

5. Solve the signaling problem

$$\begin{aligned} u_t + u_x &= u^2, \quad x > 0, \quad t \in \mathbb{R}, \\ u(0, t) &= u_0(t), \quad t \in \mathbb{R}. \end{aligned}$$

6. Consider the PDE

$$u_t + \frac{x+t}{x-t} u_x = \frac{x^2 + t^2}{u(x-t)}$$

- (a) Show that $\phi = 2xt - u^2$ and $\psi = u^2 - x^2 - t^2$ are two first integrals of the characteristic system and find the general solution.
- (b) Find the solution that satisfies $u = 0$ on the line $2t = x$.
7. Show that if $u = u(x, y)$ satisfies the equation

$$xu_x + yu_y = nu,$$

where n is a positive integer, then

$$u = x^n f\left(\frac{y}{x}\right),$$

where f is an arbitrary function.

8. Derive the general solution

$$u = xt\varphi\left(\frac{x-t}{xt}\right)$$

of the PDE

$$t^2u_t + x^2u_x = (x+t)u.$$

9. Consider

$$uu_t + u_x = 1$$

with Cauchy data $u = x/2$ on $x = t$ for $x \in (0, 1)$. Show that

$$u = \frac{4x - 2t - x^2}{2(2-x)}.$$

Find the domain of validity and sketch the characteristics.

10. Consider

$$tu_t + (x+u)u_x = u + t^2$$

with Cauchy data $u = x$ on $t = 1$ for $x \in \mathbb{R}$. Show that

$$u = \frac{x - t^2}{1 + \ln t} + t^2.$$

Find the domain where the solution is valid.

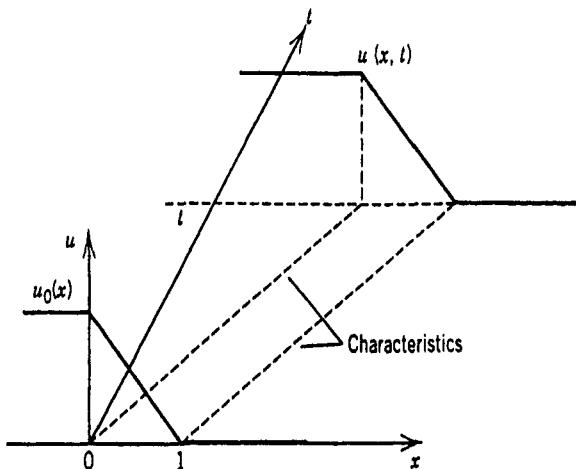


Figure 2.10 Discontinuities in derivatives propagating along characteristics.

2.4 Propagation of Singularities

In the preceding discussion and examples we assumed that the initial and boundary data were given by smooth functions. Now we consider the case where the initial or boundary data are continuous but may have discontinuities in their derivatives. The question we address is how those discontinuities on the boundary of the region are propagated into the region of interest. A simple example shows what to expect.

Example. Consider the simple advection equation

$$u_t + cu_x = 0, \quad x \in \mathbb{R}, \quad t > 0, \quad (2.4.1)$$

subject to an initial condition given by a piecewise smooth function u_0 defined by

$$u(x, 0) = u_0(x) = \begin{cases} 1, & x < 0 \\ 1 - x, & 0 < x < 1 \\ 0, & x > 1. \end{cases}$$

Because the general solution of (2.4.1) is $u(x, t) = f(x - ct)$, a right traveling wave, u is constant on the characteristic curves $x - ct = \text{const}$, which carry the initial data into the region $t > 0$ (see Figure 2.10). Thus the discontinuities in u' at $x = 0$ and $x = 1$ are carried along the characteristics into the region $t > 0$. \square

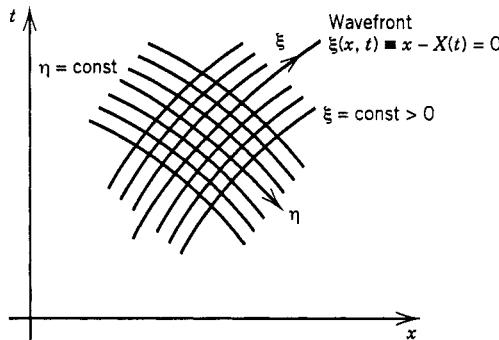


Figure 2.11 A wavefront $x = X(t)$ and associated curvilinear coordinates ξ and η .

In summary, characteristics carry data from the boundary into the region. Therefore, abrupt changes in the derivatives on the boundary produce corresponding abrupt changes in the region. In other words, discontinuities in derivatives propagate along the characteristics. For a general nonlinear equation, as we observe in Chapter 3, discontinuities in the boundary functions themselves do not propagate along characteristics; these kinds of singularities are *shocks* and they propagate along different spacetime curves.

For the simple nonlinear kinematic wave equation

$$u_t + c(u)u_x = 0, \quad (2.4.2)$$

we now demonstrate that discontinuities in the derivatives propagate along characteristics. Let $\xi = \xi(x, t)$ be a curve in xt -space separating two regions where a solution $u = u(x, t)$ is C^1 , and suppose that u has a crease in the surface; that is, that u is continuous across the curve, but there is a simple jump discontinuity in the derivatives of u across the curve. Such a curve is sometimes called a *wavefront* (see Figure 2.11). Observe that this is a different use of the term wavefront from that used in the context of TWS. To analyze the behavior of u along the wavefront, we introduce a new set of curvilinear coordinates given by

$$\xi = \xi(x, t), \quad \eta = \eta(x, t). \quad (2.4.3)$$

For example, $\eta(x, t) = \text{const}$ can be taken as the family of curves orthogonal to $\xi(x, t) = \text{const}$. Then the chain rule for derivatives implies

$$u_t = u_\xi \xi_t + u_\eta \eta_t, \quad u_x = u_\xi \xi_x + u_\eta \eta_x.$$

Here there should be no confusion in using the same variable u for the transformed function of ξ and η . The PDE (2.4.2) therefore becomes

$$(\xi_t + c(u)\xi_x)u_\xi + (\eta_t + c(u)\eta_x)u_\eta = 0. \quad (2.4.4)$$

This equation is valid in the regions $\xi > 0$ and $\xi < 0$. Also, by hypothesis

$$u(0+, \eta) = u(0-, \eta), \quad (2.4.5)$$

and therefore

$$u_\eta(0+, \eta) = u_\eta(0-, \eta). \quad (2.4.6)$$

Consequently, the tangential derivatives are continuous across the wavefront $\xi(x, t) = 0$.

Next we introduce some notation for the jump in a quantity across a wavefront. Let Q be some quantity that has a value Q_+ just ahead (to the right) of the wavefront, and a value Q_- just behind (to the left) of the wavefront. Then the *jump in the quantity Q* across the wavefront is defined by

$$[Q] = Q_- - Q_+.$$

Continuing with our calculation, we take the limit of equation (2.4.4) as $\xi \rightarrow 0+$ and then take the limit of (2.4.4) as $\xi \rightarrow 0-$. Subtracting the two results gives

$$(\xi_t + c(u)\xi_x)[u_\xi] = 0, \quad (2.4.7)$$

where we used (2.4.5) and (2.4.6). If we assume that $[u_\xi] \neq 0$, then

$$\xi_t + c(u)\xi_x = 0. \quad (2.4.8)$$

In particular, if the wavefront $\xi(x, t) = 0$ is given by $x = X(t)$, then (2.4.8) becomes

$$\frac{dX}{dt} = c(u). \quad (2.4.9)$$

That is, $\xi(x, t) = 0$ must be a characteristic. We summarize the result in the following theorem.

Theorem. Let D be a region of spacetime, and let $\xi(x, t) = x - X(t) = 0$ be a smooth curve lying in D that partitions D into two disjoint regions D^+ and D^- (see Figure 2.12). Let u be a smooth solution to (2.4.2) in D^+ and D^- that is continuous in D , and assume that the derivatives of u suffer simple jump discontinuities across $\xi(x, t) = 0$. Then $\xi(x, t) = 0$ must be a characteristic curve.

The next issue concerns the magnitude of the jump $[u_x]$ as the wave propagates via (2.4.2) along a characteristic. Under special assumptions we now

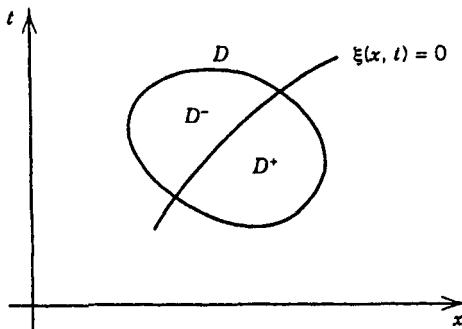


Figure 2.12

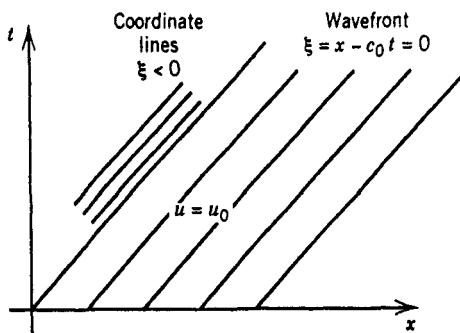


Figure 2.13 Wavefront propagating into a constant state.

derive an ordinary differential equation for the magnitude of the jump that governs its change. Let us assume that the wavefront is at the origin $x = 0$ at time $t = 0$, and that the state ahead of the wave is a constant state $u = u_0$. Then the characteristics ahead of the wavefront have constant speed $c_0 = c(u_0)$, and the wavefront itself has equation $\xi(x, t) = x - c_0 t = 0$ (see Figure 2.13). Thus $u = u_0$ for $\xi > 0$. The coordinate lines $\xi = \text{constant}$ are parallel to the wavefront, and we want to examine the behavior of the wave near the wavefront (where ξ is small and negative) as time increases. Therefore, using the assumption that ξ is small, we make the Ansatz

$$u = \begin{cases} u_0, & \xi > 0 \quad (\text{ahead}) \\ u_0 + u_1(t)\xi + \frac{1}{2}u_2(t)\xi^2 + \dots, & \xi < 0 \quad (\text{behind}) \end{cases}$$

Computing the derivatives yields

$$u_t = u_1(t)(-c_0) + \xi u'_1(t) + u_2(t)\xi(-c_0) + O(\xi^2), \quad (2.4.10)$$

$$u_x = u_1(t) + u_2(t)\xi + O(\xi^2), \quad (2.4.11)$$

and

$$\begin{aligned} c(u) &= c_0 + c'(u_0)(u - u_0) + O((u - u_0)^2) \\ &= c_0 + c'(u_0)u_1(t)\xi + O(\xi^2). \end{aligned} \quad (2.4.12)$$

Note that the approach here is to trade in the independent variables t and x for new variables t and ξ , where ξ is a curvilinear coordinate measuring the distance from the wavefront. Now substitute (2.4.10)–(2.4.12) into the PDE (2.4.2) to obtain

$$\begin{aligned} &-c_0u_1 + \xi(u'_1 - c_0u_2) + O(\xi^2) \\ &+ ((c_0 + c'(u_0)u_1\xi + O(\xi^2))(u_1 + u_2\xi + O(\xi^2)) = 0. \end{aligned}$$

Because this equation holds for all ξ , we may set the coefficients of the powers of ξ equal to zero; the $O(1)$ and $O(\xi)$ coefficients are

$$\begin{aligned} O(1) : \quad &-c_0u_1 + c_0u_1 = 0, \\ O(\xi) : \quad &u'_1 + c'(u_0)u_1^2 = 0. \end{aligned}$$

The $O(1)$ equation holds identically, and the $O(\xi)$ equation is a differential equation for the first-order correction $u_1(t)$ in the expansion for u in region just behind the wavefront ($\xi < 0$). Because u_x is given by (2.4.11), the function $u_1(t)$ also will approximate the jump $[u_x]$ to first order. [Recall that $[u_x] = u_x(\xi-, t)$ because $u_x(\xi+, t) = 0$.] Solving the $O(\xi)$ differential equation for u_1 gives

$$u_1(t) = \frac{1}{k + c'(u_0)t}, \quad (2.4.13)$$

where k is a constant of integration that can be determined by initial data, that is, the initial jump in u_x at time $t = 0$. If $c'(u_0) > 0$ and $k < 0$, that is, the wavefront initially has a negative jump at $x = 0, t = 0$, as shown in Figure 2.14, then (2.4.13) implies that the gradient jump will steepen and the wave will break at a finite time $t_b = -k/c'(u_0)$. If $k > 0$, the jump in the gradient will tend to zero as t goes to infinity.

A wavefront calculation is one of the standard tools used in nonlinear hyperbolic problems to gain information about signal propagation along characteristics. The method lies in the domain of what is often called *weakly nonlinear* theory. When we discuss hyperbolic systems in Chapter 4, we find this technique to be a valuable tool in deriving manageable equations that approximate wavefront phenomena in nonlinear problems.

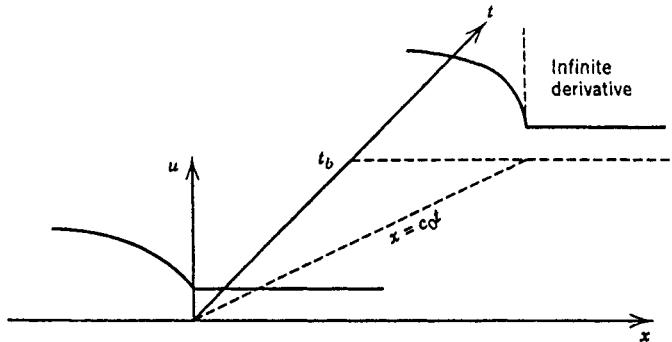


Figure 2.14 Evolution along a characteristic of a jump discontinuity in the derivative until a gradient catastrophe occurs at a breaking time t_b .

EXERCISES

1. Consider the initial value problem

$$u_t + u^2 u_x = 0, \quad x \in \mathbb{R}, \quad t > 0,$$

$$u(x, 0) = 1, \quad \text{if } x > 0; \quad u(x, 0) = \frac{2x - 1}{x - 1}, \quad \text{if } x < 0.$$

Sketch the initial signal and determine the initial jump in u_x . Sketch the characteristic diagram, and approximate the time t_b for the wave to break along the characteristic $x = t$.

2.5 General First-Order Equation

Now we consider the general first order nonlinear PDE

$$H(x, t, u, p, q) = 0, \quad p = u_x, \quad q = u_t. \quad (2.5.1)$$

on the domain $x \in \mathbb{R}, t > 0$, subject to the initial condition

$$u(x, 0) = u_0(x), \quad x \in \mathbb{R}. \quad (2.5.2)$$

In the PDE (2.5.1) there now is no obvious directional derivative that defines characteristic directions along which the PDE reduces to an ordinary differential equation. However, with some elementary analysis we can discover such directions. Let C be a curve in spacetime given by

$$x = x(s), \quad t = t(s),$$

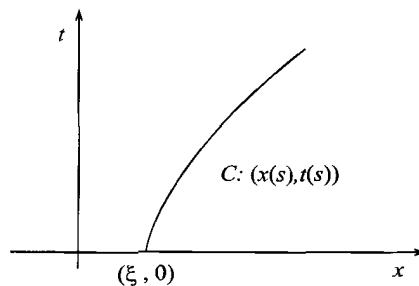


Figure 2.15 Characteristic emanating from $(\xi, 0)$.

where s is a parameter (see Figure 2.15). Then the total derivative of u along C is

$$\frac{du}{ds} = u_x x'(s) + u_t t'(s) = px' + qt'.$$

We ask whether there is a special direction (x', t') that has special significance for (2.5.1). We first calculate the total derivatives of p and q along C . To this end

$$p' = \frac{d(u_x)}{ds} = u_{xx}x' + u_{xt}t', \quad (2.5.3)$$

$$q' = \frac{d(u_t)}{ds} = u_{tx}x' + u_{tt}t'. \quad (2.5.4)$$

Now take the two partial derivatives (with respect to x and with respect to t) of the PDE (2.5.1) to obtain

$$H_x + H_u p + H_p u_{xx} + H_q u_{tx} = 0, \quad (2.5.5)$$

$$H_t + H_u q + H_p u_{xt} + H_q u_{tt} = 0. \quad (2.5.6)$$

Comparing (2.5.3) and (2.5.4) with the last two terms in (2.5.5) and (2.5.6) suggests a judicious choice for the direction (x', t') to be

$$x' = H_p, \quad t' = H_q. \quad (2.5.7)$$

In this case equations (2.5.3) and (2.5.4) combine to become

$$p' = -H_x - H_u p, \quad q' = -H_t - H_u q, \quad (2.5.8)$$

and the total derivative of u along C becomes

$$u' = pH_p + qH_q. \quad (2.5.9)$$

Let us summarize our results. If *characteristic curves* are defined by the system of differential equations (2.5.7), then (2.5.8) and (2.5.9) hold along these curves;

the latter form a system of ordinary differential equations that dictate how u , p , and q change along the curves. In other words, (2.5.8) and (2.5.9) hold along (2.5.7). The entire set of equations (2.5.7)–(2.5.9) is called the *characteristic system* associated with the nonlinear PDE (2.5.1). It is a system of five ordinary differential equations for x , t , u , p , and q .

In principle, therefore, we can develop an algorithm to solve the initial value problem (2.5.1)–(2.5.2). We emphasize that the following calculations are made assuming that a smooth solution exists. The initial condition (2.5.2) translates into

$$x = \xi, \quad u = u_0(\xi) \quad \text{at } t = 0, \quad (2.5.10)$$

and hence

$$p = u'_0(\xi), \quad (2.5.11)$$

where ξ is, as before, a real number that parameterizes the characteristics curves, representing their intersection with the x axis. To solve the characteristic system we would also need a condition on q at $t = 0$. Such a condition can be obtained from the PDE (2.5.1) itself, for we can evaluate the PDE along the $t = 0$ timeline to obtain

$$H(\xi, 0, u_0(\xi), u'_0(\xi), q) = 0. \quad (2.5.12)$$

Assuming that (2.5.12) can be solved for q (the condition $H_q \neq 0$ for $t = 0$ guarantees this), we obtain

$$q = q(\xi) \quad \text{at } t = 0. \quad (2.5.13)$$

Finally, the characteristic system (2.5.7)–(2.5.9) can be solved, subject to the initial conditions (2.5.10), (2.5.11), and (2.5.13), to obtain t , x , u , p , and q along a characteristic curve. In some cases the parameters may be eliminated to obtain an explicit form $u = u(x, t)$ for the solution.

Example. Consider the initial value problem

$$\begin{aligned} u_t + u_x^2 &= 0, \quad x \in \mathbb{R}, \quad t > 0 \\ u(x, 0) &= x, \quad x \in \mathbb{R}. \end{aligned} \quad (2.5.14)$$

Here

$$H(x, t, u, p, q) = q + p^2 = 0, \quad (2.5.15)$$

and the initial timeline can be parameterized by

$$x = \xi, \quad t = 0, \quad u = \xi, \quad p = 1. \quad (2.5.16)$$

The PDE (2.5.15) then gives at $t = 0$ a condition on q :

$$q = -1. \quad (2.5.17)$$

The characteristic system (2.5.7)–(2.5.9) corresponding to (2.5.15) is

$$x' = 2p, \quad t' = 1, \quad u' = 2p^2 + q, \quad p' = 0, \quad q' = 0. \quad (2.5.18)$$

Because (2.5.18) is autonomous, we may assume that the initial data are given at $s = 0$. Solving this system is straightforward in the present case (for other problems, solving the characteristic system may be difficult or impossible). Clearly, $t = s$ and

$$p = c_1 \quad \text{and} \quad q = c_2, \quad (2.5.19)$$

where c_1 and c_2 are constants. Therefore, from (2.5.18), we obtain

$$x' = 2c_1 \quad \text{or} \quad x = 2c_1 t + c_3, \quad (2.5.20)$$

where c_3 is another constant. Next, from (2.5.18), we have

$$u' = 2c_1^2 + c_2,$$

and consequently

$$u = (2c_1^2 + c_2)t + c_4. \quad (2.5.21)$$

Applying the initial conditions (2.5.16) and (2.5.17) allows us to determine the constants c_1, \dots, c_4 . Easily we obtain

$$c_1 = 1, \quad c_2 = -1, \quad c_3 = \xi, \quad c_4 = \xi,$$

and therefore

$$x = 2t + \xi, \quad u = t + \xi.$$

In this instance we can eliminate the parameter ξ and obtain the analytic solution

$$u(x, t) = x - t.$$

The characteristic curves in this problem are given by $x = 2t + \xi$ and are straight lines in spacetime moving with speed 2. \square

Example. Consider the initial value problem

$$H(x, t, u, p, q) = 2tq + 2xp - p^2 - 2u = 0, \quad x \in \mathbb{R}, \quad t > 0 \quad (2.5.22)$$

$$u(x, 0) = u_0(x), \quad x \in \mathbb{R} \quad (2.5.23)$$

At $t = 0$ we have $x = \xi$, $u = u_0(\xi)$, and $p = u'_0(\xi)$. To complete the initial data we need q at $t = 0$. From (2.5.22) we have

$$H(\xi, 0, u_0(\xi), u'_0(\xi), q) = 0$$

But q drops out of this equation and we are unable to determine q at $t = 0$. Thus the PDE (2.5.22) does not determine the initial derivative in the time

direction, and therefore not enough information is given to move away from the initial timeline. Notice that the characteristic direction $(x', t') = (H_p, H_q) = (2x - 2p, 2t)$ is, at $t = 0$, given by $(2\xi - 2u'_0(\xi), 0)$, which is a vector in the same direction as the initial time line. Thus the characteristic curves, which are to carry the initial data into the region $t > 0$, are tangent to the initial time line and do not have a component in the time direction to carry signals forward. Thus (2.5.22)–(2.5.23) is not a well-posed problem. \square

We summarize the comments in the preceding paragraph by making the following formal statement: A necessary condition that the initial value problem (2.5.1)–(2.5.2) have a smooth solution is that equation (2.5.1) with $t = 0$ uniquely determine q as a function of x , u , and p . A condition that guarantees this solvability is that $H_q \neq 0$ at each point $(x, 0, u, p, q)$. Under this condition it can be shown that a local solution exists about each point $(x, 0)$ on the initial timeline.

This same procedure can be extended to solve the PDE (2.5.1) subject to data given along any curve Γ in spacetime; that is, we consider the PDE (2.5.1) with the initial condition (2.5.2) replaced by the boundary condition

$$u(x, t) = f(x, t) \quad \text{on } \Gamma, \quad (2.5.24)$$

where f is a given function (see Figure 2.16). If we parameterize the curve Γ by $x = x(\xi)$, $t = t(\xi)$, condition (2.5.24) becomes

$$u = f(x(\xi), t(\xi)) = F(\xi) \quad \text{along } \Gamma.$$

As in the initial value problem, we must be able to determine p and q along the curve Γ so that we have enough data to solve the characteristic system. The PDE (2.5.1) must hold along Γ , so that

$$H(t(\xi), x(\xi), F(\xi), p(\xi), q(\xi)) = 0. \quad (2.5.25)$$

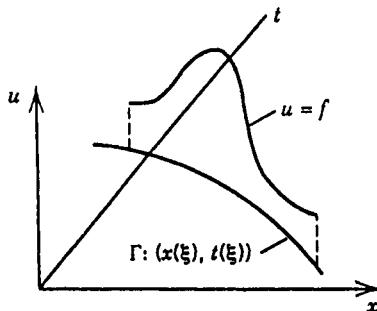


Figure 2.16 Illustration of Cauchy data where u is prescribed along a curve Γ .

Further, taking $d/d\xi$ of u (i.e., differentiating u along Γ), we have

$$\frac{dF}{d\xi} = p(\xi) \frac{dx}{d\xi} + q(\xi) \frac{dt}{d\xi}. \quad (2.5.26)$$

Equations (2.5.25) and (2.5.26) represent two equations for $p(\xi)$ and $q(\xi)$. We expect to be able to solve these equations when the Jacobian is nonzero, or

$$H_p \frac{dt}{d\xi} - H_q \frac{dx}{d\xi} \neq 0. \quad (2.5.27)$$

Geometrically, the solvability condition (2.5.27) requires that the characteristic direction $(x', t') = (H_p, H_q)$ traverse the boundary curve Γ , which has direction $(dx/d\xi, dt/d\xi)$. In other words, (2.5.27) means that the boundary curve Γ must nowhere have characteristic direction. As noted earlier, if the curve along which the data are prescribed has characteristic direction at a point on the curve, not enough information is supplied to carry data off that curve into the region where the problem is to be solved.

The initial value problem (2.5.1)–(2.5.2), or the more general problem (2.5.1) with data (2.5.24) given on a curve Γ , is called a *Cauchy problem*. The general theory of Cauchy problems (existence, uniqueness, and regularity of solutions) is discussed in many of the references [see, e.g., John 1982].

2.5.1 Complete Integral

Another class of important solutions to the general nonlinear equation

$$H(x, y, u, p, q) = 0, \quad p = u_x, \quad q = u_y \quad (2.5.28)$$

are those that depend on two independent parameters, a and b . A *complete integral* of (2.5.28) is a two-parameter family of surfaces

$$f(x, y, u, a, b) = 0$$

that implicitly defines a solution u . We are using independent variables x and y in place of x and t because of many applications to surface theory and to optics. The characteristic system is usually written in the form

$$\frac{dx}{H_p} = \frac{dy}{H_q} = \frac{du}{pH_p + qH_q} = \frac{dp}{-H_x - H_up} = \frac{dq}{-H_y - H_uq} = ds. \quad (2.5.29)$$

Notice, for example, that the ratio $dp/0$ means $p = \text{const}$. A complete integral can be used in some cases to solve the PDE subject to a condition along some given curve.

Example. The *eikonal equation* from optics is

$$u_x^2 + u_y^2 = n^2, \quad (2.5.30)$$

where n is a constant. Thus $H \equiv p^2 + q^2 = n^2 = 0$. The characteristic system is

$$\frac{dx}{2p} = \frac{dy}{2q} = \frac{du}{2p^2 + 2q^2} = \frac{dp}{0} = \frac{dq}{0}.$$

Therefore p and q are constant, and using (2.5.30), we get

$$p = a, \quad q = \sqrt{n^2 - a^2} = \text{const.}$$

Taking the differential of $u = u(x, y)$ then gives

$$du = pdx + qdy = adx + \sqrt{n^2 - a^2} dy.$$

Therefore

$$u = ax + \sqrt{n^2 - a^2} y + b,$$

where b is another arbitrary constant. Therefore we have determined a complete integral, which is a two-parameter family of planes, of the eikonal equation. Clearly this method can be easily adapted to any first-order PDE that depends only on p and q . \square

Example. Consider the Cauchy problem

$$p^2 q = 1, \quad u(x, 0) = 3x + 1.$$

A complete integral is

$$u = ax + \frac{1}{a^2} y + b.$$

Applying the initial condition gives $u(x, 0) = ax + b = 3x + 1$, giving $a = 3$ and $b = 1$. Therefore

$$u = 3x + \frac{1}{9}y + 1. \quad \square$$

EXERCISES

1. Solve the initial value problem

$$\begin{aligned} u_t + u_x &= t, \quad x \in \mathbb{R}, \quad t > 0, \\ u(x, 0) &= 0, \quad x \in \mathbb{R}. \end{aligned}$$

2. Solve the initial value problem

$$\begin{aligned} u_t + u_x &= 0, \quad x \in \mathbb{R}, \quad t > 0, \\ u(x, 0) &= -x^2, \quad x \in \mathbb{R}. \end{aligned}$$

Does the solution exist for all $t > 0$?

3. Find the characteristic system associated with the *Hamilton–Jacobi equation*

$$u_t + h(t, x, u_x) = 0,$$

where h is a given function called the *Hamiltonian*.

4. Find a first-order PDE associated with each of the following surfaces (a and b are parameters).

$$\begin{aligned} \text{a)} \quad & x^2 + y^2 + (u - a)^2 = r^2. \\ \text{b)} \quad & (x - a)^2 + (y - b)^2 + u^2 = r^2. \end{aligned}$$

5. Assuming that no independent variables occur, that is, if the PDE has the form $H(u, p, q) = 0$, show that a complete integral is

$$\int \frac{du}{f(u, a)} = x + ay + b,$$

where, from the PDE, $p = f(u, a)$.

6. Find a complete integral of the PDE

$$u_x^2 + yu_y - u = 0.$$

7. Determine the characteristic system associated with the nonlinear equation

$$u_t + c(u)u_x = 0,$$

and show that it coincides with the results obtained in Section 2.2.

8. Use the method of characteristics to determine two different solutions to the initial value problem

$$\begin{aligned} u &= p^2 - 3q^2, \quad x \in \mathbb{R}, \quad t > 0, \\ u(x, 0) &= x^2, \quad x \in \mathbb{R}. \end{aligned}$$

9. Solve

$$\begin{aligned} u_t &= u_x^3, \quad x \in \mathbb{R}, \quad t > 0, \\ u(x, 0) &= 2x^{3/2}, \quad x \in \mathbb{R}. \end{aligned}$$

10. Consider the Helmholtz equation

$$(\Delta + k^2)u = 0, \quad x \in \mathbb{R}^2,$$

where k is a large parameter. By writing $u = A(x)e^{ik\psi}$, show that to order $O(k^2)$, the function $\psi = \psi(x)$ satisfies the eikonal equation

$$\operatorname{grad} \psi \cdot \operatorname{grad} \psi = 1.$$

2.6 A Uniqueness Result

Up to this point our approach has been algorithmic in nature; that is, most efforts have gone into the actual construction of solutions, and we have essentially ignored existence and uniqueness questions. Having constructed a solution in some manner, one may legitimately ask whether a different construction could give a different solution. Both to answer this question and to give the reader a flavor of a simple theoretical result, we prove a basic uniqueness result due to A. Haar in 1928. The proof given below is an adaptation of that given in Courant & Hilbert (1962), and it applies to a specific class of nonlinear equations.

We require one definition, namely, what it means for a function to satisfy a Lipschitz condition. A function $f(x, y)$ satisfies a *Lipschitz condition with respect to y* on a domain D in the plane if, and only if, there exists a constant $k > 0$ such that

$$|f(x, y_1) - f(x, y_2)| \leq k|y_1 - y_2|$$

for all (x, y_1) and (x, y_2) in D . The constant k is called the *Lipschitz constant with respect to y* .

The reader may have encountered this concept in ordinary differential equations; it is a condition on $f(x, y)$ that guarantees uniqueness of solutions to the initial value problem $dy/dx = f(x, y), y(x_0) = y_0$. It is not difficult to show that the property of being Lipschitz is stronger than continuity, yet weaker than differentiability.

We may extend the concept to several variables. For example, a function $f(x, y, z)$ is said to be Lipschitz in y and z if there exist two positive constants k and m such that

$$|f(x, y_1, z_1) - f(x, y_2, z_2)| \leq k|y_1 - y_2| + m|z_1 - z_2|$$

for all (x, y_1, z_1) and (x, y_2, z_2) . The constants k and m are the Lipschitz constants with respect to y and z , respectively.

Now the uniqueness theorem of Haar.

Theorem. Consider a PDE of the form

$$u_t = G(x, t, u, p), \quad p = u_x, \tag{2.6.1}$$

where G is continuous and satisfies a Lipschitz condition with respect to u and p in \mathbb{R}^4 , and k is the Lipschitz constant with respect to p . Let u and v be smooth solutions of (2.6.1) such that

$$u(x, 0) = v(x, 0), \quad x_1 \leq x \leq x_2.$$

Then $u(x, t) = v(x, t)$ on the triangle

$$T = \{(x, t) : t \geq 0, t \leq k^{-1}(x - x_1), t \leq k^{-1}(x_2 - x)\}.$$

The proof is as follows. Let $w = u - v$. We need to show that $w = 0$ on T . First, using the Lipschitz property, we obtain

$$\begin{aligned} |w_t| &= |u_t - v_t| = |G(t, x, u, u_x) - G(t, x, v, v_x)|, \\ &\leq a|u - v| + k|u_x - v_x| = a|w| + k|w_x|. \end{aligned}$$

At points where $w > 0$, the last inequality can be written

$$w_t < bw + k|w_x|, \quad (2.6.2)$$

where $b > a$. Now we define the function W by $W = we^{-bt}$. It clearly suffices to show that $W = 0$ on T . We proceed by contradiction and assume without loss of generality that W is positive at some point in T . Let P be a point in T (a closed bounded set) where W assumes a local maximum, where, of course, $W > 0$. The point P cannot be on the base of T because $W = 0$ there. Thus P must be in the interior of T or on the lateral sides. In this case (see Figure 2.17) the directions $(-k, -1)$ and $(k, -1)$ both point into T . Therefore, the directional derivatives of W in these two directions must be nonpositive (W has a maximum at P):

$$\begin{aligned} (-k, -1)(W_x, W_t) &= -kW_x - W_t \leq 0, \\ (k, -1)(W_x, W_t) &= kW_x - W_t \leq 0. \end{aligned}$$

Consequently

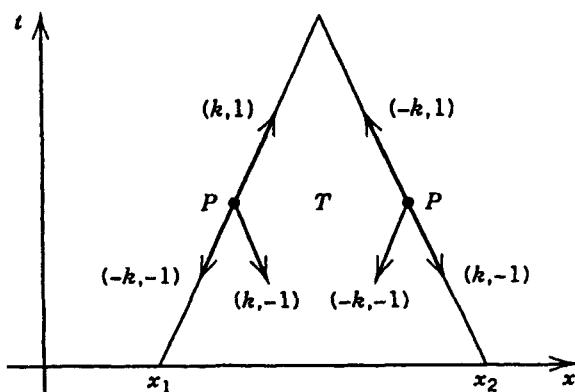


Figure 2.17 Triangle T

$$W_t \geq k|W_x| \quad \text{at } P,$$

or, in terms of w

$$w_t \geq bw + k|w_x| \quad \text{at } P.$$

This contradicts (2.6.2), which completes the proof. \square

2.7 Models in Biology

2.7.1 Age Structure

A key problem in demography is to determine how the age structure of a population evolves in time. The age structure is described by an age distribution function $u = u(a, t)$, where $u(a, t)da$ represents the approximate number of females at time t between the ages a and $a + da$. Often females are studied in demographic models because they have a well-defined beginning and end to their reproductive capacities. For simplicity, we write the age range as $0 \leq a < \infty$ even though the age at death is finite. The total female population at time t is

$$N(t) = \int_0^\infty u(a, t)da.$$

The goal is to develop a model that dictates how the age distribution responds under influence of birth and death forces.

The governing equation is a simple population balance law. We derive it using a *small-box* method. In an arbitrary small interval $[a, a + da]$ the rate of change of the number of individuals must equal the rate that they enter the interval at age a , minus the rate they leave at age $a + da$, minus the rate that die. The approximate number of females the interval is $u(a, t)da$, and therefore

$$\frac{\partial}{\partial t}[u(a, t)da] = u(a, t) - u(a + da, t) - m(a)u(a, t)da,$$

where the nonnegative function $m(a)$ is the age-specific death rate. Dividing by da and passing to the limit as $da \rightarrow 0$ gives

$$u_t = -u_a - m(a)u.$$

Thus, the balance law is an advection equation with speed 1 with a sink term given by the mortality rate; notice that the flux is $\phi = u$, or the number crossing the age a line at time t is just $u(a, t)$, which is the population density at that age. This model has its origins in the work of McKendrick in 1926 and in subsequent studies by Von Foerster in 1959, and it is called the *McKendrick–Von Foerster equation*.

At time $t = 0$ we assume a given age distribution

$$u(a, 0) = f(a), \quad a \geq 0.$$

To account for births we assume that there is a nonnegative fecundity rate $b(a, t)$, which depends on both time and age of the female, where $b(a, t)$ is the average number of offspring per female of age a at time t ; it is called the *maternity function*. We expect b to be zero until the age of onset of maturity and zero after menopause. In between these ages the fecundity rate varies according to age; for example, in humans, females of age 25 produce more offspring than do women of age 35. Because $u(a, t)da$ is the number of females between a and $a + da$, and b is the average reproduction rate, then the number of offspring produced by females from age a to age $a + da$ is $b(a, t)u(a, t)da$. Thus, the total number of offspring produced by all females is the integral (sum) over all ages, or

$$B(t) = \int_0^\infty b(a, t)u(a, t)da.$$

Observe that $B(t)$ is precisely $u(0, t)$, so this equation defines a boundary condition at $a = 0$.

Figure 2.18 depicts the evolution of age structure profiles at fixed times in three-dimensional atu space. Age-structured models can be represented visually as a conveyor belt moving at speed one (Figure 2.19). Grains of sand on the belt represent the female population density, and the belt contains holes through which the sand falls, representing mortality. A scale is located at the ages of fertility, $a_m < a < a_M$; it weighs the female population and sends an electrical signal to the valve in the funnel at $a = 0$, which releases newborns to the population at the fecundity rate. At a_L , the maximum lifetime and length of the belt, the population density is zero. [t]

In summary, the age-structured model is given by

$$u_t = -u_a - m(a)u, \quad a > 0, \quad t > 0, \quad (2.7.1)$$

$$u(a, 0) = f(a), \quad a \geq 0, \quad (2.7.2)$$

$$u(0, t) = \int_0^\infty b(a, t)u(a, t)da, \quad t > 0. \quad (2.7.3)$$

What makes this problem especially interesting, and difficult, is that the left boundary condition at age $a = 0$ is not known, but rather depends on the solution $u(a, t)$, which is also unknown. This type of condition is called a *nonlocal boundary condition* because it depends on the integrated unknown solution in the problem.

Example. (Stable Age Structure) Rather than attempting to solve (2.7.1)–(2.7.3) directly, we can ignore the initial condition and ask what happens over

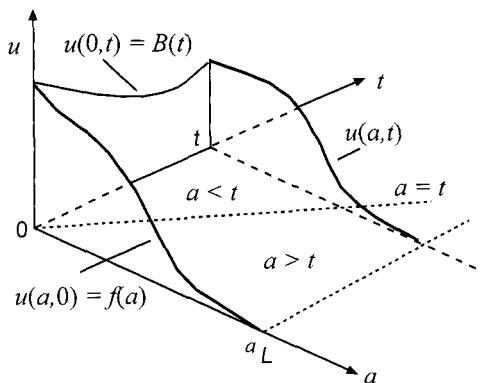


Figure 2.18 Age-structured model. Here, $f(a)$ is the initial, known age structure, and $u(0, t) = B(t)$ is the unknown offspring at age $a = 0$ and time t . The age structure $u(a, t)$ for $a > t$ is affected only by the initial population $f(a)$, whereas for $a < t$ it is affected by the entire population and its fecundity. a_L is the maximum lifetime and individuals follow paths $a = t + \text{constant}$ in age-time space.

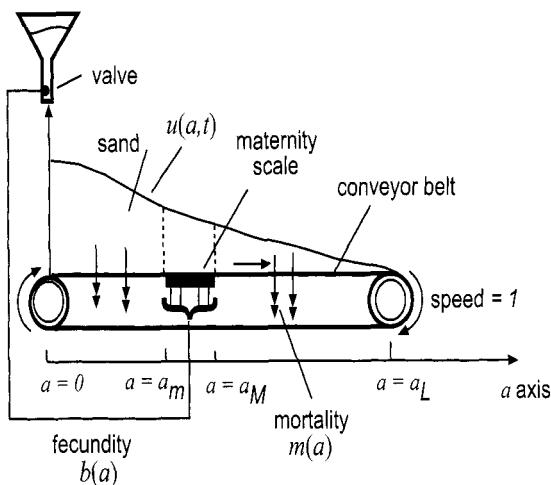


Figure 2.19 Conveyor belt visualization of an evolving age structure.

a long time. Births from the initial population $f(a)$ only affect the solution for a finite time because those individuals and their offspring die. Therefore, in the case that the maternity function is independent of time [i.e., $b = b(a)$], we can

look for an age distribution of the form

$$u(a, t) = U(a)e^{rt}, \quad t \text{ large},$$

where U is an unknown age structure and r is an unknown growth rate. Substituting into the PDE (2.7.1) and making reductions gives an ODE for U :

$$U'(a) = -(m(a) + r)U(a).$$

This equation can be solved by separation of variables to get

$$U(a) = Ce^{-ra}e^{-\int_0^a m(s)ds},$$

where C is a constant. Letting $S(a) = \exp(-\int_0^a m(s)ds)$, called the *survivorship function* (which is the probability of surviving to age a), we can write the long-time solution as

$$u(a, t) = Ce^{rt-ra}S(a). \quad (2.7.4)$$

To determine r we substitute (2.7.4) into the nonlocal boundary condition (2.7.3) to obtain

$$1 = \int_0^\infty b(a)e^{-ra}S(a)da. \quad (2.7.5)$$

This is the classic *Euler–Lotka equation*. The right side of (2.7.5) is a decreasing function of r ranging from infinity to zero, and therefore there is a unique value of r that satisfies the equation. If $r > 1$, then the population will grow; if $r < 1$, the population will die out. In the special case $m = \text{const}$, the Euler–Lotka equation is

$$1 = \int_0^\infty b(a)e^{-(r+m)a}da. \quad \square$$

Example. (The Renewal Equation) The characteristic method can be used to study (2.7.1)–(2.7.3) in the simple case when $b = b(a)$ and $m = \text{constant}$. The PDE (2.7.1) is

$$u_t = -u_a - mu, \quad a > 0, \quad t > 0. \quad (2.7.6)$$

If we change independent variables via (characteristic coordinates)

$$\xi = a - t, \quad \tau = t,$$

then (2.7.6) becomes

$$U_\tau = -mU, \quad \text{where } U = U(\xi, t).$$

Then

$$U(\xi, \tau) = C(\xi)e^{-m\tau},$$

where C is an arbitrary function. In terms of the original variables, we have

$$u(a, t) = C(a - t)e^{-mt}.$$

To determine the arbitrary function C , we consider two cases, $a > t$ and $a < t$. (See Figure 5.1.) The arbitrary function will be different in each case. The solution in $a > t$ is determined by the initial age structure, and we have

$$u(a, 0) = C(a) = f(a).$$

Therefore

$$u(a, t) = f(a - t)e^{-mt}, \quad a > t. \quad (2.7.7)$$

For $a < t$ the boundary condition gives

$$u(0, t) = B(t) = C(-t)e^{-mt},$$

or

$$C(s) = B(-s)e^{-ms}.$$

Consequently

$$u(a, t) = B(t - a)e^{-ma}, \quad a < t. \quad (2.7.8)$$

The solution to (2.7.1)–(2.7.3) in the case $m(a) = m$ and $b(a, t) = b(a)$ is given by (2.7.7)–(2.7.8), but B is still unknown. To find B , we substitute the expressions (2.7.7)–(2.7.8) into the yet unused nonlocal boundary condition (2.7.3), after breaking up the integral into two. We obtain

$$\begin{aligned} B(t) &= \int_0^\infty b(a)u(a, t)da \\ &= \int_0^t b(a)u(a, t)da + \int_t^\infty b(a)u(a, t)da, \end{aligned}$$

or

$$B(t) = \int_0^t b(a)B(t - a)e^{-ma}da + \int_t^\infty b(a)f(a - t)e^{-mt}da. \quad (2.7.9)$$

Equation (2.7.9) is a linear integral equation for the unknown $B(t)$, and it is called the *renewal equation*. Once it is solved for $B(t)$, then (2.7.7)–(2.7.8) give the age structure for the population. For long times the second integral is zero because the maternity vanishes for large ages. Generally, (2.7.9), a non-homogeneous Volterra equation, is difficult to solve and must be dealt with numerically, or by successive approximation (iteration). See, for example, Logan (2006a). \square

2.7.2 Structured Predator–Prey Model

Nonlinearities can enter demography in various ways. For example, the birth and death schedules may depend the total population $N(t)$, or there may be other populations that affect the mortality rate; for example, in the case of an animal population, predators may consume the animals. In this section we study a predator–prey model and show how, using the *method of moments* (a method akin to an energy method), the problem can be reduced to solving a system of ODEs. This is an important technique to add to our analytic toolbox for dealing with PDEs.

We consider a population of prey with age density $u(a, t)$ and constant per capita mortality rate m . Then, as above, the governing age–time dynamics is given by (2.7.6) and the initial condition (2.7.2). We assume that the maternity function has the form

$$b(a) = b_0 a e^{-\gamma a}.$$

Then the prey produce offspring (eggs) given by

$$B(t) = \int_0^\infty b_0 a e^{-\gamma a} u(a, t) da. \quad (2.7.10)$$

Now let us introduce a total predator that consumes the eggs of the prey population. We assume that the predator population is $P = P(t)$, and we do not consider age structure in this population. (To stimulate thinking about this model, recall that egg-eating predators is one of the theories posed for the demise of the dinosaurs.) Because predators eat only eggs ($a = 0$), the PDE (2.7.6) is unaffected. What is affected is the number of offspring $u(0, t)$ produced. Thus we no longer have $u(0, t) = B(t)$, but rather we must include a predation term that decreases the egg population. The simplest model is the Lotka–Volterra model (mass action), which requires that the number eggs eaten be proportional to the product of the number of eggs and the number of predators. Thus, we have

$$u(0, t) = B(t) - kB(t)P(t),$$

where k is the predation rate. Because the right side can be negative, we define $M(B, P) = \max(B - kB P, 0)$ and take the number of eggs at $a = 0$ to be

$$u(0, t) = M(B, P). \quad (2.7.11)$$

This equation provides the boundary condition for the problem. Finally, we impose Lotka–Volterra dynamics on the predator population, or

$$\frac{dP}{dt} = -\delta P + cBP, \quad (2.7.12)$$

where δ is the per capita mortality rate. Hence, in the absence of eggs, predators die out. Initially, we take $P(0) = P_0$. In summary, the model is given by the PDE (2.7.6), the initial condition (2.7.2), the boundary condition (2.7.11), and the predator equation (2.7.12).

We remark that if the predators consumed prey other than eggs, then a predation term would have to be included as a sink term on the right side of the dynamical equation (2.7.6).

The method of moments allows us to obtain a system of ordinary differential equations for the total prey and predator populations $N(t)$ and $P(t)$. In the analysis, we will also obtain equations for some additional auxiliary variables, but the end result is a system of ODEs, which is simpler than the mixed PDE–ODE system. The idea is to multiply the PDE (2.7.6) by some *moment* function $g(a)$ and then integrate over $0 \leq a \leq \infty$. The only requirement is that $u(a, t)g(a) \rightarrow 0$ as $a \rightarrow \infty$. On taking g to be different functions, we can obtain equations that lead to the differential equations that we seek. The reader will find it valuable to verify these calculations. Proceeding in general, we multiply the PDE by g and integrate to obtain

$$\frac{d}{dt} \int_0^\infty g(a)u(a, t)da = - \int_0^\infty g(a)u_a(a, t)da - m \int_0^\infty g(a)u(a, t)da.$$

The first integral on the right can be integrated by parts to get

$$\frac{d}{dt} \int_0^\infty g(a)u(a, t)da = M(B, P)g(0) + \int_0^\infty g'(a)u(a, t)da - m \int_0^\infty g(a)u(a, t)da. \quad (2.7.13)$$

Now we make different choices for g . If $g(a) = 1$, then (2.7.13) becomes simply

$$\frac{dN}{dt} = M(B, P) - mN, \quad (2.7.14)$$

an ODE involving N , P , and B . If we take $g(a) = b(a)$, the maternity function, then (2.7.13) becomes

$$\frac{dB}{dt} = -\gamma B + b_0 H - mB, \quad (2.7.15)$$

where $H = H(t)$ is defined by

$$H(t) = \int_0^\infty e^{-\gamma a} u(a, t)da.$$

But now H is yet a new variable. To obtain an equation involving H , we take $g(a) = e^{-\gamma a}$. Then (2.7.13) becomes

$$\frac{dH}{dt} = M(B, P) - (m + \gamma)H. \quad (2.7.16)$$

Therefore we have four ODEs, (2.7.14), (2.7.15), (2.7.16), and (2.7.12) for N , P , B , and H , respectively. Clearly the N equation decouples from the system and we can consider just the three ODEs

$$\frac{dP}{dt} = -\delta P + cBP, \quad \frac{dB}{dt} = -(m + \gamma)B + b_0H, \quad \frac{dH}{dt} = M(B, P) - (m + \gamma)H.$$

The initial conditions are $P(0) = P_0$, $B(0) = \int_0^\infty b_0ae^{-\gamma a}f(a)da$, and $H(0) = \int_0^\infty e^{-\gamma a}f(a)da$. We may now proceed with a numerical method to solve the system and determine the resulting dynamics. A sample calculation is requested in the Exercises.

2.7.3 Chemotherapy

We introduced two types of structure in a biological context—spatial structure and age structure. It is intuitively clear that any quantity that is characteristic of an organism's state can be used as a structural variable; these include development or maturation level, length, weight, and so on. In this section we study a simple model of leukemia cancer cell maturation and the effects of a chemotherapy regimen as a control mechanism. The model was introduced by Bischoff et al. (1971), and a more detailed motivation is given in Edelstein-Keshet (2005, pp 463ff).

We assume that malignant cells undergo a maturation process measured by a physiological variable x , where x is normalized so that $0 \leq x \leq 1$. At $x = 0$ cells are created from parent cells that divide into two at maturity, which is $x = 1$. We assume a constant rate v of cell maturation, or $dx/dt = v$, and we let $u = u(x, t)$ denote the density of cells at maturation stage x at time t ; in particular, $u(x, t)dx$ is the approximate number of cells have maturation between x and $x + dx$. The relevant quantities are summarized below:

- x = maturation level of a malignant cell
- $u(x, t)$ = maturation density of malignant cells
- v = rate of maturation of cells
- $c(t)$ = chemotherapeutic drug concentration

A conservation law describes the dynamics of cell growth and death via chemotherapy. We take the death rate to be of the form $m(t)u$, where

$$m(t) = \frac{\alpha c(t)}{\beta + c(t)}.$$

This rate has the form of Michaelis-Menten enzyme kinetics. It should be clear from previous discussion that the model equations are

$$u_t + vu_x = -m(t)u, \quad 0 < x < 1, t > 0, \quad (2.7.17)$$

$$u(x, 0) = u_0, \quad 0 \leq x \leq 1 \quad (2.7.18)$$

$$u(0, t) = 2u(1, t), \quad t > 0, \quad (2.7.19)$$

where u_0 is the initial maturation distribution of malignant cells, and where the boundary condition (2.7.19) is interpreted as parent cell division creating two daughter cells.

We can analyze the model by the characteristic method. Along the characteristics $x = vt + \xi$ the PDE reduces to $du/dt = -m(t)u$, which leads to a general solution

$$u(x, t) = K(x - vt)e^{-M(t)},$$

where K is an arbitrary function and $M(t) = \int_0^t m(s)ds$.

For $x > vt$ we have $u(x, 0) = K(x) = u_0$. Hence the solution is

$$u(x, t) = u_0 e^{-M(t)}, \quad x > vt. \quad (2.7.20)$$

For $x < vt$ the boundary condition (2.7.19) gives

$$K(-vt)e^{-M(t)} = 2K(1 - vt)e^{-M(t)},$$

which means that K satisfies the functional relation $K(z) = K(1 + z)$. It is easy to see that a solution has the form $K(z) = K_0 e^{az}$. Substituting into the relation gives $a = -\ln 2$. Therefore $K(z) = K_0 e^{-z \ln 2}$, and, using the fact that $u(0, 0) = u_0$, we have

$$u(x, t) = u_0 e^{-x \ln 2} e^{vt \ln 2} e^{-M(t)}, \quad x < vt. \quad (2.7.21)$$

Equations (2.7.20)–(2.7.21) give the solution of (2.7.17)–(2.7.19).

Observe that the total number of cells at any time $t > 1/v$ is

$$U(t) = \int_0^1 u(x, t) dx = u_0 e^{vt \ln 2} e^{-M(t)} \int_0^1 e^{-x \ln 2} dx = \frac{u_0}{2 \ln 2} e^{vt \ln 2} e^{-M(t)}.$$

See Exercise 7 for a numerical example.

2.7.4 Mass Structure

We modeled the age structure of a population by the age density $u = u(a, t)$, where $u(a, t)da$ is the approximate number of individuals between ages a and $a + da$. We found

$$u_t = -u_a - \mu u,$$

where μ is the per capita mortality rate. Age is just one of the many structure variables that demographers and ecologists study. Rather than age a , we might rather consider mass m , length x , weight w , development stage ξ , or any other physiological variable attached to the individuals.

For example, in a mass-structured population, $u(m, t)dm$ is the approximate number of individuals having mass between m and $m + dm$ at time t , where $u = u(m, t)$ is the mass density of the population, given in dimensions of individuals per mass. At a fixed t , a graph of $u(m, t)$ versus m gives the mass structure of the population. As in the age structured model, a conveyor belt picture aids in visualizing the dynamics (see Figure 2.19). To obtain a growth law we start with the basic conservation law

$$u_t = -\phi_m - \mu u,$$

where $\phi = \phi(m, t)$ measures the flux *through mass space* of individuals having mass m at time t . The flux has dimensions of individuals per time. How fast individuals move depends on their mass growth rate $g = g(m, t)$, or the rate that mass is accumulated by an individual of mass m , given in dimensions of mass per time. For an individual, we have

$$\frac{dm}{dt} = g(m, t).$$

The assumption is that all individuals of the same mass experience the same growth rate. The t dependence in g comes from environmental effects, for example, food availability or quality, which may vary over time. Therefore the flux is given by

$$\phi = gu,$$

and the balance law is

$$u_t = -(gu)_m - \mu u, \quad (2.7.22)$$

where the per capita mortality rate is $\mu = \mu(m, t)$.

Next we determine how individuals are recruited to the population. We make the simplifying assumption that all individuals are born with mass m_b , and therefore the domain of the problem is $t \geq 0$ and $m_b \leq m \leq m_f$, where m_f is the maximum possible mass that an individual can accumulate over its

lifespan. Thus, $u(m_f, t) = 0$. If we integrate (2.7.22) over the range of masses, we obtain

$$\begin{aligned}\frac{d}{dt} \int_{m_b}^{m_f} u \, dm &= - \int_{m_b}^{m_f} (gu)_m \, dm - \int_{m_b}^{m_f} \mu u \, dm \\ &= g(m_b, t)u(m_b, t) - g(m_f, t)u(m_f, t) - \int_{m_b}^{m_f} \mu u \, dm.\end{aligned}$$

The term on the left is the growth rate of the entire population. The second term on the right is zero, and the last term is the death rate. Therefore, individuals are recruited to the population at rate $g(m_b, t)u(m_b, t)$. But births are due to the reproduction of individuals during their period of fertility (as in age-structured models we restrict the analysis to the female population). Let $b(m, t)$ denote the maternity function, or birth rate, given in units of offspring per female per time. Typically, b will be zero except over the domain of masses where females are fertile. The time dependence arises from environmental factors such as food availability. The rate that females produce offspring is therefore

$$\int_{m_b}^{m_f} b(m, t)u(m, t) \, dm,$$

and thus we have the boundary condition

$$g(m_b, t)u(m_b, t) = \int_{m_b}^{m_f} b(m, t)u(m, t) \, dm. \quad (2.7.23)$$

In summary, the PDE (2.7.22), an initial condition

$$u(m, 0) = f(m), \quad m_b \leq m \leq m_f, \quad (2.7.24)$$

and the boundary condition (2.7.23) define a well-posed mathematical problem to determine how the mass structure of the entire population evolves. As in the age-structured model, the mass-structured model is punctuated by a nonlocal boundary condition.

2.7.5 Size-Dependent Predation

The central problem in population ecology is to understand the factors that regulate animal and plant populations. Therefore, models of consumer–resource interactions, and especially predation, are key in quantitatively studying some of these fundamental mechanisms. The basic models develop the dynamics of unstructured prey and predator populations $u = u(t)$ and $p = p(t)$, respectively. The simplest model is the familiar *Lotka–Volterra* model

$$u' = ru - aup, \quad p' = -dp + bup,$$

where r and d are the per capita growth and mortality rates of the prey and predator, respectively, and the predation rate, given by mass action kinetics, proportional to the product up of the populations. This model predicts oscillating populations around an equilibrium $(d/b, r/a)$. A step up in detail is, for example, the *Rosenzweig–MacArthur* model

$$u' = ru \left(1 - \frac{u}{K}\right) - \frac{au}{1 + ahu} p, \quad p' = -dp + c \frac{au}{1 + ahu} p,$$

where the prey grow logistically and the predation rate is given by a Holling type II expression

$$\frac{au}{1 + ahu},$$

where a is the attack rate and h is the handling time for a prey item. This model has rich dynamics where, for example, as the carrying capacity increases, an asymptotically stable equilibrium bifurcates into an unstable equilibrium with the appearance of a limit cycle (a Hopf bifurcation). (See, for example, Kot 2001, pp 132ff.)

Adding structure to one or both populations leads to systems of PDEs. Depending on the type of structure imposed on the populations, the system has varying degrees of difficulty. In perhaps the simplest case, in a spatially structured population we need only add diffusion or advection terms to the dynamical model, retaining the predation terms. For example, a spatially structured Lotka–Volterra model with diffusion has the form

$$u_t = D_1 u_{xx} + ru - aup, \quad p_t = D_2 p_{xx} - dp + bup,$$

where $u = u(x, t)$ is the population density of prey at location x , and $p = p(x, t)$ is the predator density at x . The assumption here is that predation is local and has the same form at each location. These types of models are discussed extensively in Murray (2002, 2003).

Size-structured populations, in which the correlation between the predator size and the size of the prey is monitored, are more complicated. Let $u = u(x, t)$ denote the density of prey at size x and $p = p(y, t)$ denote the density of predators of size y . In this discussion assume size means mass. Generally, size ranges are $0 \leq x \leq X$ and $0 \leq y \leq Y$. If $K(x)$ denotes the set of all predator sizes y that consume prey of size x , then a mass action predation term is given by

$$au(x, t) \int_{y \in K(x)} p(y, t) dy,$$

where we have integrated over the appropriate predator density to find the total number of predators that interact with prey of size x . Therefore, the prey dynamics is given by

$$u_t = -(gu)_x - \mu(x)u - au \int_{y \in K(x)} p(y, t) dy, \quad (2.7.25)$$

where $g = g(t, x)$ is the mass growth rate of the prey. In a similar way, the population law for the predator density is

$$p_t = -(\gamma p)_y - m(y) + bp \int_{x \in I(y)} u(x, t) dx, \quad (2.7.26)$$

where $I(y)$ is the set of all prey sizes x that are consumed by predators of size y , and where $\gamma = \gamma(y, t)$ is the mass growth rate of the predator. The size dependent mortality rates for prey and predators are μ and m , respectively. In this model predation decreases the density of prey and decreases the mortality rate of the predators. This model can be extended to predation events governed by Holling type II responses.

EXERCISES

- Consider a population of organisms whose per capita death rate is 3% per month and that the fecundity rate, in births per female per age in months, is given by $b(a) = 4$ for $3 \leq a \leq 8$, and $b(a) = 0$ otherwise. Use the Euler–Lotka equation to calculate the long-term growth rate r of the population. What is the long-time population distribution?
- An age-structured population in which older persons are removed at a faster rate than younger persons, and no one survives past age $x = L$, can be modeled by the initial boundary value problem

$$\begin{aligned} u_t + u_x &= -\frac{cu}{L-x}, & 0 < x < L, \quad t > 0, \\ u(0, t) &= b(t), \quad t > 0; & u(x, 0) = f(x), \quad 0 < x < L. \end{aligned}$$

Find the age-structured population density $u(x, t)$ and give physical interpretations of the positive constant c and the positive functions $b(t)$ and $f(t)$.

- Consider the model

$$\begin{aligned} u_t &= -u_a - m(U)u, \quad a > 0, \quad t > 0 \\ u(0, t) &= \int b(U)u(a, t)da, \quad t > 0 \\ u(a, 0) &= f(a), \quad a \geq 0, \end{aligned}$$

where the mortality and birth rates depend on the total population $U = U(t)$. Assume that $b'(U) \leq 0$ and $m'(U) \geq 0$.

- Show that

$$\frac{dU}{dt} = (b(U) - m(U))U. \quad (2.7.27)$$

- (b) Let \tilde{U} be a local asymptotically stable equilibrium of the total population equation (2.7.27). Show that the associated long-time population distribution is

$$u(a) = b(\tilde{U})\tilde{U}e^{-m(\tilde{U})a}.$$

4. Consider the age-structured model in the case where the mortality rate is constant ($m(a) = m$) and the maternity function is a constant ($b(a, t) = \beta$). At time $t = 0$ assume the age distribution is $f(a) = u_0$ for $0 < a \leq \delta$, and $f(a) = 0$ for $a > \delta$.

- (a) Show that the renewal equation takes the form

$$B(t) = \int_0^t \beta B(s)e^{-m(t-s)}ds + \beta \delta u_0 e^{-mt}.$$

- (b) Show that $B(t)$ satisfies the differential equation

$$B' = (\beta - m)B,$$

and determine $B(t)$ and the population density $u(a, t)$. What is the total size of the population $N(t)$ at any time t ?

5. Consider an age-structured model where the per capita mortality rate depends on the total population $N = N(t)$ and the maternity function is $b(a) = b_0 e^{-\gamma a}$:

$$\begin{aligned} u_t &= -u_a - m(N)u, \quad a > 0, t > 0, \\ u(0, t) &= \int_0^\infty b_0 e^{-\gamma a} u(a, t) da, \\ u(a, 0) &= f(a), \quad a > 0. \end{aligned}$$

- (a) Use the method of moments to obtain the system of ODEs

$$\frac{dN}{dt} = B - m(N)N, \quad \frac{dB}{dt} = (b_0 - \gamma - m(N))B,$$

for $N(t)$ and the offspring $B(t) = u(0, t)$. (Note that the maternity function in this model is unreasonable since it provides for newborns giving birth, but it may be a good approximation for the case when the population reproduces at a very young age.)

- (b) Show that the relation $B = (b_0 - \gamma)N$ gives a solution to the ODEs in the NB plane.
- (c) Show that the solution to the system cannot oscillate and, in fact, approaches a steady state.

6. Numerically solve the system of equations (2.7.14), (2.7.15), (2.7.16), (2.7.12) and plot the prey and predator populations $N(t)$ and $P(t)$ for $0 \leq t \leq 125$. Take $b_0 = 5$ and the remaining constants to be unity. According to your calculation, is there a basis for controlling pests by introducing predators that selectively eat their eggs?
7. In the chemotherapy example assume that a cell matures in about 14.4 h, and assume the rate constants in the uptake rate are $\alpha = 0.25$ (per hour) and $\beta = 0.3 \text{ } \mu\text{g/mL}$. If the concentration $c(t)$ is maintained at a constant value of 15 mg/mL per kilogram of body weight, how long does it take in a 70 kg patient for the total number of cells to decay to 0.1% of their initial population? [Also see Edelstein-Keshet 2005.]
8. Consider the age-structured model

$$\begin{aligned} u_t &= -u_a - mu, \quad a > 0, t > 0, \\ u(0, t) &= b(N_A)N_A, \quad t > 0, \\ u(a, 0) &= f(a), \quad a > 0, \end{aligned}$$

where

$$N_A(t) = \int_T^\infty u(a, t)da$$

is the total adult population.

- (a) Interpret this model and show that N_A satisfies the differential-delay equation
$$\frac{dN_A}{dt} = b(N_A(t - T))N_A(t - T)e^{-mT} - mN_A(t).$$
- (b) Show that there is a single positive equilibrium (constant) solution in (a) and examine its stability to small perturbations.
9. In the mass-structured model given by (2.7.22), (2.7.23), and (2.7.24), suppose, rather than giving birth, that the number of individuals increases by division (such as cells, e.g.). In particular, let $b_D(m, M, t)$ be the given per capita rate that individuals of mass M divide into individuals of mass m at time t . Denote

$$\beta(m, t) = \int_{m_b}^{m_f} b(m, M, t)u(M, t)dM.$$

Formulate a mass-structured model to incorporate this modification.

10. Consider a population that is structured in both age a and mass m , and define the population density by $u = u(a, m, t)$, where $u(a, m, t) da dm$ approximates the number of individuals at time t having age in the interval $(a, a + da)$ and mass in the interval $(m, m + dm)$. As time progresses, an individual moves through age space at the speed 1, and it moves through mass space at rate $g = g(a, m, t)$. Argue that the governing equation is

$$u_t = -u_a - (gu)_m - \mu u,$$

where the mortality rate may depend on age, mass, and time, or $\mu = \mu(a, m, t)$. Another way to think about this equation is in terms of a basic conservation law in two dimensions (a and m),

$$u_t = -\operatorname{div} \phi - \mu u,$$

where ϕ is the flux vector $\phi = (u, gu)$, and $\operatorname{div} = ((\partial/\partial a), (\partial/\partial m))$.

Reference Notes. Continuous, age-structured models have received a lot of attention. Cushing (1994) has an extensive bibliography that serves as an entry point to the literature for both continuous and discrete models. Relevant to the discussion in the last section, we refer the reader to Metz et al. (1988) and Logan (2008).

3

Weak Solutions to Hyperbolic Equations

Chapter 2 emphasized the role of characteristics in hyperbolic problems and the development of algorithms for determining solutions. The underlying assumption was that a continuous, smooth solution exists, at least up to some time when it ceased to be valid. Now we lay the foundation for investigation of discontinuous, or weak, solutions, and the propagation of shock waves.

By our very definition, a solution to a first-order PDE must be smooth, or have continuous first partial derivatives, so that it makes sense to calculate those derivatives and substitute them into the PDE to check if, indeed, we have a solution. Such smooth solutions are called *classical* or *genuine* solutions. Now we want to generalize the notion of a solution and admit discontinuous functions. If a discontinuity in a solution surface exists along some curve in spacetime, there must be some means of checking for a solution along that curve without calculating the partial derivatives, which do not exist along the curve. The aim is to develop such a criterion and formulate the general concept of a weak solution. In the first three sections we derive a jump condition that holds across a discontinuity and show how shocks fit into a solution. Section 3.4 examines two applications, traffic flow and chemical reactors, to gain understanding of characteristics and shock formation in practical settings. In Section 3.5 we introduce the underlying mathematical ideas and state a formal definition of a weak solution. Finally, in Section 3.6 we discuss the asymptotic, or long-time, behavior of solutions. The scenario is as follows. At time $t = 0$ a smooth profile is propagated in time until it evolves into a shock wave, or a discontinuous signal; after this discontinuous wave forms, we ask how the strength of the discontinuity behaves and along what path in spacetime it is

propagated. Understanding this process gives a complete picture of how an initial signal evolves via a first-order PDE.

3.1 Discontinuous Solutions

Characteristics play a fundamental role in understanding how solutions to first-order PDEs propagate. So far our study presupposed that solutions are smooth, or, at the worst, piecewise smooth and continuous. Now we take up the question of discontinuous solutions. As the following example shows, linear equations propagate discontinuous initial or boundary data into the region of interest along characteristics.

Example. Consider the advection equation

$$u_t + cu_x = 0, \quad x \in \mathbb{R}, \quad t > 0, \quad (c > 0)$$

subject to the initial condition $u(x, 0) = u_0(x)$, where u_0 is defined by $u_0 = 1$ if $x < 0$ and $u_0 = 0$ if $x > 0$. Because the solution to the advection equation is $u = u_0(x - ct)$, the initial condition is propagated along the characteristics $x - ct = \text{const}$, and the discontinuity at $x = 0$ is propagated along the line $x = ct$, as shown in Figure 3.1. Thus the solution to the initial value problem is given by

$$u(x, t) = 0 \quad \text{if } x > t; \quad u(x, t) = 1 \quad \text{if } x < t. \quad \square$$

Now examine a simple nonlinear problem with the same initial data. In general, a hyperbolic system with piecewise constant initial data is called a *Riemann problem*.

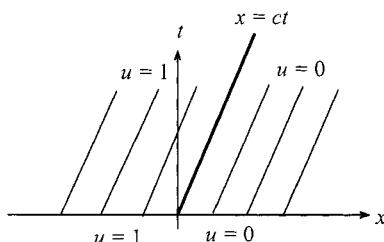


Figure 3.1 Characteristics $x - ct = \text{const}$.

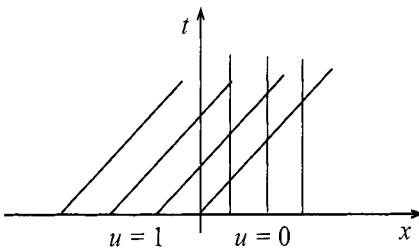


Figure 3.2 Characteristics for the initial value problem (3.1.1)–(3.1.2).

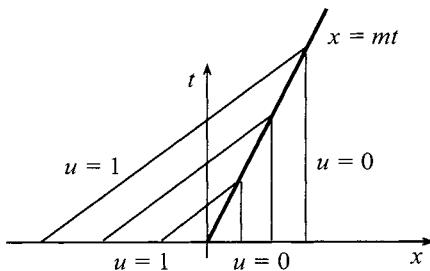


Figure 3.3 Insertion of a line $x = mt$ along which the discontinuity is carried.

Example. Consider the initial value problem

$$u_t + uu_x = 0, \quad x \in \mathbb{R}, \quad t > 0, \quad (3.1.1)$$

$$u(x, 0) = 1 \quad \text{if } x < 0; \quad u(x, 0) = 0 \quad \text{if } x > 0. \quad (3.1.2)$$

We know from Chapter 2 that $du/dt = 0$ along $dx/dt = u$, or $u = \text{const}$ on the straight-line characteristics having speed u . The characteristics emanating from the x axis have speed 0 (vertical) if $x > 0$, and they have speed unity (1) if $x < 0$. The characteristic diagram is shown in Figure 3.2. Immediately, at $t > 0$, the characteristics collide and a contradiction is implied because u must be constant on characteristics. One way to avoid this impasse is to insert a straight line $x = mt$ of nonnegative speed along which the initial discontinuity at $x = 0$ is carried. The characteristic diagram is now changed to Figure 3.3. For $x > mt$ we can take $u = 0$, and for $x < mt$ we can take $u = 1$, thus giving a solution to the PDE (3.1.1) on both sides of the discontinuity. The only question is the choice of m ; for any $m \geq 0$ it appears that a solution can be obtained away from the discontinuity and that solution also satisfies the initial condition. Shall we give up uniqueness for this problem? Is there some other solution that we have not discovered? Or, is there only one valid choice of m ? \square

The answers to these questions lie at the foundation of what a discontinuous solution is. As it turns out, in the same way that discontinuities in derivatives have to propagate along characteristics, discontinuities in the solutions themselves must propagate along special loci in spacetime. These curves are called *shock paths*, and they are not, in general, characteristic curves. What dictates these shock paths is the basic conservation law itself. As we observed earlier, PDEs arise from conservation laws in integral form, and the integral form of these laws holds true even though the functions may not meet the smoothness requirements of a PDE. The integral form of these conservation or balance laws implies a condition, called a *jump condition*, that allows a consistent shock path to be fit into the solution that carries the discontinuity. Indeed, conservation must hold even across a discontinuity. So the answer to the questions posed in the last example is that there is a special value of m (in fact, $m = \frac{1}{2}$) for which conservation holds along the discontinuity.

3.2 Jump Conditions

To obtain a restriction about how a solution across a discontinuity propagates, we consider the integral conservation law

$$\frac{d}{dt} \int_a^b u(x, t) dx = \phi(a, t) - \phi(b, t), \quad (3.2.1)$$

where u is the density and ϕ is the flux. Equation (3.2.1) states that the time rate of change of the total amount of u inside the interval $[a, b]$ must equal the rate that u flows into $[a, b]$ minus the rate that u flows out of $[a, b]$. Under suitable smoothness assumptions (e.g., both u and ϕ continuously differentiable), (3.2.1) implies

$$u_t + \phi_x = 0, \quad (3.2.2)$$

which is the differential form of the conservation law. Recall that ϕ may depend on x and t through dependence on u [i.e., $\phi = \phi(u)$], and (3.3.1) can be written

$$u_t + c(u)u_x = 0, \quad c(u) = \phi'(u). \quad (3.2.3)$$

But if u and ϕ have simple jump discontinuities, we still insist on the validity of the integral form (3.2.1).

Now assume that $x = s(t)$ is a smooth curve in spacetime along which u suffers a simple discontinuity (see Figure 3.4); that is, assume that u is continuously differentiable for $x > s(t)$ and $x < s(t)$, and that u and its derivatives have

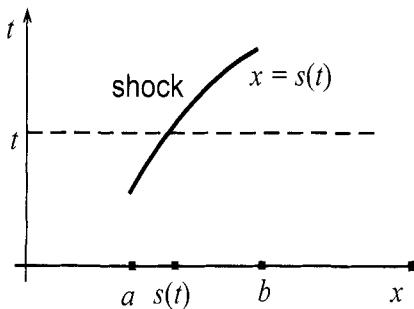


Figure 3.4 Smooth curve in spacetime along which a discontinuity is propagated. Such a curve is called a *shock path*.

finite one-sided limits as $x \rightarrow s(t)^-$ and $x \rightarrow s(t)^+$. Then, choosing $a < s(t)$ and $b > s(t)$, equation (3.2.1) may be written

$$\frac{d}{dt} \int_a^{s(t)} u(x, t) dx + \frac{d}{dt} \int_{s(t)}^b u(x, t) dx = \phi(a, t) - \phi(b, t). \quad (3.2.4)$$

Leibniz' rule for differentiating an integral whose integrand and limits depend on a parameter (here the parameter is time t) can be applied on the left side of (3.2.4), because the integrands are smooth. We therefore obtain

$$\begin{aligned} \int_a^{s(t)} u_t(x, t) dx + \int_{s(t)}^b u_t(x, t) dx + u(s^-, t)s' \\ - u(s^+, t)s' = \phi(a, t) - \phi(b, t), \end{aligned} \quad (3.2.5)$$

where $u(s^-, t)$ and $u(s^+, t)$ are the limits of $u(x, t)$ as $x \rightarrow s(t)^-$ and $x \rightarrow s(t)^+$, respectively, and $s' = ds/dt$ is the speed of the discontinuity $x = s(t)$. In (3.2.5) we now take the limit as $a \rightarrow s(t)^-$ and $b \rightarrow s(t)^+$. The first two terms to go zero because the integrand is bounded and the interval of integration shrinks to zero. Therefore, we obtain

$$-s'[u] + [\phi(u)] = 0, \quad (3.2.6)$$

where the brackets denote the jump of the quantity inside across the discontinuity (the value on the left minus the value on the right). Equation (3.2.6) is called the *jump condition*. (In fluid mechanical problems, conditions across a discontinuity are known as Rankine–Hugoniot conditions.) It relates conditions both ahead of and behind the discontinuity to the speed of the discontinuity itself. In this context, the discontinuity in u that propagates along the curve $x = s(t)$ is called a *shock wave*, and the curve $x = s(t)$ is called the *shock path*,

or just the *shock*; s' is called the *shock speed*, and the magnitude of the jump in u is called the *shock strength*. The form of (3.2.6) gives the correspondence

$$(\dots)_t \leftrightarrow -s'[(\dots)], \quad (\dots)_x \leftrightarrow [(\dots)], \quad (3.2.7)$$

between a given PDE (3.2.2) and its associated shock condition (3.2.6); equation (3.2.7) makes it easy to remember the jump conditions associated with a PDE in conservation form.

Example. Consider

$$u_t + uu_x = 0.$$

In conservation form this PDE can be written

$$u_t + \left(\frac{u^2}{2} \right)_x = 0,$$

where the flux is $\phi = u^2/2$. According to (3.2.6), the jump condition is given by

$$-s'[u] + \left[\frac{u^2}{2} \right] = 0,$$

or

$$s' = \frac{u_+ + u_-}{2}.$$

So the speed of the shock is the average of the u values ahead of and behind the shock. Returning to (3.1.1)–(3.1.2), where the initial condition is given by the discontinuous data $u = 1$ for $x < 0$ and $u = 0$ for $x > 0$, a shock can be fit with speed

$$s' = \frac{0 + 1}{2} = \frac{1}{2}.$$

Therefore, a solution consistent with the jump condition to the initial value problem (3.1.1)–(3.1.2) is

$$u(x, t) = 1 \quad \text{if } x < \frac{t}{2}; \quad u(x, t) = 0 \quad \text{if } x > \frac{t}{2}. \quad \square$$

3.2.1 Rarefaction Waves

Another difficulty can occur with nonlinear equations having discontinuous initial or boundary data.

Example. Consider the equation

$$u_t + uu_x = 0, \quad x \in \mathbb{R}, \quad t > 0,$$

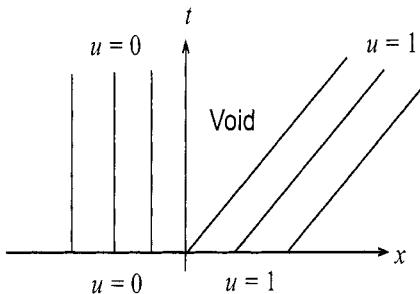


Figure 3.5 Spacetime diagram showing a characteristic void.

subject to the initial condition

$$u(x, 0) = 0 \quad \text{if } x < 0; \quad u(x, 0) = 1 \quad \text{if } x > 0.$$

The characteristic diagram is plotted in Figure 3.5. Because u is constant along characteristics, the data $u = 1$ are carried into the region $x > t$ along characteristics with speed 1, and the data $u = 0$ are carried into the region $x < t$ along vertical (speed 0) characteristics. There is a region $0 < x < t$ void of the characteristics. In this case there is a continuous solution that connects the solution $u = 1$ ahead to the solution $u = 0$ behind. We simply insert characteristics (straight lines in this case) passing through the origin into the void in such a way that u is constant on the characteristics and u varies continuously from 1 to 0 along these characteristics (see Figure 3.6). In other words, along the characteristic $x = ct$, $0 < c < 1$, take $u = c$. Consequently, the solution to the Riemann problem is

$$\begin{aligned} u(x, t) &= 0 && \text{if } x < 0, \\ u(x, t) &= \frac{x}{t} && \text{if } 0 < \frac{x}{t} < 1, \quad u(x, t) = 1 && \text{if } x > t. \end{aligned}$$

A solution of this type is called a *centered expansion wave*, or a *fan*; other terms are *release wave* or *rarefaction wave*. The idea is that the wave spreads as time increases; a wave profile is shown in Figure 3.7. \square

3.2.2 Shock Propagation

Here are some remarks regarding formation of shock waves. In air, for example, the speed that finite signals or waves propagate is proportional to the local density of air. (We are not referring to small-amplitude signals, as in acoustics, that propagate at a constant sound speed.) Thus, at local points where the

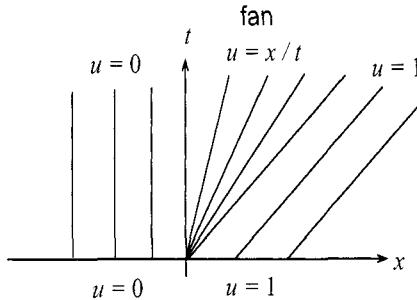


Figure 3.6 Insertion of a characteristic fan in the void in Figure 3.5.

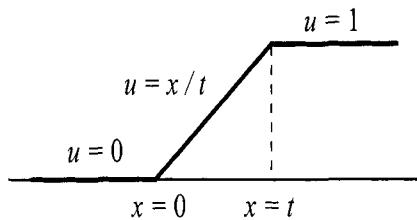


Figure 3.7 Graph of a wave profile at time t corresponding to the characteristic diagram in Figure 3.6.

density is higher, the signal (or the value of the density at that point) propagates faster. Therefore, a density wave propagating in air will gradually distort and steepen until it propagates as a discontinuous disturbance, or shock wave. Figure 3.8 depicts various time snapshots of such a density wave. The wave steepens in time because of the tendency of the medium to propagate signals faster at higher density; thus point A moves to the right faster than point B , and finally, a shock forms. This same mechanism causes rarefaction waves to form, which release the density and spread out the wave on the backside. This dependence of propagation speed on amplitude is reflected mathematically in the conservation law (3.2.3) by noting that the characteristic speed $c = c(u)$ depends on u ; this phenomenon is typically nonlinear.

Therefore, discontinuities do not need to be present initially; shocks can form from the distortion of a perfectly smooth solution. The time when the shock forms, usually signaled by an infinite spatial derivative or gradient catastrophe, is called the *breaking time*. Determining conditions under which solutions blow up in this manner is one of the important problems in nonlinear hyperbolic problems. The problem of fitting in a shock (i.e., determining the spacetime location of the shock) after the blowup occurs is a difficult problem,

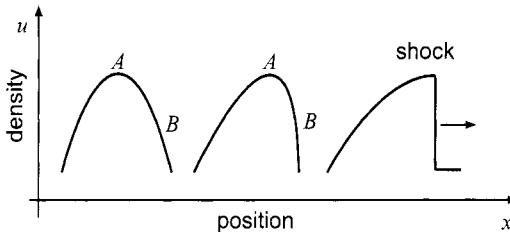


Figure 3.8 Schematic showing how a wave profile steepens into a shock wave. Points where the density is higher are propagated at a higher speed.

both analytically and numerically. We emphasize again that the distortion of wave profile and the formation of a shock is a distinctly nonlinear phenomenon.

We end this section with an example of shock fitting.

Example. (Shock Fitting) Consider the initial value problem

$$u_t + uu_x = 0, \quad x \in \mathbb{R}, \quad t > 0,$$

$$u(x, 0) = \begin{cases} 1, & x < 0, \\ -1, & 0 < x < 1, \\ 0, & x > 1. \end{cases}$$

Here $c(u) = u$ and the characteristics emanate with speed u from the x axis as shown in the characteristic diagram in Figure 3.9. The flux is $\phi(u) = u^2/2$, and the jump condition is

$$s' = \frac{u_1 + u_2}{2},$$

where u_1 is the value of u ahead of the shock and u_2 is the value of u behind the shock. Clearly, a shock must form at $t = 0$ with speed $s' = (-1 + 1)/2 = 0$ and propagate until time $t = 1$, as shown in Figure 3.10. To continue the shock beyond $t = 1$, we must know the solution ahead of the shock. Therefore, we introduce an expansion wave in the void in Figure 3.10; that is, in this region we take

$$u = \frac{x-1}{t}.$$

The straight-line characteristics issuing from $x = 1, t = 0$ have equation $x = -kt + 1$ or $(x-1)/t = k = \text{constant}$. This expansion fan takes u from the value -1 to the value 0 ahead of the wave. The shock beyond $t = 1$, according to the jump condition, has speed

$$s' = \frac{(x-1)/t + 1}{2}. \quad (3.2.8)$$

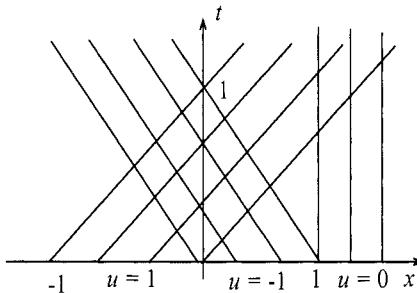


Figure 3.9 Characteristic diagram with intersecting characteristics. Figures 3.10–3.12 show how a shock is fit into the diagram.

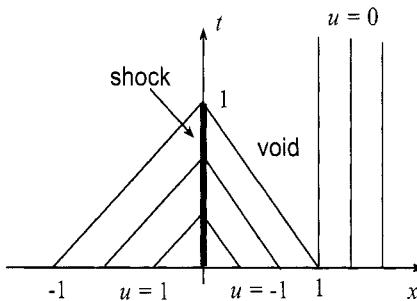


Figure 3.10 Insertion of a shock for $0 < t < 1$.

To find the shock path we note that $s' = dx/dt$, so equation (3.2.8) is a first-order ordinary differential equation for the shock path. This ODE can be expressed in the form

$$\frac{dx}{dt} - \frac{x}{2t} = \frac{1 - 1/t}{2},$$

which is a first-order linear equation. The initial condition is $x = 0$ at $t = 1$, where the shock starts. The solution is

$$x = s(t) = t + 1 - 2\sqrt{t}, \quad 1 \leq t \leq 4, \quad (3.2.9)$$

and a plot is shown in Figure 3.11. The shock in (3.2.9) will propagate until $x = 1$ at $t = 4$. At this instant the shock runs into the vertical characteristics emanating from $x > 1$, and the jump condition (3.2.8) is no longer valid. The new jump condition for $t > 4$ is

$$s' = \frac{0 + 1}{2} = \frac{1}{2}.$$

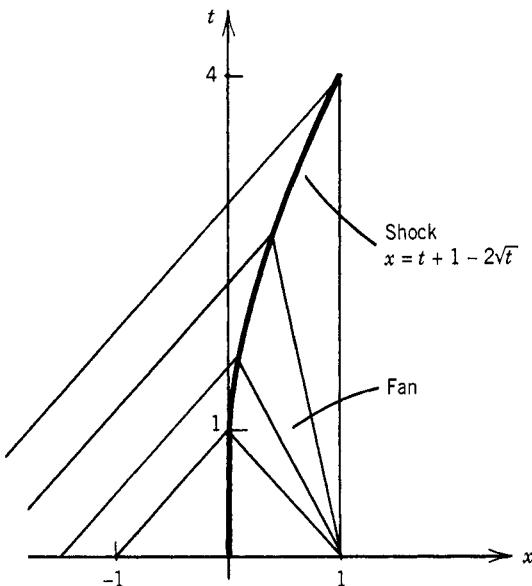


Figure 3.11 Insertion of a fan in the void region and the resulting shock for $1 < t < 4$.

Therefore, the shock is a straight line with speed $\frac{1}{2}$ for $t > 4$, and its equation is

$$x - 1 = \frac{t - 4}{2}, \quad t > 4.$$

The complete solution is indicated on the wave diagram in Figure 3.12. \square

In summary, we think of shock propagation in the following way. The speed of the shock is the average value of the solution ahead and the solution behind; the characteristics ahead and behind carry information to the shock, dictating its path. A similar interpretation may be advanced, in general, for more complicated problems such as aerodynamic flows. The location of the shock path is determined by the flow ahead and the flow behind. Information about the flow is carried by the characteristics to the shock front, fixing its position in spacetime according to the jump conditions, or conservation laws, that hold at the shock front.

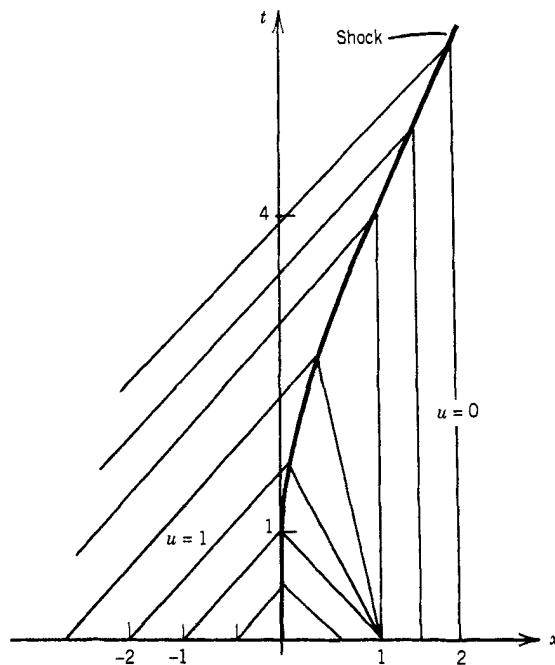


Figure 3.12 Continuation of the shock for $t > 4$.

EXERCISES

1. Consider the Riemann problem

$$u_t + c(u)u_x = 0, \quad x \in \mathbb{R}, \quad t > 0,$$

where $c(u) = \phi'(u) > 0$, and $c'(u) > 0$, with initial condition

- (a) $u(x, 0) = u_0$, $x > 0$; $u(x, 0) = u_1$, $x < 0$; u_0 and u_1 positive constants with $u_1 > u_0$ or
- (b) $u(x, 0) = u_0$, $x > 0$; $u(x, 0) = u_1$, $x < 0$; u_0 and u_1 positive constants with $u_1 < u_0$.

In each case draw a representative characteristic diagram showing the shock path, and find a formula for the solution.

2. How do the results of Exercise 1 change if $c'(u) < 0$?
 3. Consider the equation

$$u_t + c(u)u_x = 0,$$

with c twice continuously differentiable and $c(u) = \phi'(u)$, where ϕ is the flux. Given

$$\frac{\phi(u_2) - \phi(u_1)}{u_2 - u_1} = \frac{c(u_1) + c(u_2)}{2}$$

for all u_1 and u_2 , show that $\phi(u)$ is a quadratic function of u .

4. Consider the initial value problem

$$u_t + uu_x = 0, \quad x \in \mathbb{R}, \quad t > 0,$$

$$u(x, 0) = \begin{cases} 1, & x < 0, \\ 1-x, & 0 < x < 1, \\ 0, & x > 1. \end{cases}$$

Find a continuous solution for $t < 1$. For $t > 1$ fit a shock and find the form of the solution. Finally, sketch time snapshots of the wave for $t = 0$, $t = \frac{1}{2}$, $t = 1$, and $t = \frac{3}{2}$.

3.3 Shock Formation

In Section 3.1 we introduced the Riemann problem, where the initial data were discontinuous. Now we consider the problem of shock formation from smooth data, and, in particular, the question of when the blowup occurs. We focus on the initial value problem

$$u_t + c(u)u_x = 0, \quad x \in \mathbb{R}, \quad t > 0, \quad (3.3.1)$$

$$u(x, 0) = u_0(x), \quad x \in \mathbb{R}, \quad (3.3.2)$$

where $c(u) > 0$, $c'(u) > 0$, and $u_0 \in C^1$. We showed in Chapter 2 that if u_0 is a nondecreasing function on \mathbb{R} , then a smooth solution $u = u(x, t)$ exists for all $t > 0$ and is given implicitly by the formulas

$$u(x, t) = u_0(\xi), \quad \text{where } x - \xi = c(u_0(\xi))t. \quad (3.3.3)$$

Therefore, for singularities to occur in (3.3.1)–(3.3.2), necessarily u_0 must be strictly decreasing in some open interval.

Consider the case when $u_0(x) > 0$ and $u'_0(x) < 0$ on \mathbb{R} . The characteristic equations are

$$u = \text{const on } \frac{dx}{dt} = c(u).$$

Therefore, the characteristics, which are straight lines, issuing from two points ξ_1 and ξ_2 on the x axis with $\xi_1 < \xi_2$, have speeds $c(u_0(\xi_1))$ and $c(u_0(\xi_2))$, respectively. Because u_0 is decreasing and c is increasing, it follows that

$$c(u_0(\xi_1)) > c(u_0(\xi_2)).$$

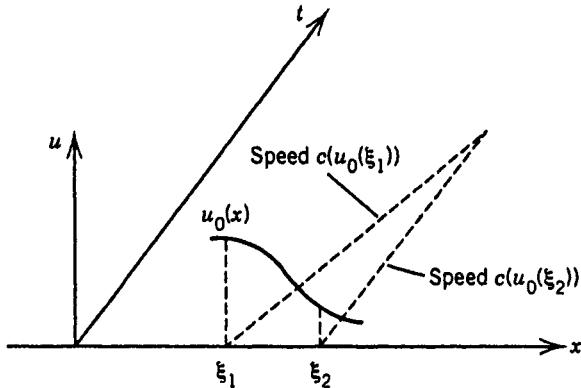


Figure 3.13

In other words, the characteristic emanating from ξ_1 is faster than the one emanating from ξ_2 (see Figure 3.13). Therefore, the characteristics cross and a contradiction results because u is a different constant on each characteristic. So a smooth solution cannot exist for all $t > 0$.

To determine the breaking time, or the time when a gradient catastrophe occurs, we calculate u_x along a characteristic, which has equation

$$x - \xi = c(u_0(\xi))t. \quad (3.3.4)$$

Let $g(t) = u_x(x(t), t)$ denote the gradient of u along the characteristic $x = x(t)$ given by (3.3.4). Then

$$\frac{dg}{dt} = u_{tx} + c(u)u_{xx}.$$

But differentiating the PDE (3.3.1) with respect to x gives

$$u_{tx} + c(u)u_{xx} + c'(u)u_x^2 = 0.$$

Comparing the last two equations gives

$$\frac{dg}{dt} = -c(u)g^2 \quad (3.3.5)$$

along the characteristic. Equation (3.3.5) can be solved easily to obtain

$$g = \frac{g(0)}{1 + g(0)c'(u_0(\xi))t}, \quad (3.3.6)$$

where $g(0)$ is the initial gradient at $t = 0$. Translating (3.3.6) into alternate notation gives

$$u_x = \frac{u'_0(\xi)}{1 + u'_0(\xi)c'(u_0(\xi))t}, \quad (3.3.7)$$

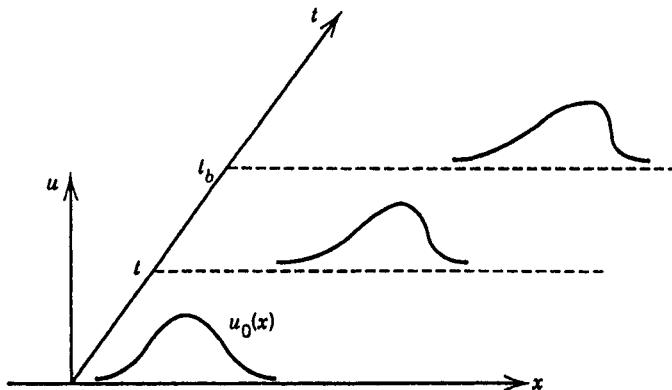


Figure 3.14 Diagram showing how a wave evolves into a shock wave. The breaking time t_b is the time when the gradient catastrophe occurs.

which is a formula for the gradient u_x along the characteristic (3.3.4). Because of the assumptions (u_0 nonincreasing and c increasing), it follows that u'_0 and c' in the denominator of (3.3.7) have opposite signs, and therefore u_x will blow up at a finite time along the characteristic (3.3.4). Consequently, if we examine u_x along *all* the characteristics, the breaking time for the wave will be on the characteristic, parameterized by ξ , where the denominator in (3.3.7) first vanishes. Denoting

$$F(\xi) = c(u_0(\xi)), \quad F'(\xi) = u'_0(\xi)c'(u_0(\xi)),$$

we conclude that the wave first breaks along the characteristic $\xi = \xi_b$ for which $|F'(\xi)|$ is maximum. Then the time of the first breaking is

$$t_b = -\frac{1}{F'(\xi_b)}.$$

The positive time t_b is called the *breaking time* of the wave. Figure 3.14 illustrates several time snapshots of a wave that breaks at time t_b .

If the initial data u_0 are not monotone, breaking will first occur on the characteristic $\xi = \xi_b$, for which $F'(\xi) < 0$ and $|F'(\xi)|$ is maximum.

Example. Consider the initial value problem

$$\begin{aligned} u_t + uu_x &= 0, \quad x \in \mathbb{R}, \quad t > 0, \\ u(x, 0) &= e^{-x^2}, \quad x \in \mathbb{R}. \end{aligned}$$

The initial signal is a bell-shaped curve, and characteristics emanate from the x axis with speed $c(u) = u$. A characteristic diagram is shown in Figure 3.15.

It is clear that the characteristics collide and a shock will form on the front portion of the wave along the characteristics emanating from the positive x axis. To determine the breaking time of the wave, we compute

$$F(\xi) = c(u_0(\xi)) = e^{-\xi^2}.$$

Then

$$F'(\xi) = -2\xi e^{-\xi^2} \text{ and } F''(\xi) = -(4\xi^2 - 2)e^{-\xi^2}.$$

Thus $F''(\xi) = 0$ when $\xi = \xi_b = \sqrt{0.5}$. This is where F' is a maximum. Therefore, the breaking time is

$$t_b = -\frac{1}{F'(\xi_b)} \cong 1.16.$$

So breaking first occurs along the characteristic $\xi = \xi_b$ at time t_b . The solution of the initial value problem up to time t_b is given implicitly by

$$u(x, t) = e^{-\xi^2},$$

where $\xi = \xi(x, t)$ is the solution of $x - \xi = e^{-\xi^2}t$. \square

There is an alternate way to look at the problem of determining the breaking time of a wave for the initial value problem (3.3.1)–(3.3.2). The solution is given

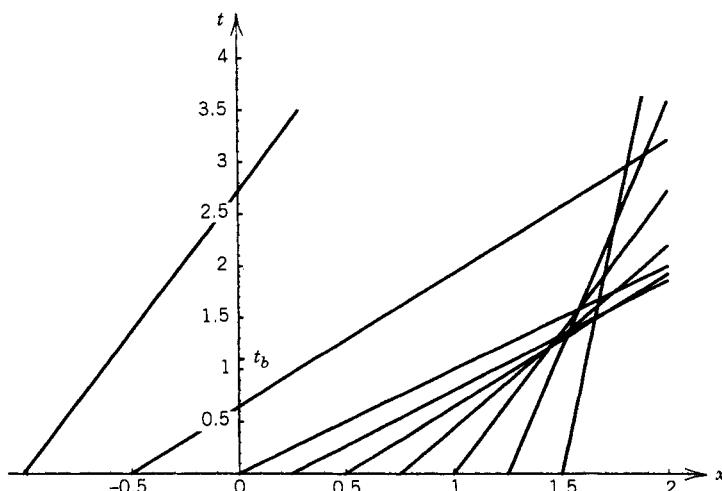


Figure 3.15 Characteristic diagram. The breaking time is approximately $t_b = 1.16$, and breaking first occurs on the characteristic $\xi_b = \sqrt{0.5}$.

implicitly by the formulas (3.3.3). To determine the solution explicitly, we must be able to solve the equation

$$x - \xi = c(u_0(\xi))t \quad (3.3.8)$$

for ξ and substitute the result into the equation $u = u_0(\xi)$ to find $u = u(x, t)$ (i.e., u as a function of x and t). We may ask when it is possible to solve (3.3.8) for ξ . This is a standard problem involving the implicit function theorem. If we write (3.3.8) as

$$J(x, t, \xi) = x - \xi - c(u_0(\xi))t = 0, \quad (3.3.9)$$

then the condition that allows us to locally solve (3.3.9) for ξ is that the partial derivative of J with respect to ξ must be nonzero. Thus we must have

$$-1 - c'(u_0(\xi))u'_0(\xi)t \neq 0. \quad (3.3.10)$$

This is the same condition that the denominator of (3.3.7) be nonzero (i.e., u_x is not infinite). Consequently, if condition (3.3.10) holds, then for each (x, t) we can uniquely determine the characteristic, designated by ξ , that leads backward in time from (x, t) to the point $(\xi, 0)$ on the initial timeline (x axis).

EXERCISES

1. Consider the initial value problem

$$u_t + uu_x = 0, \quad x \in \mathbb{R}, \quad t > 0,$$

$$u(x, 0) = \frac{1}{1+x^2}, \quad x \in \mathbb{R}.$$

Find the breaking time t_b of the wave and write the solution for $0 < t < t_b$ in implicit form. Sketch the characteristic diagram.

2. Discuss the existence of a global solution ($x \in \mathbb{R}$, $t > 0$) for the initial value problem

$$u_t = -\frac{t}{u}, \quad x \in \mathbb{R}, \quad t > 0,$$

$$u(x, 0) = u_0(x), \quad x \in \mathbb{R}.$$

Sketch the characteristic diagram.

3. Consider the conservation law

$$u_t + \phi(u)_x = 0, \quad x \in \mathbb{R}, \quad t > 0,$$

$$u(x, 0) = u_0(x) > 0, \quad x \in \mathbb{R},$$

where ϕ' , $\phi'' > 0$, and where $u'_0 > 0$. Show there exists a constant E for which $u_x < E/t$ for all $t > 0$.

4. Consider the initial-boundary value problem

$$\begin{aligned} u_t + uu_x &= 0, \quad x > 0, \quad t > 0, \\ u(x, 0) &= 1, \quad x > 0; \quad u(0, t) = t + 1, \quad t > 0. \end{aligned}$$

Sketch a characteristic diagram and find the breaking time t_b . Determine the solution up to $t = t_b$. After the shock forms, calculate its path in spacetime. *Solution:* $s(t) = (t+3)(3t+1)/16$.

5. Consider the initial value problem

$$\begin{aligned} u_t + uu_x &= -ku^2, \quad x \in \mathbb{R}, \quad t > 0 \quad (k > 0), \\ u(x, 0) &= u_0(x) > 0, \quad x \in \mathbb{R}. \end{aligned}$$

- (a) Find u implicitly.
- (b) Determine conditions on k and u_0 such that a shock will *not* form.
- (c) If a shock forms, what is t_b ?

6. Consider the PDE

$$u_t + uu_x = B(x - vt),$$

where v is a positive constant and the function B satisfies the conditions $B(y) = 0$ if $|y| \geq y_0$ and $B(y) > 0$ otherwise. (This PDE is a conservation law with a moving source term.)

- (a) Show there exists a traveling wave solution of the form $u = u(x - vt)$ satisfying $u(+\infty) = u_0 < v$ if, and only if

$$(u_0 - v)^2 > 2 \int_R B(y) dy.$$

- (b) Let $u(x, 0) = u_0$, $x \in \mathbb{R}$. By transforming to a moving coordinate system (moving with speed v), show that

$$\frac{du}{d\tau} = B(z) \quad \text{on} \quad \frac{dz}{d\tau} = u - v,$$

with $u = u_0$ and $z = x$ at $\tau = 0$.

- (c) In part (b) take $B(y) = B = \text{constant}$ for $|y| \leq a$ and $B(y) = 0$ otherwise. Draw a characteristic diagram in $z\tau$ space in both the case $v < u_0$ and the case $v > u_0$. When do shocks form? Discuss the case $v = u_0$.

7. Consider the initial value problem

$$u_t + uu_x = 0, \quad x \in \mathbb{R}, \quad t > 0,$$

$$u(x, 0) = 2 - x^2 \quad \text{for } x < 1; \quad u(x, 0) = 1 \quad \text{for } x \geq 1.$$

Sketch the characteristic diagram and find the breaking time t_b . Write the solution for $t < t_b$.

8. Consider the initial value problem

$$u_t + uu_x = 0, \quad x \in \mathbb{R}, \quad t > 0,$$

$$u(x, 0) = 1 \quad \text{for } x > 1; \quad u(x, 0) = |x| - 1 \quad \text{for } x < 1.$$

Find a solution containing a shock and determine the shock path.

9. Consider the initial value problem

$$u_t + uu_x = 0, \quad x \in \mathbb{R}, \quad t > 0,$$

$$u(x, 0) = 2 \quad \text{if } x < 0, \quad u(x, 0) = 1 \quad \text{if } 0 < x < 2, \quad u(x, 0) = 0 \quad \text{if } x > 2.$$

Find a solution containing a shock and determine the shock path. (In this case two shocks merge into a single shock.)

10. Consider the initial value problem

$$u_t + uu_x = 0, \quad x \in \mathbb{R}, \quad t > 0,$$

$$u(x, 0) = \cos x, \quad 0 \leq x \leq \frac{\pi}{2},$$

$$u(x, 0) = 0, \quad x > \frac{\pi}{2},$$

$$u(x, 0) = 1, \quad x < 0.$$

Show that u_x blows up at $t_b = 1$ and at location $x_b = \pi/2$. Write the solution for $0 < t < t_b$ in implicit form and sketch the characteristic diagram.

3.4 Applications

There are many applications of first-order PDEs to problems in science and engineering. Rather than present an exhaustive list, we focus on two applications, one involving traffic flow and one involving a plug flow chemical reactor. In Chapter 4, where we consider systems of first-order equations, there is more fertile ground for applications.

3.4.1 Traffic Flow

What was once a novel application of first-order PDEs—namely, the application to traffic flow—has become commonplace in treatments of first order PDEs. Nevertheless, the application is interesting and provides the novice with a simple, concrete example where intuition and mathematics can come together in a familiar setting.

Imagine a single-lane roadway occupied by cars whose density $u = u(x, t)$ is given in vehicles per unit length, and the flux of vehicles $\phi(x, t)$ is given in vehicles per unit time. The length and time units are usually miles and hours. We are making a continuum assumption here by regarding u and ϕ as continuous functions of the distance x . (A more satisfying treatment might be to regard the vehicles as discrete entities and write down finite difference equations to model the flow; this approach has been developed, but we shall not follow it here.) The basic conservation law requires that the time rate of change of the total number of vehicles in any interval $[a, b]$ equals the inbound flow rate at $x = a$ minus the outbound flow rate at $x = b$, or, in integral form,

$$\frac{d}{dt} \int_a^b u(x, t) dx = \phi(a, t) - \phi(b, t). \quad (3.4.1)$$

If u and ϕ are smooth functions, then in the standard way we can obtain a PDE relating u and ϕ that models the flow of the cars. This equation is the conservation law

$$u_t + \phi_x = 0. \quad (3.4.2)$$

Now we need to assume a constitutive relation for the flux ϕ . The flux represents the rate that vehicles go by a given point; it seems desirable for the flux to depend on the traffic density u , or $\phi = \phi(u)$. If $u = 0$, the flux should be zero; if u is large, the flux should also be zero because the traffic is jammed. Therefore, we assume that there is a positive value of u , say, u_j , such that $\phi(u_j) = 0$. Otherwise, we assume that ϕ is positive and concave down on the interval $(0, u_j)$, and ϕ has a unique maximum at u_{\max} . Thus $u = u_{\max}$ is the value of the density where the greatest number of vehicles go by. In summary, the assumptions on the flux are

$$\begin{aligned} \phi(u) &> 0 \quad \text{and} \quad \phi''(u) < 0 \text{ on } (0, u_j), \\ \phi(0) &= \phi(u_j) = 0, \quad \phi'(u) > 0 \text{ on } [0, u_{\max}), \\ \phi'(u) &< 0 \text{ on } (u_{\max}, u_j], \quad \phi'(u_{\max}) = 0. \end{aligned} \quad (3.4.3)$$

See Figure 3.16. As before, we may write the PDE (3.4.2) as

$$u_t + c(u)u_x = 0, \quad \text{where} \quad c(u) = \phi'(u), \quad c'(u) < 0. \quad (3.4.4)$$

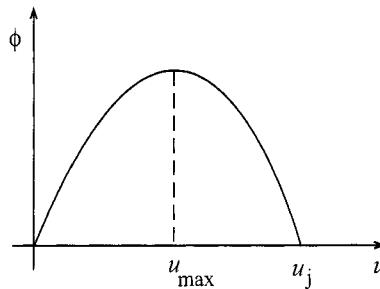


Figure 3.16 Experimentally measured flux ϕ as a function of the vehicle density u .

Equation (3.4.4) has characteristic form

$$u = \text{const} \quad \text{on} \quad \frac{du}{dt} = c(u). \quad (3.4.5)$$

Therefore, signals that travel at characteristic speed $c(u)$ propagate forward into the traffic if $u < u_{\max}$, and signals propagate backward if $u > u_{\max}$. Because $c'(u) < 0$, waves will show a shocking-up effect on their backside. For example, if there is a density wave in the form of a bell-shaped curve, with the density everywhere less than u_{\max} , the wave will propagate as shown in Figures 3.17 and 3.18, which show the characteristic diagram and time snapshots of the wave, respectively.

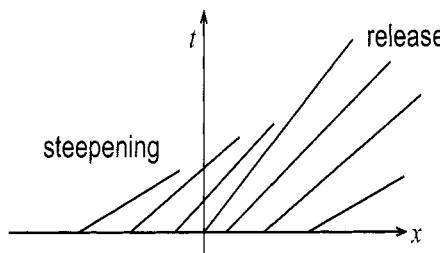


Figure 3.17 Characteristic diagram for an initial density in the form of a bell-shaped curve.

To see how a driver experiences and responds to such a density wave, we define the *velocity* $V = V(u)$ of the traffic flow defined by the equation

$$\phi(u) = uV(u). \quad (3.4.6)$$

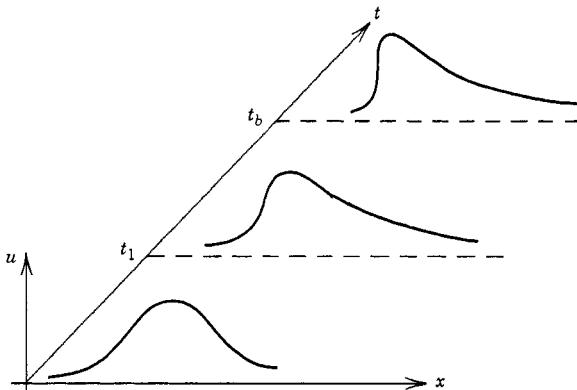


Figure 3.18 Diagram showing how the density wave profile “shocks up” on the backside when $c'(u) < 0$.

Like ϕ and u , V is a point function of (x, t) , giving a continuum of values at each location x in the flow. With this interpretation, V is the actual velocity of the vehicle at (x, t) . Equation (3.4.6) states that density times velocity equals flux, which is a correct statement. We now place one additional restriction on ϕ that will guarantee V is a decreasing function of u ; in particular, we assume that $u\phi'(u) < \phi(u)$ (see Exercise 1). Using (3.4.6) we find

$$c(u) = \phi'(u) = uV'(u) + V(u).$$

Then, because $V' < 0$, it follows that $c(u) < V(u)$. Consequently, the traffic moves at a speed greater than the speed signals are propagated in the flow. For example, a driver approaching a density wave like the one shown in Figure 3.18 will decelerate rapidly through the steep rear of the wave and then accelerate slowly through the rarefaction on the front side.

Typical experimental numbers for values of the various constants are quoted in Whitham (1974). For example, on a highway, $u_j = 225$ vehicles per mile, $u_{\max} = 80$ vehicles per mile, and the maximum flux is 1500 vehicles per hour. Interestingly enough, the velocity that gives this maximum flux is a relatively slow 20 miles per hour. In another case study, in the Lincoln Tunnel linking New York and New Jersey, a flux given by $\phi(u) = au \ln(u_j/u)$, with $a = 17.2$ and $u_j = 228$, was found to fit the experimental data.

Another common phenomenon is a stream of traffic of constant density $u_0 < u_{\max}$ suddenly encountering a traffic light that turns red at time $t = 0$. Figure 3.19 shows the resulting shock wave that is propagated back into the flow. Ahead of the shock, the cars are jammed at density $u = u_j$ and velocity is zero; behind the shock the cars have density u_0 and corresponding velocity

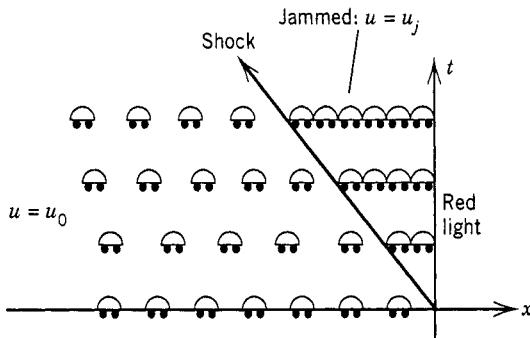


Figure 3.19 Spacetime diagram showing the effect of a red light stopping a uniform stream of traffic. A shock is propagated back into the traffic. This is a shock that we experience regularly when we drive a vehicle.

$V(u_0) > c(u_0) > 0$. Figure 3.20 illustrates the characteristic diagram and a typical vehicle trajectory. The speed of the shock is given by the jump condition

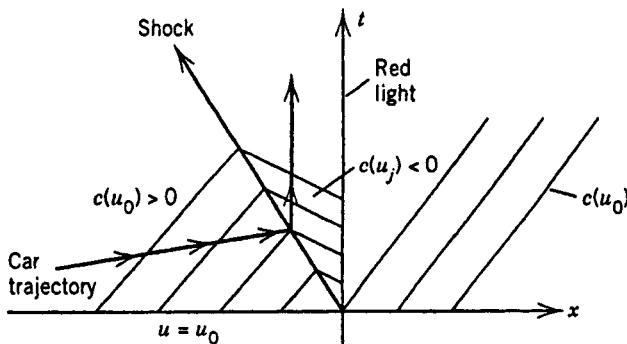


Figure 3.20 Characteristic diagram associated with a red light stopping a uniform stream of traffic.

$$s' = \frac{\phi_{\text{ahead}} - \phi_{\text{behind}}}{u_{\text{ahead}} - u_{\text{behind}}} = \frac{0 - \phi(u_0)}{u_j - u_0} = -\frac{\phi(u_0)}{u_j - u_0} < 0.$$

If the traffic light turns green at some time $t = t_0$, a centered rarefaction wave is created that releases the density as the stopped vehicles move forward. The lower value of the density created by the rarefaction ahead of the shock wave will force the shock to change direction and accelerate through the intersection. A characteristic diagram of this situation is shown in Figure 3.21. Observe that

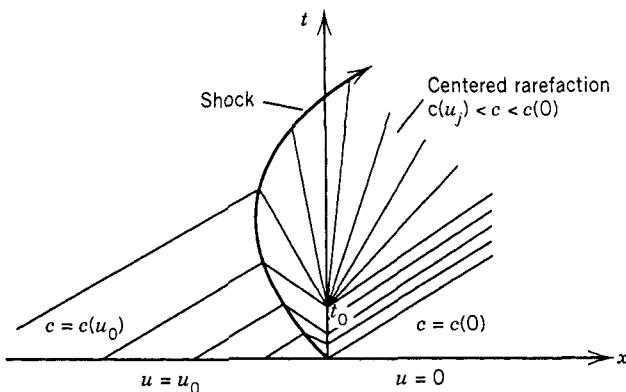


Figure 3.21 When the light turns green a rarefaction wave releases the jammed traffic and the shock turns to go through the intersection itself. At a busy intersection, the shock does not get through the intersection before the light turns red again.

the characteristics are straight lines, and the fan must take characteristics of speed $c(u_j)$ behind the rarefaction to characteristics of speed $c(0)$ ahead of the rarefaction. The density u must change from the jammed value u_j behind to the fan to the value zero ahead (because there are no cars ahead). The characteristic speed $c(u)$ in the rarefaction is therefore given by the equation

$$c(u) = \frac{x}{t - t_0}.$$

Note that u is constant on straight lines, so $c(u)$ is constant on straight lines; the straight lines in the fan emanating from $(0, t_0)$ are given by $x/(t - t_0) = \text{const.}$ The curved shock path can be computed by the jump condition

$$s' = \frac{dx}{dt} = \frac{\phi(u) - \phi(u_0)}{u - u_0}.$$

The Exercises provide further practice in traffic flow problems.

3.4.2 Plug Flow Chemical Reactors

In industry, many types of chemical reactors produce products from given reactants. One of these is called a *plug flow reactor* (also called a *piston flow*, *slug flow*, or *ideal tubular reactor*). A model of this reactor is a long tube where a reactant enters at a fixed rate and a chemical reaction takes place as the

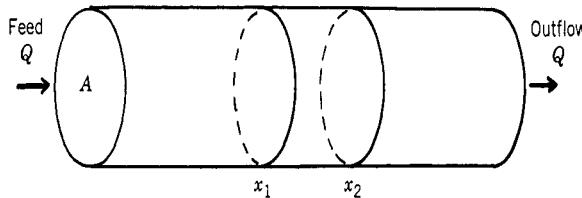


Figure 3.22 Plug flow reactor with material entering and leaving at the same volumetric flow rate Q .

material moves through the tube, producing the final product as it leaves the other end. In such a reactor the flow is assumed to be one-dimensional—that is, all state variables are assumed to be constant in any cross section; the variation takes place only in the axial direction. A model of a plug flow reactor of cross-sectional area A and length L is shown in Figure 3.22. A reactant **A** is fed in the left end at $x = 0$ at the constant volumetric rate Q (volume/second). On entry, **A** reacts chemically and produces a product **B**; that is, $\mathbf{A} \rightarrow \mathbf{B}$. The rate of production, or the chemical reaction rate, is given by r , measured in mass/volume/time. The material leaves the reactor at $x = L$ at the same volumetric flow rate Q . If $a = a(x, t)$ denotes the concentration of the chemical species **A**, given in mass per unit volume, then the flux is then given by Qa . The model equation governing the evolution of the concentration $a(x, t)$ can be expressed from a basic conservation law, which states that the time rate of change of the total amount of **A** in the section $[x_1, x_2]$ must balance the net flux of **A** through x_1 and x_2 , plus the rate at which **A** is consumed by the chemical reaction in $[x_1, x_2]$. This is expressed symbolically as

$$\frac{d}{dt} \int_{x_1}^{x_2} Aa(x, t) dx = Qa(x_1, t) - Qa(x_2, t) + \int_{x_1}^{x_2} Ar dx. \quad (3.4.7)$$

Assuming that a is sufficiently differentiable and using the arbitrariness of the interval $[x_1, x_2]$, (3.4.7) reduces to the PDE

$$a_t + va_x = r, \quad v = \frac{Q}{A}. \quad (3.4.8)$$

Equation (3.4.8) is the advection equation with a source term. We next make some assumptions regarding the reaction rate r . By the law of mass action, the reaction rate depends on the concentration a as well as the temperature at which the reaction occurs. For first-order kinetics, $r = -ka$, where k is a temperature-dependent rate factor (called the *rate constant*, which is a slight misnomer). Rather than introduce another variable (temperature), however, we assume that k is a constant; so, we consider the problem

$$a_t + va_x = -ka. \quad (3.4.9)$$

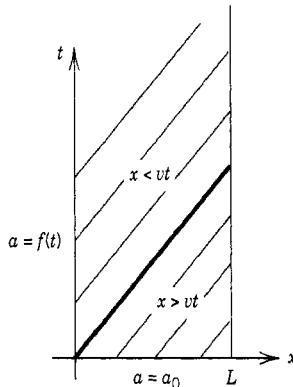


Figure 3.23 Characteristic diagram for the initial-boundary value problem (3.4.9)–(3.4.11).

To equation (3.4.9) we append the initial condition

$$a(x, 0) = a_0, \quad 0 < x < L, \quad (3.4.10)$$

and the boundary condition

$$a(0, t) = f(t), \quad t > 0. \quad (3.4.11)$$

The initial concentration in the reactor is a_0 , a constant, and the concentration of the feed at $x = 0$ is $f(t)$; therefore, the flux at $x = 0$ is $Qf(t)$. We can solve this problem by characteristics. The characteristic form of (3.4.9) is

$$\frac{da}{dt} = -ka \quad \text{on } x = vt + c_1, \quad (3.4.12)$$

where c_1 is a constant. Hence

$$a = c_2 e^{-kt} \quad \text{on } x = vt + c_1, \quad (3.4.13)$$

where c_2 is another constant. We note (see Figure 3.23) that the initial and boundary data are carried along straight-line characteristics at speed v to the boundary at $x = L$, which is the right end of the reactor. Therefore, we divide the problem into two regions, $x > vt$ and $x < vt$. For $x > vt$ we parameterize the initial data by $a = a_0, x = \xi$, at $t = 0$. Then, from (3.4.13), the constants are given by

$$c_1 = a_0, \quad c_2 = \xi. \quad (3.4.14)$$

Thus

$$a(x, t) = a_0 e^{-kt}, \quad x > vt. \quad (3.4.15)$$

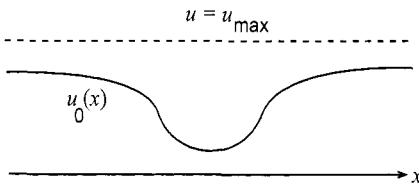


Figure 3.24 Exercise 2.

In $x < vt$ we parameterize the boundary data (3.4.11) by $a = f(\tau)$, $t = \tau$, at $x = 0$. Then the constants c_1 and c_2 in (3.4.13) are given by

$$c_1 = -v\tau \quad c_2 = f(\tau)e^{k\tau}.$$

Then

$$a = f(\tau)e^{-k(t-\tau)} \quad \text{on } x = v(t-\tau),$$

or

$$a(x, t) = f\left(t - \frac{x}{v}\right)e^{-kx/v}, \quad x < vt. \quad (3.4.16)$$

Equations (3.4.15)–(3.4.16) solve the problem (3.4.9)–(3.4.11). Physically, the concentration decays from its initial or boundary value as it moves through the reactor at constant speed $v = Q/A$.

More realistic reactors are nonisothermal. In that case an energy balance equation is required to go with the chemical species equation, and a coupled system for the species concentration and the temperature is obtained.

EXERCISES

1. In traffic flow the flux is given by $\phi(u) = uV(u)$, where V is the velocity, and conditions (3.4.3) hold for the flux. Determine a growth condition on ϕ that is implied by the assumption $V' < 0$.
2. In the traffic flow problem assume that the flux is given by $\phi(u) = au(u_j - u)$, where a and u_j are positive constants. Find the vehicle velocity $V(u)$ and the characteristic speed $c(u)$. Describe how a density wave like the one shown in Figure 3.24 evolves in time. On a characteristic diagram indicate the path of a vehicle approaching such a wave from behind. How does the situation change if the density wave is like the one in Figure 3.25?
3. It was determined experimentally that traffic flow through the Lincoln Tunnel is approximately described by the conservation law $u_t + \phi_x = 0$ with flux

$$\phi(u) = au \ln \frac{u_j}{u},$$

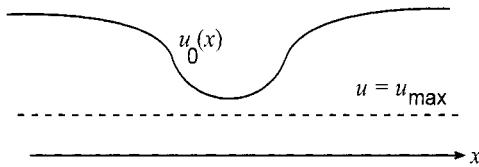


Figure 3.25 Exercise 2.

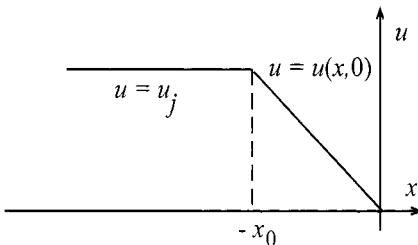


Figure 3.26 Exercise 3.

where a and u_j are known positive constants. Suppose that the initial density u of the traffic varies linearly from bumper-to-bumper traffic (behind $x = -x_0$) to no traffic (ahead of $x = 0$), as sketched in Figure 3.26. Two hours later, where does $u = u_j/2$?

4. (*Flood waves*) The height $h = h(x, t)$ of a flood wave is modeled by the advection equation

$$h_t + q_x = 0,$$

where the water flux is given by the Chezy law $q = vh$, where the average stream velocity v is given by $v = a\sqrt{h}$, where $a > 0$ is a proportionality constant. Show that flood wave propagate 1.5 times faster than the average stream velocity.

3.5 Weak Solutions: A Formal Approach

We took a heuristic approach to introduce the concept of a discontinuous solution. The implication was that a discontinuous solution is a function that is smooth and satisfies the PDE on both sides of a curve, the shock path; along the shock path the function suffers a simple discontinuity and, as well, satisfies a jump condition relating its values on both sides of the discontinuity to the

speed of the discontinuity itself. We want to develop a more formal, mathematical definition of these concepts. The result will be an extension of the notion of a solution to a PDE to include nonclassical solutions.

For definiteness, consider the initial value problem

$$u_t + \phi(u)_x = 0, \quad x \in \mathbb{R}, \quad t > 0, \quad (3.5.1)$$

$$u(x, 0) = u_0(x), \quad x \in \mathbb{R}, \quad (3.5.2)$$

where ϕ is a continuously differentiable function on \mathbb{R} . By a *classical*, or *genuine*, solution to (3.5.1)–(3.5.2) is meant a smooth function $u = u(x, t)$ that satisfies (3.5.1)–(3.5.2). For motivation, assume for the moment that $u = u(x, t)$ is a classical solution, and let $f(x, t)$ be any smooth function that vanishes outside some closed, bounded set in the plane. Closed, bounded sets in the plane are *compact* sets, and the closure of the set of points where a function f is nonzero is called the *support* of f , denoted by $\text{supp } f$. Therefore, the assumption is that f is a function with compact support. The set of all smooth functions with compact support is denoted by C_0^1 . Therefore, we may choose a rectangle $D = \{(x, t) : a \leq x \leq b, 0 \leq t \leq T\}$, where $f = 0$ along $x = a$, $x = b$, and $t = T$ (see Figure 3.27). If we multiply (3.5.1) by f and integrate over D , we obtain

$$0 = \int_0^T \int_a^b (f u_t + f \phi_x) dx dt. \quad (3.5.3)$$

Integrating both terms in (3.5.3) by parts then gives

$$\begin{aligned} \int_0^T \int_a^b f u_t dx dt &= \int_a^b \int_0^T f u_t dt dx = \int_a^b \left[f u \Big|_0^T - \int_0^T f_t u dt \right] dx \\ &= - \int_a^b f(x, 0) u_0(x) dx - \int_a^b \int_0^T f_t u dt dx \end{aligned}$$

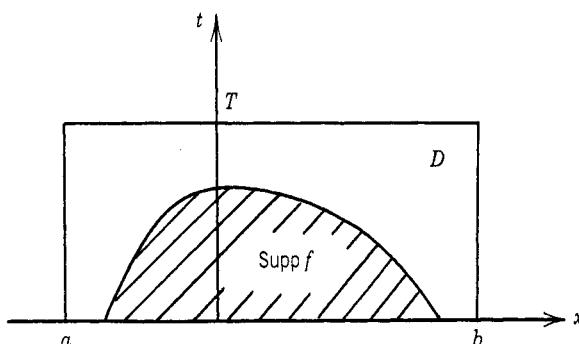


Figure 3.27 Rectangular region containing the support of the function f .

and

$$\int_0^T \int_a^b f \phi_x dx dt = \int_0^T \left[f \phi |_a^b - \int_a^b f_x \phi dx \right] dt = - \int_0^T \int_a^b f_x \phi dx dt.$$

Therefore, from (3.5.3), we have

$$\int_{t \geq 0} \int (u f_t + f_x \phi) dx dt + \int_{t=0} u_0 f dx = 0. \quad (3.5.4)$$

To summarize, if u is a genuine solution to (3.5.1)–(3.5.2) and f is any smooth function with compact support, the integral relation (3.5.4) holds. We now observe that the integrands in (3.5.4) do not involve any derivatives of either u or ϕ . Thus (3.5.4) remains well defined even if u and $\phi(u)$ or their derivatives have discontinuities. We use equation (3.5.4), therefore, as a basis of a definition for a generalized solution of the initial value problem. The formal definition is given below.

Definition. A bounded piecewise smooth function $u(x, t)$ is called a *weak solution* of the initial value problem (3.5.1)–(3.5.2), where u_0 is assumed to be bounded and piecewise smooth if, and only if, (3.5.4) is valid for all smooth functions f with compact support.

This is the general definition of a solution that does not require smoothness. (In other contexts using the Lebesgue integral, we can replace the piecewise smooth condition by *measurable*.)

To verify that the notion of weak solution is indeed an extension of the usual notion of a classical solution, we should actually check that if u is smooth and satisfies (3.5.4) for all functions f with compact support, then u is a solution to the initial value problem (3.5.1)–(3.5.2). We leave this verification as an Exercise.

Does the definition of a weak solution, involving equation (3.5.4), result in a formula for the shock speed that is consistent with the jump condition derived in Section 3.3? The answer is, of course, affirmative, and we now present that argument. This gives an alternative way to derive the jump condition. To carry out the analysis, we recall Green's.

Theorem. (Green's Theorem) Let C be a piecewise smooth, simple closed curve in the xt plane, and let D denote the domain enclosed by C . If $p = p(x, t)$ and $q = q(x, t)$ are smooth functions in $D \cup C$, then

$$\int_C p dx + q dt = \int \int_D (q_x - p_t) dx dt,$$

where the line integral over C is taken in the counterclockwise direction.

Let Γ be a smooth curve in spacetime given by $x = s(t)$ along which u has a simple jump discontinuity, and let D be a ball centered at some point on Γ and lying in the $t > 0$ plane (Figure 3.28). Further, let D_1 and D_2 denote the disjoint subsets of D on each side of Γ . Now, choose $f \in C_0^1$ in D . From condition (3.5.4) we have

$$\begin{aligned} 0 &= \int_D \int (u f_t + \phi(u) f_x) dx dt \\ &= \int_{D_1} \int (u f_t + \phi(u) f_x) dx dt + \int_{D_2} \int (u f_t + \phi(u) f_x) dx dt. \end{aligned} \quad (3.5.5)$$

The product rule for derivatives and the fact that u is smooth in D_2 allows the second integral in (3.5.5) to be written as

$$\begin{aligned} \int_{D_2} \int (u f_t + \phi(u) f_x) dx dt &= \int_{D_2} \int ((u f)_t + (f \phi)_x) dx dt \\ &= \int_{\text{Bd } D_2} -u f dx + f \phi dt, \end{aligned}$$

where in the last step we used Green's theorem. But the line integral is nonzero only along Γ because of the choice of f . Therefore, denoting by u_1 and u_2 the limiting values of u on the front- and backsides of Γ , respectively (i.e., $u_1 = u(s(t)+, t)$ and $u_2 = u(s(t)-, t)$), we have

$$\int_{D_2} \int (u f_t + \phi(u) f_x) dx dt = \int_{P_1}^{P_2} f(-u_2 dx + \phi(u_2) dt). \quad (3.5.6)$$

In a similar manner, we obtain

$$\int_{D_1} \int (u f_t + \phi(u) f_x) dx dt = \int_{P_2}^{P_1} f(-u_1 dx + \phi(u_1) dt). \quad (3.5.7)$$

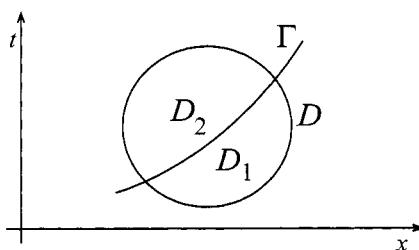


Figure 3.28 Ball D centered on a curve Γ along which a discontinuity in u occurs.

Equations (3.5.5)–(3.5.7) imply

$$\int_{\Gamma} f(-[u] dx + [\phi(u)] dt) = 0, \quad (3.5.8)$$

where the brackets denote the jump in the quantity inside (e.g., $[u] = u_2 - u_1$). Finally, the fact that f may be chosen arbitrarily leads to the conclusion that

$$-[u] dx + [\phi(u)] dt = 0,$$

or

$$-s'(t)[u] + [\phi(u)] = 0, \quad (3.5.9)$$

which is the jump condition obtained in Section 3.3. In conclusion, we showed that the definition of a weak solution leads to the jump condition across a shock.

The weak form of solutions given in integral form by (3.5.4) ties down the idea; where the solution is smooth, the PDE holds, and where the solution is discontinuous, the shock condition is implied. The next example shows that care must be taken when going from a PDE to the integrated form of the weak solution.

Example. Consider the PDE

$$u_t + uu_x = 0. \quad (3.5.10)$$

Any positive smooth solution of this equation is a smooth solution of both the conservation laws

$$u_t + \left(\frac{u^2}{2}\right)_x = 0 \quad \text{and} \quad \left(\frac{u^2}{2}\right)_t + \left(\frac{u^3}{3}\right)_x = 0, \quad (3.5.11)$$

each of which satisfies different jump conditions, because the flux is different. Thus, to one PDE there are associated different weak solutions, depending on the form in which the equation is written. \square

Therefore we cannot tell what the correct shock condition is if the only thing given is the PDE. This means that the most basic, relevant piece of information is the integral form of the conservation law or the PDE in conservation form, where the flux is identified. Fortunately, in real applications in engineering and sciences, the integral form of the conservation law usually comes out initially from the modeling process. If the underlying physical process is unknown, a PDE by itself does not uniquely determine the conservation law or shock conditions.

Exercise 6 shows that the uniqueness problem is in fact even worse than expected. Weak solutions are not unique, and another condition is required if we want otherwise. Such a condition is called an *entropy condition*, which

dictates how the speed of the characteristics ahead of and behind a shock wave must relate to the shock speed itself. For the PDE (3.5.1) with $\phi''(u) > 0$, it can be proved that there is a unique solution to the initial value problem (3.5.1)–(3.5.2) that also satisfies the condition

$$u(x+a,t) - u(x,t) \leq \frac{aE}{t}, \quad x \in \mathbb{R}, \quad t > 0, \quad (3.5.12)$$

where $a > 0$ and E is a positive constant independent of x , t , and a . Equation (3.5.12) is called the *entropy condition*, and it requires that a solution can jump downward only as we traverse a discontinuity from left to right. In terms of characteristic speeds, (3.5.12) implies (Exercise 7) the entropy inequality

$$\phi'(u_2) > s' > \phi'(u_1), \quad (3.5.13)$$

where s' is the shock speed and u_1 and u_2 are the states ahead of and behind the shock, respectively. Equation (3.5.13) establishes the shock speed as intermediate between the speeds of the characteristics ahead and behind. The entropy condition originates from a condition in gas dynamics that requires the entropy increase (as required in the second law of thermodynamics) across a shock wave, the latter of which is basically an irreversible process. This condition precludes the existence of rarefaction shocks, where there is a positive jump from left to right across the shock wave. The proof of existence and uniqueness of a weak solution satisfying the entropy condition is beyond the scope of this text, and we refer the reader to Lax (1973) or Smoller (1994) for accessible treatments.

EXERCISES

- Let u and u_0 be smooth functions that satisfy (3.5.4) for all $f \in C_0^1$. Prove that u and u_0 satisfy the initial value problem (3.5.1)–(3.5.2). *Hint:* Choose f judiciously, use integration by parts, and use the fact that if $\int h f dx = 0$ for all $f \in C_0^1$, where h is continuous, then $h = 0$.
- Find a weak solution to each of the following initial value problems:

(a)

$$u_t + (e^u)_x = 0, \quad x \in \mathbb{R}, \quad t > 0,$$

$$u(x, 0) = 2 \quad \text{if } x < 0; \quad u(x, 0) = 1 \quad \text{if } x > 0.$$

(b)

$$u_t + 2uu_x = 0, \quad x \in \mathbb{R}, \quad t > 0,$$

$$u(x, 0) = \sqrt{-x} \quad \text{if } x < 0; \quad u(x, 0) = 0 \quad \text{if } x > 0.$$

3. (*Porous Media*) Consider the flow of an incompressible fluid through a porous medium (e.g., water flowing through soil). The soil is assumed to contain a large number of pores through which the water can infiltrate downward, driven by gravity. With the positive x axis oriented downward, let $u = u(x, t)$ denote the ratio of the water-filled pore space to the total pore space at the depth x , at time t . Thus, $u = 0$ is dry soil and $u = 1$ is *saturated* soil, where all the pore space is filled with water. The basic conservation law holds for volumetric water content u , namely

$$u_t + \phi_x = 0,$$

where ϕ is the vertical flux. In one model of partially saturated, gravity-dominated flow, the flux is assumed to have the form

$$\phi = ku^n,$$

where k and n are positive constants depending on soil characteristics; k is called the *hydraulic conductivity*. In the following, take $n = 2$ and $k = 1$. If the moisture distribution in the soil is $u = \frac{1}{2}$ initially, and if at the surface $x = 0$ the water content is $u = 1$ (saturated) for $0 < t < 1$ and $u = \frac{1}{2}$ for $t > 1$, what is the water content $u(x, t)$ for $x, t > 0$? Plot the jump in u across the shock as a function of time.

4. (*Traffic Flow*) In the model of traffic flow presented in Section 3.4, the velocity V is found experimentally to be related to the density u by the constitutive relation $V = v_0(1 - u/U)$, where v_0 is the maximum speed and U is the maximum density. At time $t = 0$ cars are traveling along the roadway with constant velocity $2v_0/3$, and a truck enters the roadway at $x = 0$ traveling at speed $v_0/3$. It continues at this constant speed until it exits at $x = L$.

(a) On a spacetime diagram sketch the path of the truck and the path of the shock while the truck is on the road.

(b) After the truck leaves the road, find the time that the car following the truck takes to catch the traffic ahead of the truck.

(c) Describe the qualitative behavior of the weak solution for all $t > 0$.

5. Determine the jump conditions associated with the equations (3.5.11).

6. Consider the initial value problem

$$\begin{aligned} u_t + uu_x &= 0, \quad x \in \mathbb{R}, \quad t > 0, \\ u(x, 0) &= 0 \quad \text{if } x < 0; \quad u(x, 0) = 1 \quad \text{if } x > 0, \end{aligned}$$

and let u and v be defined by

$$\begin{aligned} u(x, t) &= 0 \quad \text{if } x < \frac{t}{2}, & u(x, 0) &= 1 \quad \text{if } x > \frac{t}{2} \\ v(x, t) &= 0 \quad \text{if } x < 0, & v(x, t) &= \frac{x}{t} \quad \text{if } 0 < x < t, \\ v(x, t) &= 1 \quad \text{if } x > t. \end{aligned}$$

Sketch u and v and the characteristics. Are both u and v weak solutions to the initial value problem? If the entropy condition (3.5.12) is assumed, which is the correct solution?

7. Show the entropy condition (3.5.12) implies the entropy inequality (3.5.13).
8. Consider the initial value problem

$$\begin{aligned} u_t + \left(\frac{u^2}{2} \right)_x &= 0, \quad x \in \mathbb{R}, \quad t > 0, \\ u(x, 0) &= 0 \quad \text{if } x > 0 \text{ and } u(x, 0) = U(x) \quad \text{if } x < 0, \end{aligned}$$

where U is unknown. Determine conditions on a shock $x = s(t)$, $t > 0$, in the upper half-plane that are consistent with definition of a weak solution, and then, under these conditions, calculate U in terms of $s(t)$. What is $U(x)$ if $s(t) = \sqrt{t+1} - 1$?

9. Consider the conservation law

$$\frac{d}{dt} \int_a^b u(x, t) dx = \frac{1}{2} [u(a, t)^2 - u(b, t)^2] + \int_a^b g(x) dx,$$

where g is a discontinuous source term given by $g(x) = 1$ if $x > \frac{1}{2}$ and $g(x) = C = \text{const}$ if $x < \frac{1}{2}$. Find the weak solution when the initial condition is $u(x, 0) = U = \text{const}$ for all $x \in \mathbb{R}$.

10. Consider the initial-boundary value problem

$$\begin{aligned} u_t + u_x &= -u + u(1, t), & 0 < x < 1, \quad t > 0, \\ u(x, 0) &= u_0(x), \quad 0 < x < 1; \quad u(0, t) = u_b(t), \quad t > 0. \end{aligned}$$

This problem has an unknown source term evaluated on the boundary of the domain, and it arises in a problem in detonation theory.

- (a) Find an analytic expression for $u(1, t)$, $0 < t < 1$, and determine the solution in $t < x < 1$.
- (b) Derive a differential-difference equation for $u(1, t)$, $t > 1$, and a formula for the solution in terms of $u(1, t)$ in the region $0 < x < t$, $x < 1$.
Solution: $u'(1, t) + au(1, t-1) = a[u_b(t-1) + u'_b(t-1)]$, $a = 1/e$.

- (c) If $u_b(0) = u_0(0)$, show that the jump condition across the line $x = t$ is given by

$$u_+(t) - u(t) = e^{-t}[u_b(0) - u_0(0)].$$

- (d) Find the solution in the case $u_b(t) = 0$ and $u_0(x) = x$.
11. Show that $u(x, t) = x/t$ if $-\sqrt{t} < x < \sqrt{t}$, and $u(x, t) = 0$ otherwise, is a weak solution to the equation $u_t + (u^2/2)_x = 0$ on the domain $x \in \mathbb{R}$, $t \geq t_0 > 0$. Where are the shocks, and what are their speeds? Sketch a characteristic diagram and a typical wave profile for $t > t_0$, and show that the entropy condition holds.
12. Solve the Riemann problem

$$u_t + \phi_x = 0, \quad u(x, 0) = 1 \quad \text{if } x < 0, \text{ and } u(x, 0) = 0 \quad \text{if } x > 0,$$

where the flux ϕ is

$$\phi(u) = \frac{u^2}{u^2 + (1-u)^2/2}.$$

3.6 Asymptotic Behavior of Shocks

3.6.1 Equal-Area Principle

Next, in a simple context, we study the evolution of an initial waveform over its entire history, as propagated by the nonlinear equation $u_t + c(u)u_x = 0$. The scenario is as follows. An initial signal $u_0(x)$ at $t = 0$ begins to distort and “shocks up” at breaking time $t = t_b$; after the shock forms, the shock discontinuity follows a path $x = s(t)$ in spacetime. One of the fundamental problems in nonlinear analysis is to determine the shock path and how the strength of the shock varies along that path for large times t .

Central to the calculation of the asymptotic form is an interesting geometric result called the *equal-area principle*. To formulate this principle in a simple manner, we limit the types of initial data that we consider. We say that an initial profile $u = u_0(x)$ satisfies property (P) if $u_0(x)$ is a smooth, nonnegative pulse; that is, u_0 belongs to $C^1(\mathbb{R})$ and has a single maximum (say, at $x = x_0$); u_0 is nondecreasing for $x < x_0$ and nonincreasing for $x > x_0$, and u_0 approaches a nonnegative constant value as $|x| \rightarrow \infty$. Therefore, we consider the initial value problem

$$u_t + c(u)u_x = 0, \quad x \in \mathbb{R}, \quad t > 0, \tag{3.6.1}$$

$$u(x, 0) = u_0(x), \quad x \in \mathbb{R}, \tag{3.6.2}$$

where u_0 satisfies property (P) and $c(u) = \phi'(u) > 0, c'(u) \geq 0$. We know from Section 3.3 that a classical solution exists only up to some time $t = t_b$ when a gradient catastrophe occurs, and it is given implicitly by

$$u = u_0(\xi), \quad (3.6.3)$$

$$x = \xi + F(\xi)t, \quad F(\xi) = c(u_0(\xi)). \quad (3.6.4)$$

Equation (3.6.4) defines straight-line characteristic curves propagating from $x = \xi$ at $t = 0$ into the spacetime domain with speed $F(\xi)$.

Accordingly, as the initial signal propagates, the wave breaks at the first time t_b where the characteristics collide. For $t > t_b$, instead of forming the weak solution with a shock, let us assume that the wave actually breaks (like an ocean wave), forming a multivalued wavelet where the intersecting characteristics are carrying their constant values of u . We interpret the formation of this wavelet as shown in Figure 3.29. A point $P : (\xi, u)$ on the initial wave at $t = 0$ is propagated to a point $P' : (x, u)$ on the multivalued wavelet at time t , where the coordinate x at time t is given by (3.6.4); points having height u are propagated at speed $c(u)$. We may now state:

Now we define the equal-area principle: The location of the shock $x = s(t)$ at time t is the position at which a vertical line cuts off equal area lobes of the multivalued wavelet. Figure 3.30 shows the geometric interpretation of this rule.

Two simple facts are required to demonstrate the equal-area principle, and these are recorded formally in the following two lemmas. The first states that the area under a section of the wavelet remains unchanged as that section

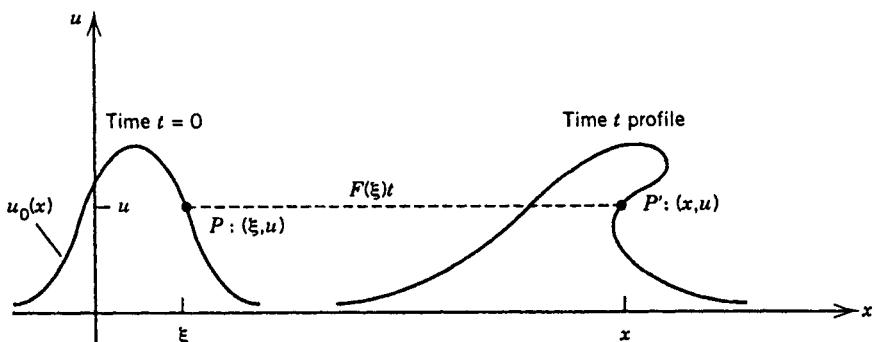


Figure 3.29 Initial wave profile evolving into a multivalued wavelet. Points at height u on the wave are propagated at speed $F(\xi) = c(u_0(\xi))$.

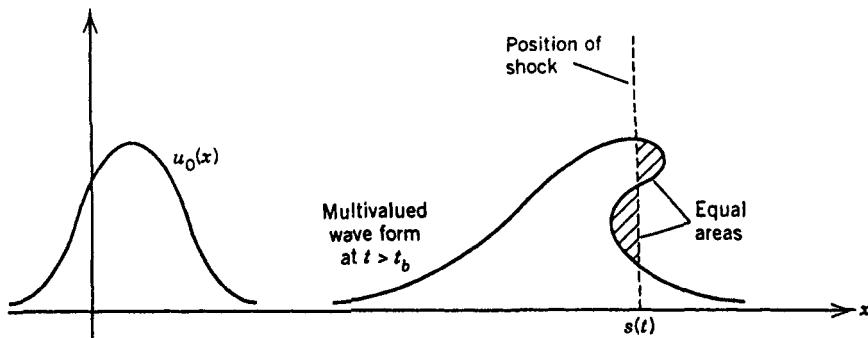


Figure 3.30 Equal-area rule.

is propagated in time (see Figure 3.31), and the second states that the area under the weak solution (i.e., the discontinuous solution with a shock) remains unchanged as it is propagated in time (see Figure 3.32).

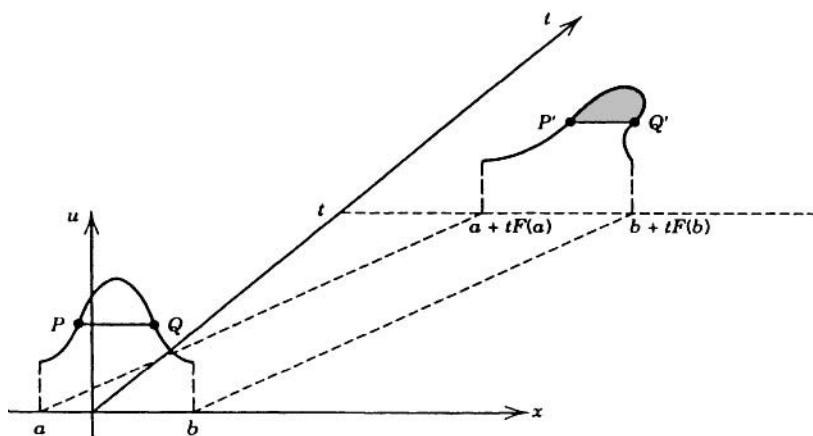


Figure 3.31

Lemma. Consider a section of the initial waveform between $x = a$ and $x = b$, with $u_0(a) = u_0(b)$, with u_0 satisfying property (P). Then, for any $t > 0$ the area under this section of the wave remains constant as it propagates in time.

The proof of this fact is geometrically obvious (see Figure 3.31). From (3.6.4), the x coordinates a and b are moved in time t to $a + F(a)t$ and $b + F(b)t$, respectively. Any horizontal line segment PQ at $t = 0$ has the same length as

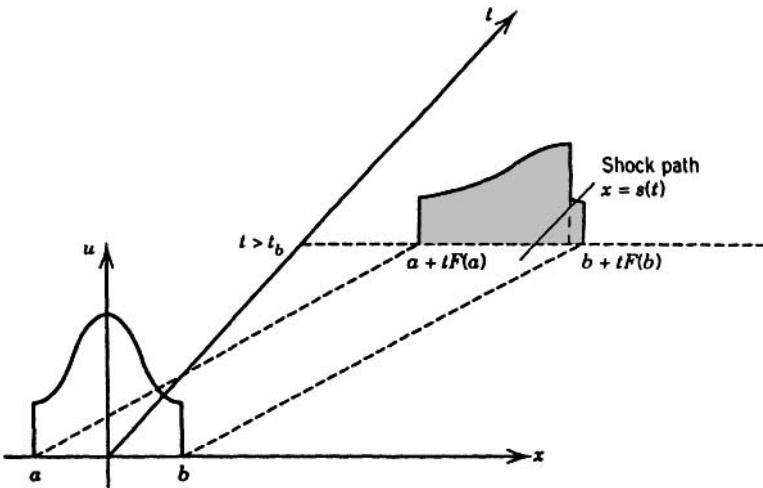


Figure 3.32

$P'Q'$ at time t because points on the wave at the same height u move at the same speed $c(u)$.

Lemma. Let a and b be chosen such that $u_0(a) = u_0(b)$, where u_0 satisfies property (P), and assume that the shock locus is given by $x = s(t)$ with $a + tF(a) < s(t) < b + tF(b)$ for t in some open interval I . Then

$$\int_{a+tF(a)}^{b+tF(b)} u(x, t) dx = \text{const}, \quad t \in I, \quad (3.6.5)$$

where $u = u(x, t)$ is the weak solution.

The proof of (3.6.5) can be carried out by showing the derivative of the left side of (3.6.5) is zero. We leave this calculation as an exercise.

The equal-area principle now follows easily from the two lemmas because the multivalued waveform shown in Figure 3.31 and the weak solution shown in Figure 3.32 both encompass the same area, the area under the initial waveform from $x = a$ to $x = b$. We remark that these lemmas and the equal-area rule hold for more general initial data than that satisfying property (P). Furthermore, in the argument above we assumed that $c(u) > 0$; equally valid is the case $c(u) < 0$, with the appropriate adjustments.

3.6.2 Shock Fitting

The general problem of determining the position of the shock locus is fundamental in hyperbolic systems. The equal-area principle leads to a quantitative condition whereby the shock path $x = s(t)$ can be calculated directly for certain simple equations. In this subsection we consider the initial value problem for the inviscid Burgers' equation:

$$u_t + uu_x = 0, \quad x \in \mathbb{R}, \quad t > 0, \quad (3.6.6)$$

$$u(x, 0) = u_0(x), \quad x \in \mathbb{R}, \quad (3.6.7)$$

where u_0 satisfies property (P). Thus $c(u) = u$; the case for arbitrary $c(u)$ is discussed in Whitham (1974).

First suppose that the shock is at some yet undetermined location $x = s(t)$ at some time $t > t_b$, and assume that the two characteristics carrying data to the shock from ahead and behind are specified by $\xi_1 = \xi_1(t)$ and $\xi_2 = \xi_2(t)$, respectively (Figure 3.33). By construction the equations of these two characteristics are

$$x = \xi_1 + u_0(\xi_1)t, \quad x = \xi_2 + u_0(\xi_2)t,$$

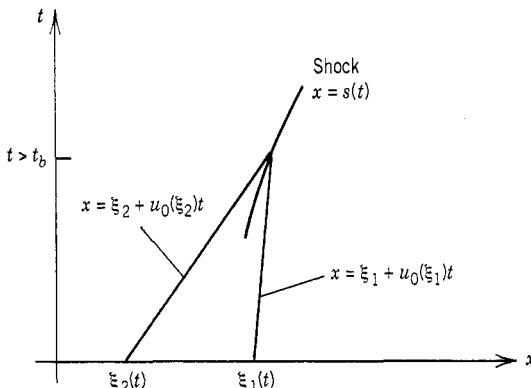


Figure 3.33 Two characteristics from behind and ahead that intersect the shock at time t .

and therefore

$$s(t) = \xi_1 + u_0(\xi_1)t \quad s(t) = \xi_2 + u_0(\xi_2)t. \quad (3.6.8)$$

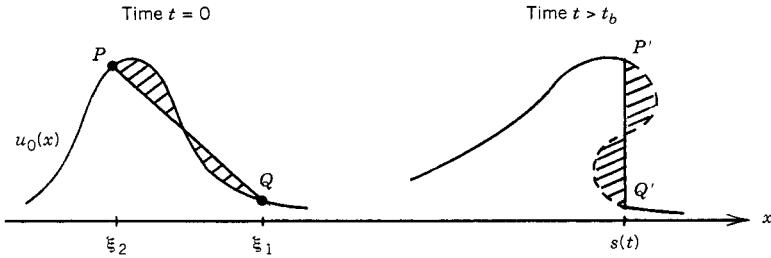


Figure 3.34 Whitham's rule. The chord PQ on the initial curve that cuts off equal areas has slope $-1/t$, where t is the time at the shock.

It follows immediately from (3.6.8) that

$$t = -\frac{\xi_1 - \xi_2}{u_0(\xi_1) - u_0(\xi_2)} > 0, \quad (3.6.9)$$

which is an equation for the time at the shock in terms of the known initial signal and the still unknown values of ξ_1 and ξ_2 . The ratio on the right side of (3.6.9) is the reciprocal of the slope of the chord PQ on the initial wave profile u_0 connecting the points $(\xi_2, u_0(\xi_2))$ and $(\xi_1, u_0(\xi_1))$, as shown in Figure 3.34. Because the movement of the wavelet in xu space is governed by the linear transformation $(x, u) \rightarrow (x + ut, u)$, straight lines are mapped to straight lines, and therefore the chord PQ is mapped to a vertical segment $P'Q'$ at time t . [This argument cannot be made in the case for arbitrary $c(u)$.] This vertical segment is the shock at time t , and the shock strength is clearly given by the formula

$$\text{shock strength} = u_0(\xi_2) - u_0(\xi_1). \quad (3.6.10)$$

Remember, ξ_1 and ξ_2 are still undetermined. Equation (3.6.8) gives two equations for ξ_1 , ξ_2 , and $s(t)$, and therefore we need one additional relation to determine these quantities and thus to determine the shock path. This additional relation comes from the equal-area principle. Because the shock cuts off equal-area lobes in the multivalued wavelet, the chord PQ must cut off equal-area lobes on the initial wave profile, as shown in Figure 3.34. This means that the area under the initial wave profile u_0 from ξ_2 to ξ_1 must equal the area of the trapezoid whose upper boundary is the chord PQ . Or, quantitatively, this is

$$\int_{\xi_2}^{\xi_1} u_0(\xi) d\xi = \frac{1}{2}[u_0(\xi_1) + u_0(\xi_2)](\xi_1 - \xi_2). \quad (3.6.11)$$

Therefore, in principle, we have developed an algorithm to determine the shock position $s(t)$: For each fixed $t > t_b$, solve the three equations in (3.6.8) and (3.6.11) to determine $s(t)$, $\xi_1(t)$, and $\xi_2(t)$. It is clear that one cannot always

carry out this calculation analytically, and consequently, a numerical procedure may be the only recourse.

However, one can proceed geometrically by *Whitham's rule*, a simple construction on the initial waveform to determine the location and time evolution of the shock. This rule is based on equation (3.6.9).

Whitham's Rule. For a succession of values of time t , draw a chord on the initial profile u_0 of slope $-1/t$ cutting off equal-area lobes. Then shift the endpoints of the chord to the right by the amounts $tu_0(\xi_2)$ and $tu_0(\xi_1)$, respectively, obtaining the location of the shock at the time t .

Some practice in applying this rule is given in the Exercises.

3.6.3 Asymptotic Behavior

Limiting the discussion to the initial value problem (3.6.6)–(3.6.7) for the inviscid Burgers' equation, we next investigate the long-time behavior of the shock path and the shock strength. For definiteness, we consider an initial wave profile u_0 satisfying property (P) with the additional stipulation that $u_0(x) = u^*$ for $x \leq 0$ and for $x \geq a > 0$. Here u^* is a fixed positive constant. Thus $u_0(x)$ is a single hump, and we assume that the area under the hump, that is, the area bounded by $u_0(x)$, $u = u^*$, $x = 0$, and $x = a$, is A . From (3.6.9) the shock strength must vanish as t increases; the question is the rate of decay and the path of the shock. We demonstrate that for large t

$$\text{Shock strength} \sim \left(\frac{2A}{t} \right)^{1/2}, \quad (3.6.12)$$

$$s(t) \sim u^* t + (2At)^{1/2}. \quad (3.6.13)$$

Thus the shock strength decays like $t^{-1/2}$ and the shock path is eventually parabolic. We also show that for large t the wave profile behind the shock is given by

$$u(x, t) \sim \frac{x}{t}. \quad (3.6.14)$$

The symbol \sim means in a limiting or asymptotic sense; we use the notation $t \gg 1$ to indicate that t is large.

The initial wave profile is shown in Figure 3.35. Using the equal-area rule applied to the initial profile, along with equation (3.6.9), we observe that $\xi_1(t)$ will eventually be larger than a because the slope of the chord must go to zero for large t . Therefore

$$u_0(\xi_1(t)) = u^* \quad \text{for } t \gg 1. \quad (3.6.15)$$

Now, subtraction of $u^*(\xi_1 - \xi_2)$ from both sides of equation (3.6.11) yields

$$\int_{\xi_2}^{\xi_1} [u_0(\xi) - u^*] d\xi = \frac{1}{2}(\xi_1 - \xi_2)[u_0(\xi_2) - u^*], \quad \xi_1 > a.$$

Using (3.6.9) to replace $\xi_1 - \xi_2$ gives

$$\int_{\xi_2}^{\xi_1} [u_0(\xi) - u^*] d\xi = \frac{t}{2}[u_0(\xi_2) - u^*]^2, \quad \xi_1 > a \quad (3.6.16)$$

But the left side of (3.6.16) can be written

$$\int_{\xi_2}^a [u_0(\xi) - u^*] d\xi + \int_a^{\xi_1} [u_0(\xi) - u^*] dx = \int_{\xi_2}^a [u_0(\xi) - u^*] d\xi.$$

Hence (3.6.16) becomes

$$\int_{\xi_2}^a [u_0(\xi) - u^*] d\xi = \frac{t}{2}[u_0(\xi_2) - u^*]^2, \quad \xi_1 > a. \quad (3.6.17)$$

For $t \gg 1$ the left side of (3.6.17) approaches the area A because ξ_2 approaches 0 for $t \gg 1$. Further, the quantity $u_0(\xi_2) - u^*$ is the shock strength. Thus, for $t \gg 1$, the estimate (3.6.12) holds.

To deduce (3.6.13), we use equation (3.6.8) in the form

$$s(t) = \xi_2(t) + tu_0(\xi_2(t)).$$

But from (3.6.12) we have $u_0(\xi_2(t)) \sim u^* + (2A/t)^{1/2}$ for $t \gg 1$, and we also have $\xi_2(t) \sim 0$ for $t \gg 1$. Relation (3.6.13) follows.

To prove that the wave has the form (3.6.14), we note that the point on the wavelet at $x = 0$ at $t = 0$ moves at speed u^* and hence is located at $x = u^*t$ at time t . Clearly, $u = u^*$ for $x < u^*t$. At the same time the shock is

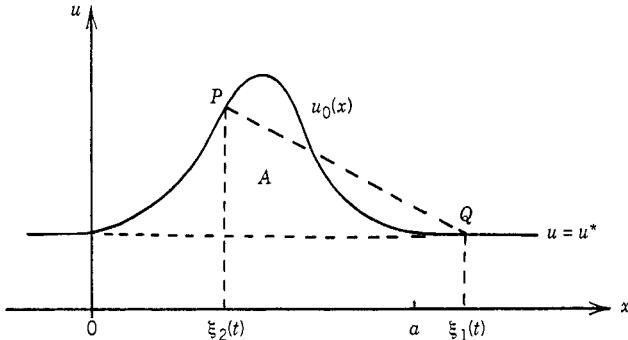


Figure 3.35 Initial wave profile $u_0(x)$ having area A .

at $s(t) = u^*t + (2At)^{1/2}$, and $u = u^*$ for $x > s(t)$. So we need the solution in the interval $u^*t < x < u^*t + (2At)^{1/2}$. But the solution in this interval is as follows, from (3.6.3) and (3.6.4),

$$u = u_0(\xi), \quad \text{where } x = \xi + tu_0(\xi).$$

Therefore

$$u = \frac{x - \xi}{t}.$$

But for $t \gg 1$ we have $\xi = \xi_2 \sim 0$, and so

$$u(x, t) \sim \frac{x}{t}, \quad u^*t < x < u^*t + (2At)^{1/2}, \quad t \gg 1. \quad (3.6.18)$$

Figure 3.36 shows the triangular wave that develops for long times.

EXERCISES

1. Prove the second lemma following the equal-area principle. *Hint:* Break up the integral into two parts, each having a smooth integrand. Differentiate using Leibniz' rule, and then use the jump conditions that hold across the shock.
2. Use Whitham's rule to determine the shock position geometrically and the wave profile for the initial value problem

$$u_t + uu_x = 0, \quad u(x, 0) = \frac{2}{1+x^2}$$

at the times $t = 1, 3, 5, 7$. When does this wave break?

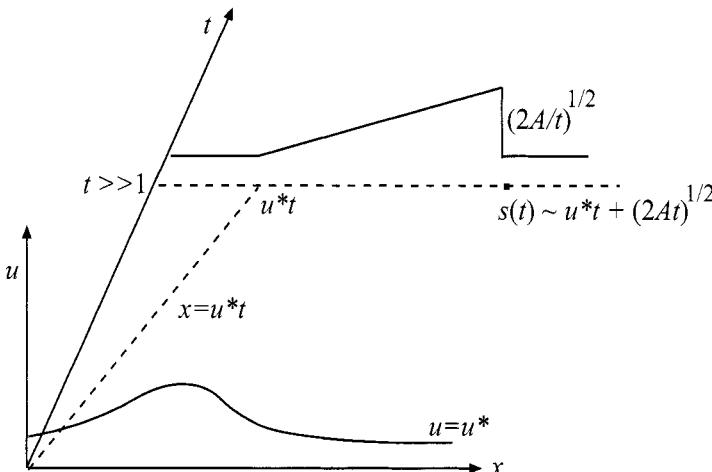


Figure 3.36 Evolution of a wave profile into a triangular wave for long times.

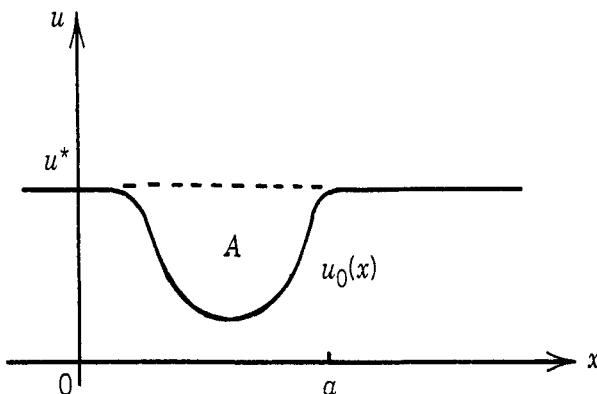


Figure 3.37 Initial wave profile $u_0(x)$ with a dip of area A .

3. Consider the initial value problem (3.6.6)–(3.6.7) where the initial wavelet has the form shown in Figure 3.37. Determine the form of the wavelet, the strength of the shock, and the shock locus for $t \gg 1$.

Reference Notes. Practically all of the general texts mentioned in the reference notes at the end of Chapter 1 discuss first-order partial differential equations and characteristics. Particularly relevant is Whitham (1974), especially with regard to the asymptotic behavior of shocks and the equal-area rule. Also recommended is Lax (1973) and Lighthill (1978). The latter has a detailed discussion of the equal-area rule. The classic treatise on fluid mechanics by Landau & Lifshitz (1987) discusses these topics from an intuitive, physical viewpoint.

The two volumes on first-order equations by Rhee et al. (2001) contain a wealth of information and examples.

An example of a plug flow reactor is the digestive tract in many animals. Logan et al. (2002) examine simple models of digestion with temperature and location dependence. Texts in chemical engineering focusing on reaction kinetics are a good source of nontrivial examples.

4

Hyperbolic Systems

Most physical systems involve several unknown functions. For example, the complete description of a fluid mechanical system might require knowledge of the density, pressure, temperature, and the particle velocity, so we would expect to formulate a system of PDEs to describe the flow. The central idea for hyperbolic systems, as for a single equation, is that of characteristics, and it is this thread that weaves through the entire subject.

Section 4.1 develops, from first principles, two physical models that serve to illustrate the concepts. First, we derive the PDEs that govern waves in shallow water; this derivation is simple and serves as a precursor to the formulation of the conservation laws of gas dynamics; the latter is the paradigm of applied nonlinear hyperbolic PDEs. Section 4.2 extends the notion of characteristics to systems of equations, and out of this development evolves a formal classification of hyperbolicity. Section 4.3 gives several examples of Riemann's method, or the method of characteristics, applied to the shallow water equations with various boundary conditions. Sections 4.4 and 4.5 discuss the theory of weakly nonlinear waves and conditions along a wavefront propagating into a region of spacetime. Because Burgers' equation is fundamental in PDEs, we give its derivation from the equations of gas dynamics in a weakly nonlinear limit.

4.1 Shallow-Water Waves; Gas Dynamics

4.1.1 Shallow-Water Waves

The partial differential equations governing waves in shallow water can be obtained from a limiting form of the general equations of fluid mechanics when the ratio of the water depth to the wavelength of a typical surface wave is small. In the present treatment, however, we obtain these equations *a priori* without the benefit of knowing the conservation laws of general fluid flow. The shallow-water equations, which represent mass and momentum balance in the fluid, are highly typical of other hyperbolic systems, and thus they provide a solid example of the methods of nonlinear analysis in a simple context.

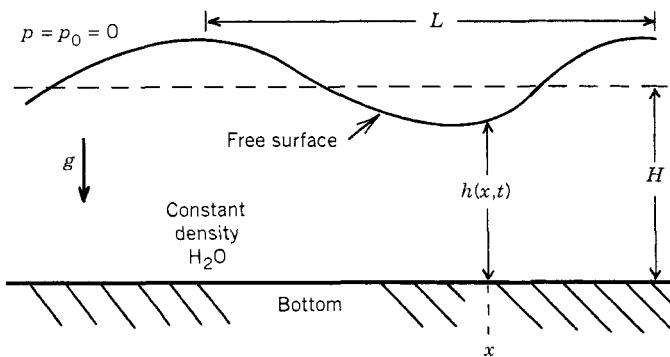


Figure 4.1

Figure 4.1 shows the geometry. Water, which has constant density ρ , lies above a flat bottom. We measure the distance along the bottom by a coordinate x , and y measures the height above the bottom. There is assumed to be no variation in the z direction, which is into the paper. Let H be the height of the undisturbed surface, and assume that the height of the free surface is given at any time t and location x by $y = h(x, t)$. At the free surface the pressure is p_0 , the ambient air pressure. With no loss of generality, we may assume that $p_0 = 0$. If a typical wave on the surface has wavelength L , the shallow water assumption is that H is small compared to L (i.e., $H \ll L$). In this case, we may ignore vertical motions and assume that there is a flow velocity $u(x, t)$ in the x direction that is an average velocity over the depth.

Because vertical motions are ignored, each volume $\Delta x \Delta y \Delta z$ of fluid is in hydrostatic balance in the vertical direction. That is, the upward pressure force on a fluid volume equals the downward pressure force plus the force caused

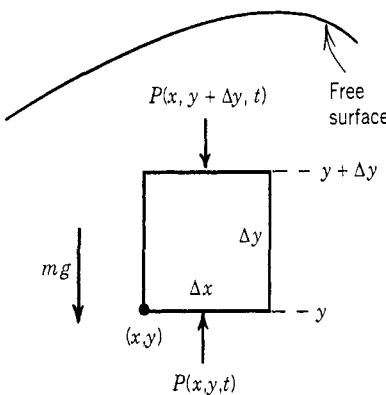


Figure 4.2

by gravity. Figure 4.2 shows a typical fluid element. If pressure is denoted by $P(x, y, t)$, we can write the hydrostatic balance equation as

$$P(x, y, t) \Delta x \Delta z = P(x, y + \Delta y, t) \Delta x \Delta z + \rho g \Delta x \Delta y \Delta z.$$

Dividing by $\Delta x \Delta y \Delta z$ and taking the limit as $\Delta y \rightarrow 0$ yields

$$P_y(x, y, t) = -\rho g.$$

Integrating this equation from a depth y to the free surface h gives an equation for the pressure P :

$$P(x, y, t) = p_0 + \rho g(h(x, t) - y) \quad (p_0 = 0). \quad (4.1.1)$$

Next we derive an equation for *mass balance*, which requires that the time rate of change of the total mass in a region between $x = a$ and $x = b$ equal the mass flux into the region at $x = a$ minus the mass flux out of the region at $x = b$. In the present case, the mass flux at x is given by the density times the velocity u at x times the cross-sectional area at x ; the latter is given by $h \Delta z$, where the width in the z direction is taken to be Δz . Thus mass flux = $\rho u h \Delta z$, in units of mass per unit time. Therefore, the mass balance law is

$$\frac{d}{dt} \int_a^b h(x, t) dx = u(a, t)h(a, t) - u(b, t)h(b, t), \quad (4.1.2)$$

where the constant quantity $\rho \Delta z$ canceled from each term. As usual, if u and h are smooth, then equation (4.1.2) may be cast into the PDE

$$h_t + (uh)_x = 0. \quad (4.1.3)$$

Equations (4.1.2) and (4.1.3) represent mass conservation in integral and differential form, respectively. We recall that the former must be used when u and h have discontinuities.

The next equation is *momentum balance in the x direction*. [Recall that we assumed hydrostatic balance in the y direction, giving (4.1.2).] Momentum balance requires that the time rate of change of the total momentum in the zone between $x = a$ and $x = b$ equal the net momentum flux (the rate at which momentum flows into the region at $x = a$ minus the rate of momentum flowing out of the region at $x = b$), plus the forces (pressure) acting on the region (at $x = a$ and $x = b$). Momentum is mass times velocity, so the total momentum in $a \leq x \leq b$ is given by

$$\text{Total momentum} = \int_a^b \rho u(x, t) h(x, t) \Delta z \, dx.$$

Momentum flux through a section at x is the momentum multiplied by the velocity at x , or $(\rho u h \Delta z)u$. Therefore, the net momentum flux is given by

$$\text{Net momentum flux} = \rho h(a, t) u(a, t)^2 \Delta z - \rho h(b, t) u(b, t)^2 \Delta z.$$

The force acting on the face of area $h(a, t) \Delta z$ is the pressure $P(a, y, t)$. But the pressure varies with depth according to (4.1.2), and therefore we must integrate over the depth to obtain the total force acting at $x = a$, which is

$$\begin{aligned} (\text{Pressure force at } x = a) &= \int_0^{h(a, t)} P(a, y, t) \Delta z \, dy \\ &= \int_0^{h(a, t)} \rho g(h(a, t) - y) \Delta z \, dy \\ &= \frac{\rho g}{2} h(a, t)^2 \Delta z. \end{aligned}$$

Similarly

$$(\text{Pressure force at } x = b) = - \int_0^{h(b, t)} P(b, y, t) \Delta z \, dy = - \frac{\rho g}{2} h(b, t)^2 \Delta z.$$

Consequently, the integral form of the momentum balance equation is

$$\begin{aligned} \frac{d}{dt} \int_a^b h(x, t) u(x, t) \, dx \\ = h(a, t) u(a, t)^2 - h(b, t) u(b, t)^2 + \frac{g}{2} [h(a, t)^2 - h(b, t)^2], \end{aligned} \tag{4.1.4}$$

where $\rho \Delta z$ canceled from each term. A differential form of this conservation law may be obtained if u and h have the requisite degree of smoothness. Then

$$(hu)_t + \left(hu^2 + \frac{g}{2} h^2 \right)_x = 0. \tag{4.1.5}$$

Equations (4.1.3) and (4.1.5) are coupled nonlinear PDEs for the velocity u and the height h of the free surface, and these equations are the governing evolution equations for shallow-water waves. Clearly, equation (4.1.3) can be expanded to give

$$h_t + uh_x + hu_x = 0. \quad (4.1.6)$$

A simple exercise shows that the momentum equation (4.1.5) can be written, with the aid of (4.1.6), simply as

$$u_t + uu_x + gh_x = 0. \quad (4.1.7)$$

Equations (4.1.6) and (4.1.7) are model equations for shallow-water theory. They are quasilinear and resemble the nonlinear model equations from Chapters 2 and 3.

4.1.2 Small-Amplitude Approximation

If we restrict the analysis to small-amplitude waves on the surface that do not deviate much from the undisturbed depth H , and if the velocity is small, then (4.1.6) and (4.1.7) can be linearized. To this end, assume that

$$h(x, t) = H + \eta(x, t), \quad u(x, t) = 0 + v(x, t), \quad (4.1.8)$$

where η and v , and their derivatives, are small (η is small compared to H). The small quantities η and v are perturbations, and they represent quantities that deviate slightly from the constant equilibrium state $h = H$, $u = 0$. Substituting (4.1.8) into (4.1.6) and (4.1.7) give nonlinear equations for these deviations:

$$\eta_t + v\eta_x + (H + \eta)v_x = 0, \quad v_t + vv_x + g\eta_x = 0. \quad (4.1.9)$$

These equations, called the *nonlinear perturbation equations*, govern the evolution of the small deviations from equilibrium, assuming that some initial and boundary conditions are specified. Equations (4.1.9) are no simpler than the original shallow-water equations; however, if we retain only the linear terms in (4.1.9), we have

$$\eta_t + Hv_x = 0, \quad v_t + g\eta_x = 0. \quad (4.1.10)$$

The reasoning is that quadratic terms such as $v\eta_x$ are small compared to linear terms. Equations (4.1.10) are called the *linearized perturbation equations*. Now, it is easy to eliminate one of the unknowns in (4.1.10) and obtain a single, linear equation in one unknown. Differentiating the first equation in (4.1.10) with respect to t and the second with respect to x , and then using the equality of v_{xt} and v_{tx} , we obtain

$$\eta_{tt} - gH\eta_{xx} = 0. \quad (4.1.11)$$

We recognize this as the classical linear wave equation for η whose solution consists of right and left traveling waves

$$\eta(x, t) = F_1(x - \sqrt{gH}t) + F_2(x + \sqrt{gH}t),$$

where F_1 and F_2 are arbitrary functions. Thus small-amplitude surface waves are governed by the wave equation, and they travel at speed \sqrt{gH} . It is straightforward to show that the velocity perturbation v also satisfies the wave equation

$$v_{tt} - gH v_{xx} = 0. \quad (4.1.12)$$

We emphasize that (4.1.11) and (4.1.12) are small-amplitude equations and hold only in the limit as the perturbations go to zero.

4.1.3 Gas Dynamics

The subject of gas dynamics is at the core of nonlinear PDEs, and it is difficult to overstate its role in the history and development of the subject. Many of the analytical and numerical methods in nonlinear analysis were spawned by investigations in fluid mechanics, and in particular, in aerodynamics and the flow of gases. Any serious student of PDEs must be well grounded in basic concepts of fluid flow.

The equations of gas dynamics consist of conservation laws and constitutive relations that define the properties of the gas. The conservation laws are nonlinear PDEs that express conservation of mass, momentum, and energy. For constitutive equations we consider the simplest case and assume an ideal gas law.

To fix the notion, we consider a tube of constant cross-sectional area A through which a gas is flowing in such a way that the physical variables are constant over any cross section. This assumption gives the motion its one-dimensional character. Further, we assume the gas is a continuum; that is, the physical parameters may be regarded as point functions of time t and a fixed, laboratory, spatial coordinate x in the longitudinal direction (Figure 4.3). Let $u = u(x, t)$ denote the velocity of a cross section of gas, $\rho = \rho(x, t)$ denote the density of the gas, and $p = p(x, t)$ denote the pressure. By convention, $p(x, t)A$ is the force on the cross section at x caused by the gas to the left of x . All physical quantities are measured in the fixed laboratory frame of reference; thus the velocity $u(x, t)$ is the velocity measured by an observer located at laboratory position x . This velocity is the velocity of different material as the gas passes by. Laboratory measurements of this type are called *Eulerian measurements*, and the laboratory coordinate x is called an *Eulerian coordinate*. There is an alternative description, the Lagrangian description, where all the physical

quantities are measured with respect to an observer moving with the gas (Lin & Segel 1974, Logan 2006a).

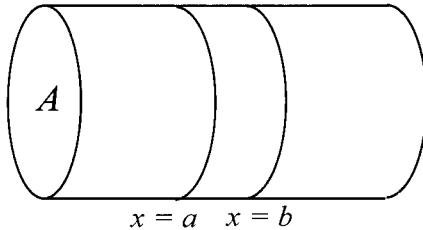


Figure 4.3 Cylindrical tube of cross-sectional area A containing a gas flowing in the axial direction.

In the section of the tube between $x = a$ and $x = b$, *mass conservation* requires that the time rate of change of the total mass inside $[a, b]$ equal the rate that mass flows in at $x = a$ minus the rate that mass flows out at $x = b$. Symbolically,

$$\frac{d}{dt} \int_a^b \rho(x, t) A \, dx = u(a, t)\rho(a, t)A - u(b, t)\rho(b, t)A, \quad (4.1.13)$$

which is the integral form of the mass conservation law. If u and ρ are smooth, we deduce the local form of the conservation law:

$$\rho_t + (\rho u)_x = 0. \quad (4.1.14)$$

The PDE (4.1.14) representing mass conservation is often called the *continuity equation*.

Momentum balance demands that the rate of change of momentum of the gas inside the region $[a, b]$ equal the rate that momentum flows into the region at $x = a$, minus the rate that momentum flows out of the region at $x = b$, plus the net force on the region (caused by the pressure at $x = a$ and $x = b$). We express this balance law mathematically as

$$\begin{aligned} \frac{d}{dt} \int_a^b \rho(x, t)u(x, t)A \, dx &= \rho(a, t)u(a, t)Au(a, t) - \rho(b, t)u(b, t)Au(b, t) \\ &\quad + p(a, t)A - p(b, t)A. \end{aligned} \quad (4.1.15)$$

Assuming smoothness, the integral form may be expressed in local form as

$$(\rho u)_t + (\rho u^2 + p)_x = 0, \quad (4.1.16)$$

An exercise shows that with the aid of (4.1.14), equation (4.1.16) can be expressed as

$$\rho(u_t + uu_x) + p_x = 0. \quad (4.1.17)$$

Equations (4.1.14) and (4.1.17), expressing mass and momentum balance, are quasilinear PDEs governing gas flow in the tube. However, there are three unknowns, u , ρ , and p . An additional equation is required to have a well determined system. We turn at this point to an equation of state that defines the material properties of the gas. The simplest assumption is to take a *barotropic equation of state*

$$p = f(\rho), \quad f', f'' > 0, \quad (4.1.18)$$

where f is a given function. For example, if there are no temperature changes, some gases can be modeled by the equation $p = k\rho^\gamma$, where k is a positive constant and $\gamma > 1$ (for air, $\gamma = 1.4$). Equations (4.1.14), (4.1.17), and (4.1.18) then form a well-determined system and govern *barotropic flow*.

In general, however, there are temperature changes in a system, and the pressure p depends on both the density ρ and the temperature $T = T(x, t)$. For example, an ideal gas satisfies the well-known equation

$$p = R\rho T, \quad (4.1.19)$$

where R is the universal gas constant. This equation of state introduces another unknown, the temperature T , and yet another equation is needed. This additional equation is the conservation of energy law, which must be considered when temperature changes occur.

We may write down an *energy balance* law in the same manner as for mass and momentum. There are two kinds of energy in a system, the kinetic energy of motion and the internal energy; the latter is due to molecular movement, and so on. We denote the specific internal energy function by $e = e(x, t)$, given in energy units per unit mass. Thus the total energy in the region $[a, b]$ is given by

$$\begin{aligned} \text{Total energy in } [a, b] &= \text{Kinetic energy of } [a, b] + \text{Internal energy inside } [a, b] \\ &= \int_a^b \frac{1}{2}\rho(x, t)u(x, t)^2 A dx + \int_a^b \rho(x, t)e(x, t)A dx. \end{aligned}$$

How can the total energy change? Energy, both kinetic and internal, can flow in and out of the region, and energy can change by doing work. It is the forces caused by the pressure that do the work, and the rate that a force does work is the force times the velocity. In other words, the rate of change of total energy in the region $[a, b]$ must equal the rate at which total energy flows into the region at $x = a$, minus the rate at which energy flows out of the region at $x = b$, plus

the rate at which the pressure force does work at the face $x = a$, minus the rate at which pressure does work at the face $x = b$. Expressed symbolically, this is

$$\begin{aligned} \frac{d}{dt} \int_a^b \rho \left(\frac{u^2}{2} + e \right) A dx &= \left[\frac{1}{2} \rho(a, t) u(a, t)^2 A + \rho(a, t) e(a, t) A \right] u(a, t) \\ &\quad - \left[\frac{1}{2} \rho(b, t) u(b, t)^2 A + \rho(b, t) e(b, t) A \right] u(b, t) \\ &\quad + p(a, t) u(a, t) A - p(b, t) u(b, t) A, \end{aligned} \quad (4.1.20)$$

which is the integral form of the conservation of energy law. Assuming smoothness, we conclude that

$$\left[\rho \left(\frac{u^2}{2} + e \right) \right]_t + \left[\rho u \left(\frac{u^2}{2} + e \right) + pu \right]_x = 0, \quad (4.1.21)$$

which is the *energy balance equation*.

Equations (4.1.14), (4.1.17), and (4.1.21) are the equations of gas dynamics, expressing conservation of mass and balance of momentum and energy. There are four unknowns: the density ρ , the velocity u , the pressure p , and the specific internal energy e . We point out that there are some physical effects that have been neglected in these equations. In the momentum equation we have assumed that there are no external forces present (e.g., gravity or an electromagnetic field) and that there are no viscous forces present; both these types of forces can change momentum. Flows without viscous effects are called *inviscid*. Furthermore, in the energy balance equation we neglected diffusion effects (heat transport) resulting from temperature gradients; if viscous forces were present, they would also do work and would have to be included as well.

Generally, equations (4.1.14), (4.1.17), and (4.1.21) are supplemented by two equations of state of the form

$$p = p(\rho, T) \quad (4.1.22)$$

and

$$e = e(\rho, T), \quad (4.1.23)$$

called a *thermal* equation of state and a *caloric* equation of state, respectively. Under these constitutive assumptions there are five equations in five unknowns: ρ , u , p , T , and e . A special case is the *ideal gas law*,

$$p = R\rho T, \quad e = c_v T, \quad (4.1.24)$$

where c_v is the constant specific heat at constant volume. Under the assumption (4.1.24) of an ideal gas, the energy equation (4.1.21) can be written in several

ways, depending on the choice of variables. For example, we leave it as an exercise to show that the energy equation can be written

$$\rho c_v(T_t + uT_x) + pu_x = 0. \quad (4.1.25)$$

In the same way that linearized equations for small disturbances are obtained from the fully nonlinear shallow-water equations, we can obtain linearized equations governing small-amplitude waves in gas dynamics. In this case the linearized theory is called *acoustics*, and the resulting equations govern ordinary sound waves in a gas. As in the case of shallow-water waves, the linearized equations of acoustics are linear wave equations (Exercise 5).

EXERCISES

1. Show that (4.1.7) follows from (4.1.6) and (4.1.5).
2. Derive (4.1.12).
3. Derive (4.1.25) under the ideal gas assumption (4.1.24).
4. In the case of an ideal gas, equations (4.1.24), show that the energy balance equation can be written

$$p_t + up_x - \frac{p}{\rho}(\rho_t + u\rho_x) + (\gamma - 1)pu_x = 0,$$

where $\gamma = 1 + R/c_v$.

5. (*Acoustics*) Consider the barotropic flow of a gas governed by (4.1.14), (4.1.17), and (4.1.18). Let $c^2 = f'(\rho)$, which is the square of the sound speed. Assume that the gas is in a constant equilibrium state $u = 0$, $\rho = \rho_0$, $p = p_0$, and let $c_0^2 = f'(\rho_0)$. Let

$$u = 0 + \tilde{u}(x, t), \quad \rho = \rho_0 + \tilde{\rho}(x, t),$$

where \tilde{u} is a small, velocity perturbation and $\tilde{\rho}$ is a small (compared to ρ_0) density perturbation, and derivatives of \tilde{u} and $\tilde{\rho}$ are small as well. Show that, under the assumption that quadratic terms in the perturbations may be ignored compared to linear terms, deviations \tilde{u} and $\tilde{\rho}$ from the equilibrium state each satisfy the wave equation

$$\psi_{tt} - c_0^2 \psi_{xx} = 0,$$

and therefore acoustic signals travel at speed c_0 .

4.2 Hyperbolic Systems and Characteristics

A single first-order PDE is essentially wave-like, and we identified a family of spacetime curves (characteristics) along which the PDE reduces to an ODE. Characteristics are spacetime loci along which signals are transmitted. Now we ask whether such curves exist for a *system* of first-order PDEs, and the answer leads to a general classification of systems.

To fix the idea and motivate the general definition and procedure, we analyze the shallow-water equations

$$h_t + u h_x + h u_x = 0, \quad (4.2.1)$$

$$u_t + u u_x + g h_x = 0. \quad (4.2.2)$$

The technique we apply is applicable to other systems. We search for a direction in spacetime in which both equations, simultaneously, contain directional derivatives of h and u in that direction. To begin, we take a linear combination of (4.2.1) and (4.2.2) to get

$$\gamma_1(h_t + u h_x + h u_x) + \gamma_2(u_t + u u_x + g h_x) = 0,$$

or, on rearranging terms to get the derivatives of h and of u together, we obtain

$$\gamma_1 \left[h_t + \left(u + \frac{g\gamma_2}{\gamma_1} \right) h_x \right] + \gamma_2 \left[u_t + \left(u + \frac{\gamma_1}{\gamma_2} h \right) u_x \right] = 0, \quad (4.2.3)$$

where γ_1 and γ_2 (not necessarily constant) are to be determined. From (4.2.3) the total derivative of h and the total derivative of u will be in the same direction if we force

$$\frac{g\gamma_2}{\gamma_1} = \frac{h\gamma_1}{\gamma_2},$$

or

$$\frac{\gamma_2}{\gamma_1} = \pm \sqrt{\frac{h}{g}}. \quad (4.2.4)$$

Therefore, take $\gamma_1 = 1$ and $\gamma_2 = \sqrt{h/g}$. Then (4.2.3) becomes

$$h_t + (u + \sqrt{gh})h_x + \sqrt{\frac{h}{g}}[u_t + (u + \sqrt{gh})u_x] = 0. \quad (4.2.5)$$

If $\gamma_1 = 1$ and $\gamma_2 = -\sqrt{h/g}$, then (4.2.3) becomes

$$h_t + (u - \sqrt{gh})h_x - \sqrt{\frac{h}{g}}[u_t + (u - \sqrt{gh})u_x] = 0. \quad (4.2.6)$$

In summary, we replaced the original shallow-water equations by alternate equations obtained by taking linear combinations, in which the derivatives of

h and u occur in the same direction, namely, $u + \sqrt{gh}$ in (4.2.5) and $u - \sqrt{gh}$ in (4.2.6). In other words, along curves defined by

$$\frac{dx}{dt} = u + \sqrt{gh}, \quad (4.2.7)$$

the PDE (4.2.5) reduces to

$$\frac{dh}{dt} + \frac{\sqrt{h}}{g} \frac{du}{dt} = 0, \quad (4.2.8)$$

which is an ODE. Also, along the curves defined by

$$\frac{dx}{dt} = u - \sqrt{gh}, \quad (4.2.9)$$

the PDE (4.2.6) reduces to

$$\frac{dh}{dt} - \sqrt{\frac{h}{g}} \frac{du}{dt} = 0. \quad (4.2.10)$$

Equations (4.2.7)–(4.2.10) become

$$\frac{dh}{dt} \pm \sqrt{\frac{h}{g}} \frac{du}{dt} = 0 \quad \text{along} \quad \frac{dx}{dt} = u \pm \sqrt{gh}. \quad (4.2.11)$$

Equations (4.2.11) are the *characteristic equations*, or characteristic form, corresponding to (4.2.1) and (4.2.2).

To review, for the two-PDE system (4.2.1) and (4.2.2), there are *two* families of characteristic curves defined by equations (4.2.7) and (4.2.9). Along these curves the PDEs reduce to ODEs.

4.2.1 Classification

The classification of a system of n first-order PDEs is based on whether there are n directions along which the PDEs reduce to n ODEs. To be more precise, assume that we are given a system of n equations in n unknowns u_1, \dots, u_n , which we write in matrix form as

$$\mathbf{u}_t + A(x, t, \mathbf{u})\mathbf{u}_x = \mathbf{b}(x, t, \mathbf{u}), \quad (4.2.12)$$

where $\mathbf{u} = (u_1, \dots, u_n)^t$, $\mathbf{b} = (b_1, \dots, b_n)^t$, and $A = (a_{ij}(x, t, \mathbf{u}))$ is an $n \times n$ matrix. The superscript t denotes the transpose operation, and boldface letters denote column vectors. In the subsequent analysis we prefer matrix notation

rather than the more cumbersome index notation. For example, in matrix form the shallow-water equations are,

$$\begin{pmatrix} h \\ u \end{pmatrix}_t + \begin{pmatrix} u & h \\ g & u \end{pmatrix} \begin{pmatrix} h \\ u \end{pmatrix}_x = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

In this case $\mathbf{u} = (h, u)^t$, $\mathbf{b} = (0, 0)^t$, and the matrix A is

$$A = \begin{pmatrix} u & h \\ g & u \end{pmatrix}. \quad (4.2.13)$$

Now we ask whether there is a family of curves along which the PDEs (4.2.12) reduce to a system of ODEs, that is, in which the directional derivative of each u_i occurs in the same direction. We proceed as in the example of the shallow-water equations and take a linear combination of the n equations in (4.2.12). This amounts to multiplying (4.2.12) on the left by a row vector $\gamma^t = (\gamma_1, \dots, \gamma_n)$, which is to be determined. Thus

$$\gamma^t \mathbf{u}_t + \gamma^t A \mathbf{u}_x = \gamma^t \mathbf{b}. \quad (4.2.14)$$

We want (4.2.14) to have the form of a linear combination of total derivatives of the u_i in the same direction λ ; that is, we want (4.2.14) to have the form

$$\mathbf{m}^t (\mathbf{u}_t + \lambda \mathbf{u}_x) = \gamma^t \mathbf{b} \quad (4.2.15)$$

for some \mathbf{m} . Consequently, we require

$$\mathbf{m} = \gamma, \quad \mathbf{m}^t \lambda = \gamma^t A,$$

or

$$\gamma^t A = \lambda \gamma^t. \quad (4.2.16)$$

This means that λ is an eigenvalue of A and γ^t is a corresponding left eigenvector. Note that λ as well as γ can depend on x , t , and \mathbf{u} . So, if (λ, γ^t) is an eigenpair, then

$$\gamma^t \frac{d\mathbf{u}}{dt} = \gamma^t \mathbf{b} \quad \text{along} \quad \frac{dx}{dt} = \lambda(x, t, \mathbf{u}), \quad (4.2.17)$$

and the system of PDEs (4.2.12) is reduced to a single ODE along the family of curves, called *characteristics*, defined by $dx/dt = \lambda$. The eigenvalue λ is called the *characteristic direction*. Clearly, because there are n unknowns, it would appear that n ODEs are required; but if A has n distinct real eigenvalues, there are n ordinary differential equations, each holding along a characteristic direction defined by an eigenvalue. In this case we say that the system is hyperbolic. We record the following definition.

Definition. The quasilinear system (4.2.12) is *hyperbolic* if A has n real eigenvalues and n linearly independent left eigenvectors. If (4.2.12) is hyperbolic

and the eigenvalues are distinct, it is *strictly hyperbolic*. The system (4.2.12) is *elliptic* if A has no real eigenvalues, and it is *parabolic* if A has n real eigenvalues but fewer than n independent left eigenvectors. In the hyperbolic case, the system of n equations (4.2.17), obtained by selecting each of the n distinct eigenpairs, is called the *characteristic form* of (4.2.12).

No exhaustive classification is made in the case that A has both real and complex eigenvalues. Recall that if a matrix has n distinct, real eigenvalues, it has n independent left eigenvectors, because distinct eigenvalues have independent eigenvectors. It is, of course, possible for a matrix to have fewer than n distinct eigenvalues (multiplicities can occur), yet have a full complement of n independent eigenvectors. Finally, we recall that the eigenvalues can be calculated directly from the equation

$$\det(A - \lambda I) = 0,$$

a condition that follows immediately from the stipulation that the homogeneous linear system of equations (4.2.16) must have nontrivial solutions.

More general systems of the form

$$B(x, t, \mathbf{u})\mathbf{u}_t + A(x, t, \mathbf{u})\mathbf{u}_x = \mathbf{b}(x, t, \mathbf{u}) \quad (4.2.18)$$

can be considered as well. If B is nonsingular, (4.2.18) can be transformed into a system of the form (4.2.12) by multiplying through by B inverse. However, let us follow through with the analysis of (4.2.18) as it stands. We ask whether there is a vector γ such that the equation

$$\gamma^t(B\mathbf{u}_t + A\mathbf{u}_x) = \gamma^t\mathbf{b} \quad (4.2.19)$$

can take the form

$$\mathbf{m}^t(\beta\mathbf{u}_t + \alpha\mathbf{u}_x) = \gamma^t\mathbf{b}, \quad (4.2.20)$$

where α and β are scalar functions. If so, there is a single direction (α, β) where the directional derivative of each u_i in the equation is in that same direction. To elaborate, let $x = x(s), y = y(s)$ be a curve Γ such that $dx/ds = \alpha$ and $dy/ds = \beta$. Along this curve $d\mathbf{u}/ds = \alpha\mathbf{u}_x + \beta\mathbf{u}_t$, so (4.2.20) may be written

$$\mathbf{m}^t \frac{d\mathbf{u}}{ds} = \gamma^t\mathbf{b}, \quad (4.2.21)$$

which is a differential equation along Γ . Now, the conditions that (4.2.19) and (4.2.20) are equivalent are

$$\gamma^t B = \mathbf{m}^t \beta, \quad \gamma^t A = \mathbf{m}^t \alpha, \quad (4.2.22)$$

or, on eliminating \mathbf{m} , we obtain

$$\gamma^t(B\alpha - A\beta) = \mathbf{0}. \quad (4.2.23)$$

A necessary and sufficient condition that the homogeneous equation (4.2.23) has a nontrivial solution γ^t is

$$\det(B\alpha - A\beta) = 0. \quad (4.2.24)$$

Equation (4.2.24) is a condition on the direction (α, β) of the curve Γ . Such a curve is said to be *characteristic*, and (4.2.21) is said to be in *characteristic form*. As before, we say that a system (4.2.18) satisfying (4.2.24) is *hyperbolic* if the linear algebraic system (4.2.23) has n linearly independent solutions γ^t , where the directions (α, β) are real and not both α and β are zero. The directions need not be distinct. It may be the case that either B or A is singular, but we assume that A and B are not both singular so that the system becomes degenerate. If $\det B = 0$, then $\beta = 0$ is a solution of (4.2.24) for any α , and therefore the x direction is characteristic (i.e., the horizontal lines $t = \text{const}$ are characteristics); similarly, if $\det A = 0$, the vertical coordinate lines $x = \text{const}$ are characteristics. If no real directions (α, β) exist, the system (4.2.9) is called *elliptic*; the system (4.2.18) is called *parabolic* if there are n real directions defined by (4.2.24) but fewer than n linearly independent solutions of (4.2.23).

Example. For the shallow-water equations (4.2.1)–(4.2.2) the coefficient matrix A given by (4.2.13). The eigenvalues of (4.2.13) are found from $\det(A - \lambda I) = 0$, or

$$\det \begin{pmatrix} u - \lambda & h \\ g & u - \lambda \end{pmatrix} = (u - \lambda)^2 - gh = 0.$$

The eigenvalues are $\lambda = u \pm \sqrt{gh}$, which are real and distinct. Thus the shallow-water equations are strictly hyperbolic. Notice that the eigenvalues define the characteristic directions (cf. (4.2.11)). The eigenvectors are $(1, \sqrt{h/g})$ and $(1, -\sqrt{h/g})$, which define the appropriate linear combination of the PDEs. \square

Now we see why a single first-order quasilinear PDE is hyperbolic. The coefficient matrix for the equation

$$u_t + c(x, t, u)u_x = b(x, t, u)$$

is just the real scalar function $c(x, t, u)$, which has the single eigenvalue $c(x, t, u)$. In this direction, that is, if $dx/dt = c(x, t, u)$, the PDE reduces to $du/dt = b(x, t, u)$. \square

Example. Consider the system

$$\begin{aligned} u_t - v_x &= 0, \\ v_t - cu_x &= 0, \end{aligned} \tag{4.2.25}$$

where c is a constant. The equivalent matrix form is

$$\begin{pmatrix} u \\ v \end{pmatrix}_t + \begin{pmatrix} 0 & -1 \\ -c & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}_x = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

The coefficient matrix

$$\begin{pmatrix} 0 & -1 \\ -c & 0 \end{pmatrix}$$

has eigenvalues $\lambda = \pm\sqrt{c}$. Therefore, if $c > 0$, the system is hyperbolic, and if $c < 0$, the system is elliptic. It is straightforward to eliminate the unknown v from the two equations (4.2.25) and obtain the single second-order PDE

$$u_{tt} - cu_{xx} = 0. \tag{4.2.26}$$

If $c > 0$, we recognize (4.2.26) as the wave equation, which is hyperbolic. If $c < 0$, then (4.2.26) is elliptic [if $c = -1$, then (4.2.26) is Laplace's equation]. Therefore, the classification scheme for systems is consistent with the classification in Chapter 1 for second-order linear equations. For the system (4.2.25) in the hyperbolic case ($c > 0$), the characteristic directions are $\pm\sqrt{c}$, and the characteristic curves are

$$\frac{dx}{dt} = \pm\sqrt{c} \quad \text{or} \quad x = \pm\sqrt{c}t + \text{const.}$$

These two families of characteristics are straight lines of speed \sqrt{c} and $-\sqrt{c}$. To find the characteristic form of (4.2.25), we proceed as in the example of the shallow-water equations and take a linear combination of the two equations to obtain

$$\gamma_1(u_t - v_x) + \gamma_2(v_t - cu_x) = 0.$$

Rearranging to put the derivatives of u and the derivatives of v in different terms, we get

$$\gamma_1 \left(u_t - \frac{c\gamma_2}{\gamma_1} u_x \right) + \gamma_2 \left(v_t - \frac{\gamma_1}{\gamma_2} v_x \right) = 0.$$

Therefore the derivatives of u and v are in the same direction if $c\gamma_2/\gamma_1 = \gamma_1/\gamma_2$, or

$$\frac{\gamma_1}{\gamma_2} = \pm\sqrt{c}.$$

Thus we take $\gamma_1 = \pm\sqrt{c}$ and $\gamma_2 = 1$. Consequently, the characteristic form of the PDEs (4.2.25) is

$$\pm\sqrt{c}(u_t \mp \sqrt{c}u_x) + (v_t \mp \sqrt{c}v_x) = 0.$$

We can write

$$\pm\sqrt{c}\frac{du}{dt} + \frac{dv}{dt} = 0 \quad \text{along} \quad \frac{dx}{dt} = \mp\sqrt{c},$$

or

$$\pm\sqrt{c}u + v = \text{const} \quad \text{along} \quad x = \mp\sqrt{c}t + \text{const}.$$

The expressions $\pm\sqrt{c}u + v$ are called *Riemann invariants*,¹ they are expressions that are constant along the characteristic curves. It follows immediately that

$$\sqrt{c}u + v = f(x + \sqrt{c}t), \quad -\sqrt{c}u + v = g(x - \sqrt{c}t),$$

where f and g are arbitrary functions. We can obtain formulas for u and v by solving these equations simultaneously. Thus we obtained the general solution of (4.2.25) in terms of arbitrary functions; the latter are determined from initial or boundary data. \square

There are basically two ways to find the characteristic form of a system of hyperbolic equations. We can calculate the eigenvalues and eigenvectors directly and then use (4.2.17); or we can proceed directly and take linear combinations of the equations as we did for the shallow-water equations at the beginning of this section and for the wave equation. For a small number of equations ($n \leq 3$) the latter method may be preferable. Additional examples are given in the Exercises.

In summary, a system of hyperbolic equations can be transformed a system of equations where each equation in the new system involves directional derivatives in only one direction, that is, in the direction of the characteristics. This transformed set of characteristic equations reduce to ODEs along the characteristic curves. Often, they can be solved and used to obtain the solution to the original system. This method is called the *method of characteristics*, and it is fundamental in the analysis of hyperbolic problems.

Example. The characteristic form of the shallow-water equations (4.2.1)–(4.2.2) is given by (4.2.11). We can write (4.2.11) as

$$\sqrt{\frac{g}{h}}\frac{dh}{dt} \pm \frac{du}{dt} = 0 \quad \text{on} \quad \frac{dx}{dt} = u \pm \sqrt{gh}.$$

The differential equations on the characteristic curves may be integrated directly to obtain

$$R_{\pm} = 2\sqrt{gh} \pm u = \text{const} \quad \text{on} \quad \frac{dx}{dt} = u \pm \sqrt{gh}.$$

¹ B. Riemann, who is most known for his work in analysis, was one of the early investigators of wave phenomena and he pioneered some of the techniques and methods for nonlinear equations and shock waves, particularly in the area of fluid dynamics.

The quantities R_{\pm} are *Riemann invariants*, and they are constant along characteristics; R_+ is constant along a characteristic of the family $dx/dt = u + \sqrt{gh}$, and R_- is constant along a characteristic of the family $dx/dt = u - \sqrt{gh}$. The constant may vary from one characteristic to another. In the shallow-water equations, the characteristics with direction $u + \sqrt{gh}$ are called the *positive characteristics* (or C^+ characteristics), and the characteristics with direction $u - \sqrt{gh}$ are called the *negative characteristics* (or C^- characteristics). Consequently, as for a single PDE, characteristics carry signals or information; the C^+ characteristics carry the constancy of R_+ , and the C^- characteristics carry the constancy of R_- . Armed with this fact, we are able to solve the shallow-water equations for some types of initial and boundary data. This is discussed in Section 4.3.

Example. (Diffusion Equation) The diffusion equation $u_t = u_{xx}$ may be written as the first-order system

$$u_t = v_x, \quad u_x = v, \quad (4.2.27)$$

which, in matrix form, is

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}_t + \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}_x = \begin{pmatrix} 0 \\ v \end{pmatrix}.$$

This is in the form (4.2.18) with $\det B = 0$. Here, $\det(B\alpha - A\beta) = \beta^2 = 0$ forces $\beta = 0$, so the characteristic direction $(\alpha, 0)$ is in the direction of the x -axis, confirming previous observations. The eigenvalue equation $\gamma^t(B\alpha - A\beta) = \mathbf{0}$ becomes

$$(\gamma_1 \quad \gamma_2) \begin{pmatrix} \alpha & 0 \\ 0 & 0 \end{pmatrix} = (0 \quad 0),$$

which has only one solution, namely, $\gamma^t = (0, c)$, where c is a constant. Therefore, the system (4.2.27) is parabolic, consistent with the classification of the second-order diffusion equation. If we interpret the characteristics, here horizontal straight lines with speed infinity, as curves in spacetime that carry signals, then the diffusion equation transmits signals at infinite speed. Recall that the diffusion equation with an initial point source (zero everywhere except at a single point) has a solution that is nonzero for all real x for $t > 0$. Thus signals travel infinitely fast. \square

EXERCISES

1. The electrical current $i = i(x, t)$ and voltage $v = v(x, t)$ at a position x at time t along a transmission line satisfy the first-order system

$$Cv_t + i_x = -Gv, \quad Li_t + v_x = -Ri,$$

where C , G , L , and R are positive constants denoting the capacitance, leakage, inductance, and resistance, respectively, all per unit length in the line. Show that this system is hyperbolic and write it in characteristic form. Sketch the two families of characteristics on an xt diagram.

2. (*Gas Dynamics*). The equations governing barotropic flow are

$$\rho_t + (\rho u)_x = 0, \quad (\rho u)_t + (\rho u^2 + p)_x = 0, \quad p = f(\rho),$$

where $f', f'' > 0$.

- (a) Show that these equations may be written

$$p_t + up_x + c^2 \rho u_x = 0, \quad \rho u_t + \rho uu_x + p_x = 0,$$

where $c^2 = f'(\rho)$.

- (b) Derive the characteristic form for the equations in part (a). You should get

$$\frac{dp}{dt} \pm \rho c \frac{du}{dt} = 0 \quad \text{on} \quad C^\pm : \frac{dx}{dt} = u \pm c.$$

- (c) Prove that

$$R_\pm = \int \frac{c(\rho)}{\rho} d\rho \pm u = \text{constant} \quad \text{on} \quad C^\pm.$$

- (d) Determine the Riemann invariants R_\pm in part (c) assuming that the equation of state is $f(\rho) = k\rho^\gamma$, where $k > 0$ and $\gamma > 1$ are constants.

Solution:

$$R_\pm = \frac{2c}{\gamma - 1} \pm u.$$

3. Consider the linear nonhomogeneous system

$$\begin{pmatrix} -1 & a^2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}_t + \begin{pmatrix} u \\ v \end{pmatrix}_x = \begin{pmatrix} f(t) \\ g(t) \end{pmatrix}$$

on $x > 0, t > 0$, where $0 < a < 1$, and where f and g are given continuous functions.

- (a) Show that the system is hyperbolic and find the characteristic form of the equations.
- (b) Sketch the characteristics on an xt diagram.
- (c) Determine expressions that are constants on the characteristics.

(d) To the system of PDEs append the initial and boundary conditions

$$u(0, t) = 0, \quad t > 0,$$

$$u(x, 0) = \alpha v(x, 0), \quad x > 0, \text{ where } \alpha > 0.$$

For any $t > 0$, prove that the boundary condition on v is given by

$$v(0, t) = \delta v(0, rt) + \frac{\delta[F(rt) + aG(rt)]}{a(a-1)} + \frac{F(t) - G(t)}{a(1-a)},$$

where F and G are the antiderivatives of f and g , and

$$r = \frac{1-a}{1+a}, \quad \delta = \frac{a-\alpha}{a+\alpha}.$$

(Logan 1989).

4. Consider the linear, strictly hyperbolic system

$$\mathbf{u}_t + A(x, t)\mathbf{u}_x = \mathbf{0}.$$

Show that the transformation $\mathbf{u} = P\mathbf{w}$, where $P^{-1}AP = D$, and D is a diagonal matrix having the eigenvalues of A on its diagonal, transforms the system into a linear system

$$\mathbf{w}_t + D(x, t)\mathbf{w}_x = C(x, t)\mathbf{w}$$

with the derivatives of the components of \mathbf{w} decoupled. What are the characteristic directions? Why does such a matrix P exist? Write the system in characteristic form.

5. Solve the linear initial value problem

$$u_t + v_t + 3u_x + 2v_x = 0,$$

$$-u_t + v_t + 5u_x + 2v_x = 0, \quad x \in \mathbb{R}, \quad t > 0,$$

$$u(x, 0) = \sin x, \quad v(x, 0) = e^x, \quad x \in \mathbb{R}.$$

Sketch the characteristics in the xt plane. *Solution:* $u = \sin(x+t)$, $v = [8\sin(x-2t) + 6\exp(x-2t) - 8\sin(x+t)]/6$.

6. Consider the system

$$u_t + 4u_x - 6v_x = 0,$$

$$v_t + u_x - 3v_x = 0.$$

(a) Show that the system is strictly hyperbolic.

(b) Transform the system to characteristic form by finding the eigenvalues and left eigenvectors.

- (c) Find the general solution in terms of arbitrary functions.
- (d) Determine the characteristics and Riemann invariants.
7. A simple analog of a detonation is given by model equations

$$u_t + f(u, z)_x = 0, \quad z_t = -r(u, z),$$

where $u = u(x, t)$ is a temperature-like quantity, and $z = z(x, t)$ is the mass fraction of the reactant **A** in an exothermic, irreversible, chemical reaction **A** \rightarrow **B** (i.e., z is the mass of **A** divided by the sum of the masses of **A** and **B**); $f(u, z)$ defines an equation of state and $r(u, z)$ defines the chemical reaction rate. Assume that $f_u, f_{uu} > 0$ and $f_z < 0$. These equations represent an idealized model chemical-fluid mechanic interactions in a chemically reactive medium. Show that this system is hyperbolic, find the characteristic directions, and transform the system to characteristic form.

8. Consider the hyperbolic system

$$\begin{aligned} u_t + uv_x &= u\sqrt{v}, \\ v_t + vu_x &= v\sqrt{u}, \end{aligned}$$

and $u, v > 0$. Show that the characteristic directions are $\pm 1/\sqrt{uv}$, and the left eigenvectors are $(\sqrt{v}, \pm\sqrt{u})$. On the positive characteristic show that

$$\sqrt{u} + \sqrt{v} - x = \text{const.}$$

4.3 The Riemann Method

Using the shallow water equations as the model, we illustrate how to solve a variety of initial-boundary value problems by the method of characteristics, or *Riemann's method*. As we observed in Section 4.2, under some circumstances it is possible to determine expressions, called *Riemann invariants*, that are constant along the characteristic curves. Knowing these invariants and relating them back to the initial and boundary conditions permits the determination of their values on the characteristics; this information then points the way toward the solution of the problem.

4.3.1 Jump Conditions for Systems

Before proceeding with representative problems, we pause to consider shock relations for a system of hyperbolic equations in conservation form. The situation is much the same as in the single-equation case discussed in Section 3.3.

We assume that the governing equations governing are integral conservation laws of the form

$$\frac{d}{dt} \int_a^b \Psi_k dx = -\Phi_k|_{x=a}^{x=b} + \int_a^b f_k dx, \quad (k = 1, \dots, n), \quad (4.3.1)$$

where the Ψ_k , Φ_k , and f_k are physical quantities depending on x , t , and n unknown functions, or densities, $u_1(x, t), \dots, u_n(x, t)$. The Φ_k are flux terms and the f_k are local source terms. Under the hypothesis that the u_k are continuously differentiable, we can proceed in the usual way to obtain a system of PDEs of the form

$$(\Psi_k(x, t, \mathbf{u}))_t + (\Phi_k(x, t, \mathbf{u}))_x = f_k(x, t, \mathbf{u}) \quad (k = 1, \dots, n), \quad (4.3.2)$$

where $\mathbf{u} = (u_1, \dots, u_n)$. We say that the system (4.3.2) is in conservation form. If the functions u_k have simple discontinuities along a smooth curve $x = s(t)$ in spacetime, the same argument as in Section 3.3 can be applied to obtain the *n jump conditions*

$$-s'[\Psi_k] + [\Phi_k] = 0 \quad (k = 1, \dots, n), \quad (4.3.3)$$

where $s' = ds/dt$, and where $[Q] = Q^- - Q^+$ denotes the jump in the quantity Q across $x = s(t)$. The curve $x = s(t)$ is the *shock path*, s' is the *shock velocity*, and the waveforms u_k themselves, as their discontinuities propagate along $x = s(t)$, are collectively called a *shock wave*. The same proviso holds as for a single scalar, conservation law, namely, the correct jump conditions can be obtained from integral forms of the conservation laws because there is no unique way to obtain PDEs in conservation form from a system of arbitrary PDEs.

Example. The shallow-water equations derived in Section 4.1 were obtained from integral conservation laws, which led, in turn, to a system of PDEs in conservation form

$$h_t + (hu)_x = 0, \quad (4.3.4)$$

$$(hu)_t + (hu^2 + \frac{1}{2}h^2)_x = 0, \quad (4.3.5)$$

where we have set $g = 1$. (The fact that the shallow-water equations can be rescaled to set the acceleration due to gravity equal to unity is explored in Exercise 1.) Thus, according to (4.3.3), the jump conditions are

$$s' = \frac{[hu]}{[h]} \quad \text{and} \quad s' = \frac{[hu^2 + h^2/2]}{[hu]}. \quad (4.3.6)$$

We regard (4.3.6) as two equations relating five quantities consisting of the values of h and of u ahead of and behind the shock, and the shock velocity s' . In the context of shallow-water theory, a propagating shock is called a *bore*.

□

4.3.2 Breaking Dam Problem

We imagine water at height 1 held motionless in $x < 0$ by a dam placed at $x = 0$. Ahead of the dam ($x > 0$) there is no water. At time $t = 0$ the dam is suddenly removed, and the problem is to determine the height and velocity of the water for all $t > 0$ and $x \in \mathbb{R}$, according to the shallow-water theory. Consequently, the governing equations are (4.3.4) and (4.3.5) subject to the initial conditions

$$\begin{aligned} u(x, 0) &= 0 \quad \text{for } x \in \mathbb{R}; \quad h(x, 0) = 1 \quad \text{for } x < 0, \\ h(x, 0) &= 0 \quad \text{for } x > 0. \end{aligned} \quad (4.3.7)$$

The solution of this problem, using Riemann's method, is a blend of physical intuition and analytic calculations, the former guiding the latter. We prefer this approach, which interrelates the physics and the mathematics, rather than purely analytic reasoning that ignores the origins of the problem. From Section 4.2 we know that the shallow-water equations (4.3.4) and (4.3.5) can be put in the characteristic form

$$\frac{dh}{dt} \pm \sqrt{h} \frac{du}{dt} = 0 \quad \text{on} \quad C^\pm : \frac{dx}{dt} = u \pm \sqrt{h}. \quad (4.3.8)$$

Direct integration gives

$$R^\pm = 2\sqrt{h} \pm u = \text{constant} \quad \text{on} \quad C^\pm : \frac{dx}{dt} = u \pm \sqrt{h}, \quad (4.3.9)$$

where R^+ and R^- are the two Riemann invariants, which are expressions that are constant on the C^+ and C^- characteristics, respectively. We first examine the C^- curves emanating from the negative x axis, where the water is located. The speed of the negative characteristics is $u - \sqrt{h}$, and therefore the C^- curves leave the negative x axis (where $u = 0$ and $h = 1$) with speed -1 . Along a C^- characteristic we have

$$R^- = 2\sqrt{h} - u = 2\sqrt{h(x, 0)} - u(x, 0) = 2. \quad (4.3.10)$$

Similarly, a C^+ characteristic emanating from the negative x axis has initial speed $+1$, and on such a characteristic we have

$$R^+ = 2\sqrt{h} + u = 2. \quad (4.3.11)$$

Therefore, at a point P in the region $x < -t$, both (4.3.10) and (4.3.11) hold, giving $u = 0$ and $h = 1$ in the domain $x < -t$. Because h and u are constant in this entire region, the speeds of the C^+ and C^- characteristics are constants ($+1$ and -1 , respectively), so these characteristics are straight lines in the region $x < -t$. This portion of the characteristic diagram is shown in Figure 4.4.

The region $x < -t$ is the region of spacetime that cannot be affected by the removal of the dam. The first signal back into the water is the leading edge of rarefaction wave that releases the lower water height ahead of the wave. This signal travels along $x = -t$ at speed -1 .

We also expect the water to move into $x > 0$ as a bore, subject to the jump conditions (4.3.6). Ahead of the bore (see Figure 4.4) we have $u = h = 0$, and therefore the jump conditions (4.3.6) become

$$s' = u_-, \quad s' = \frac{u_-^2 + h_-/2}{u_-}, \quad (4.3.12)$$

where h_- and u_- are the values of h and u just behind the bore. It follows immediately from (4.3.12) that $h_- = 0$; that is, there is no jump in height across the bore. To determine u_- , and hence the bore velocity s' , we need to know the solution $u(x, t)$ in the triangular region behind the bore and in front of the leading edge of the rarefaction $x = -t$.

One way to proceed is to observe that the C^+ characteristics carry information forward in time; they originate on the negative x axis and end on the bore. Because equation (4.3.11) holds on *every* C^+ characteristic, it must in fact hold *everywhere* behind the bore (this is because the C^+ come from a constant state, $u = 0$ and $h = 1$, in this problem; if the initial state were not constant, this conclusion could not be made). Hence, (4.3.11) must hold just behind the bore, or

$$2\sqrt{h_-} + u_- = 2.$$

Consequently, because $h_- = 0$, we must have $u_- = 2$ and $s' = 2$. We conclude that the bore is a straight line $x = 2t$; the jump in the height h across the bore is zero, and the jump in the velocity u is 2.

To find the solution $u(x, t)$ and $h(x, t)$ in the region behind the bore we expect, from earlier experience with rarefactions in traffic flow, to fit in a fan of

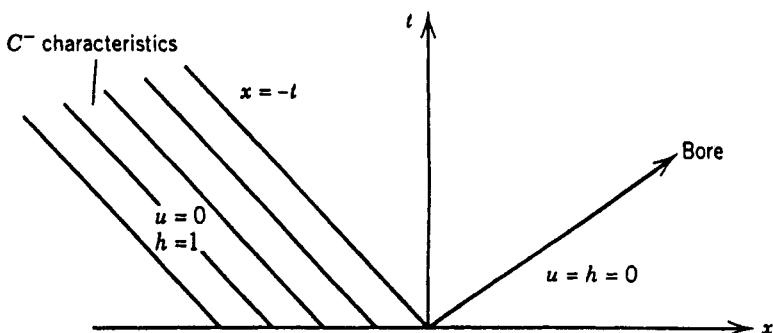


Figure 4.4 Negative (C^-) characteristics for the breaking dam problem

C^- characteristics connecting the solution along $x = -t$ to the solution along $x = 2t$. Therefore, we insert straight-line C^- characteristics into this region, passing through the origin, with equations

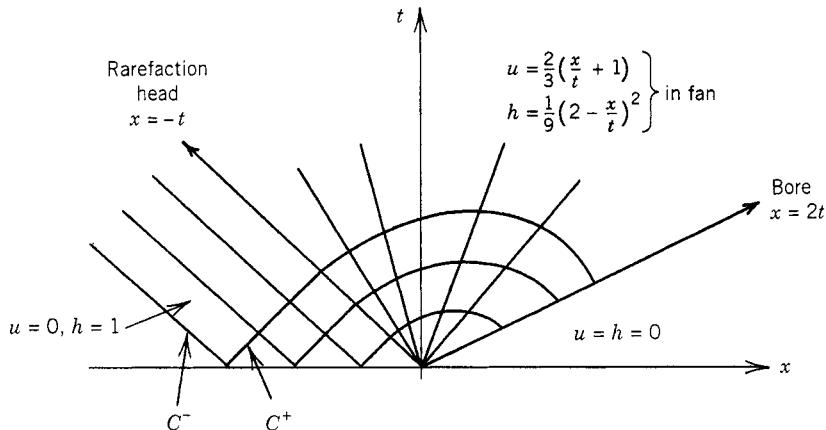


Figure 4.5 Characteristic diagram for the breaking dam problem

$$x = (u - \sqrt{h})t, \quad (4.3.13)$$

where $u - \sqrt{h}$ is the speed of a C^- characteristic. Then, in this triangular region, we obtain

$$u = \frac{x}{t} + \sqrt{h}. \quad (4.3.14)$$

Using (4.3.14) along with (4.3.11), where the latter holds at all points of the region, gives

$$u(x, t) = \frac{2(x/t + 1)}{3}, \quad h(x, t) = \frac{(2 - x/t)^2}{9} \quad (-t < x < 2t). \quad (4.3.15)$$

We already know that $u = h = 0$ for $x > 2t$, and $u = 0, h = 1$ for $x < -t$. Thus we have a complete solution to the problem. Figure 4.5 shows a complete characteristic diagram, and Figure 4.6 shows typical time snapshots of the velocity and height.

4.3.3 Receding Wall Problem

Imagine water in $x > 0$ being held motionless at height $h = 1$ by a wall at $x = 0$. At time $t = 0$ the wall is pulled back along a given spacetime path

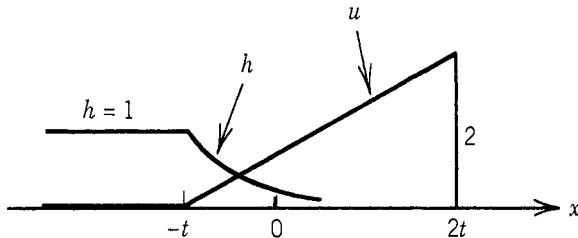


Figure 4.6 Velocity and water height profiles for the dam breaking problem

$x = X(t)$, $t > 0$, where X satisfies the conditions $X'(0) = 0$, $X'(t) < 0$, and $X''(t) < 0$. Thus the wall has negative velocity and continues to accelerate backward as time increases. The problem is to determine the velocity and height of the water for $X(t) < x$ and $t > 0$. Shallow-water theory is assumed to model the evolution of the system. The initial and boundary conditions are expressed analytically by

$$u(x, 0) = 0, \quad h(x, 0) = 1 \quad \text{for } x > 0, \quad (4.3.16)$$

and

$$u(X(t), t) = X'(t), \quad t > 0. \quad (4.3.17)$$

Equation (4.3.17) expresses the fact that the velocity of the water adjacent to the wall is the same as the velocity of the wall. As it turns out, we cannot independently impose a condition on height of the water at the wall; intuitively, we should be able to determine h at the wall as a part of the solution to the problem.

Again we resort to the characteristic form of the shallow-water equations:

$$R^+ = 2\sqrt{h} + u = \text{const} \quad \text{on} \quad C^+ : \frac{dx}{dt} = u + \sqrt{h}, \quad (4.3.18)$$

$$R^- = 2\sqrt{h} - u = \text{const} \quad \text{on} \quad C^- : \frac{dx}{dt} = u - \sqrt{h}. \quad (4.3.19)$$

There is enough information in the characteristic form of the equations to solve the problem. The idea is to let the C^+ and C^- characteristics carry the constancy of R^+ and R^- from the boundaries into the region of interest. First we consider the C^- characteristics emanating from the positive x axis. Because $u = 0$ and $h = 1$ along the x axis, the C^- characteristics leave the x axis with speed -1 . Along each of these characteristics R^- has the same constant value:

$$2\sqrt{h} - u = 2. \quad (4.3.20)$$

However, because (4.3.20) holds along every C^- characteristic, *it must hold everywhere* in $x > X(t)$; that is, the negative Riemann invariant is constant

throughout the region of interest. Because the speed of the C^- characteristics is more negative than the speed of the wall along the path of the wall, the C^- curves must intersect the wall path as shown in Figure 4.7. Equation (4.3.20) permits us to calculate the water height h at the wall, because u at the wall is known. Thus the height of the water at the wall is given by

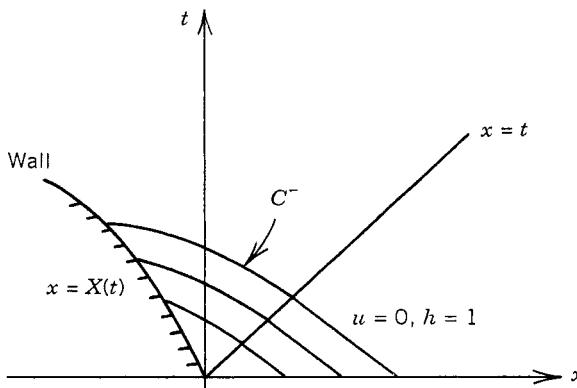


Figure 4.7 Negative characteristics for the retracting wall problem.

$$h(X(t), t) = \left(\frac{2 + X'(t)}{2} \right)^2. \quad (4.3.21)$$

The C^+ characteristics emanating from the x axis have initial speed 1, and it is easy to see that in the region $x > t$ we have the constant state $u = 0$ and $h = 1$; this is the region of spacetime not influenced by the motion of the wall. The leading signal into the water ahead travels at speed 1 along the limiting C^+ characteristic $x = t$. Because both u and h are constant ahead of this initial signal, both sets of characteristics are straight lines in this region. Now consider a C^+ characteristic emanating from the wall, as shown in Figure 4.8. It is straightforward to see that C^+ must be a straight line; along a C^+ we have

$$2\sqrt{h} + u = \text{const.} \quad (4.3.22)$$

The constant, of course, will vary from one C^+ to another because u is changing along the wall. Adding and subtracting (4.3.22) and (4.3.20), the latter holding everywhere, shows that on a specific C^+ characteristic both u and h are constant. Hence, the speed $u + \sqrt{h}$ of a specific C^+ must be constant, and the characteristic is therefore a straight line. The C^+ characteristic emanating

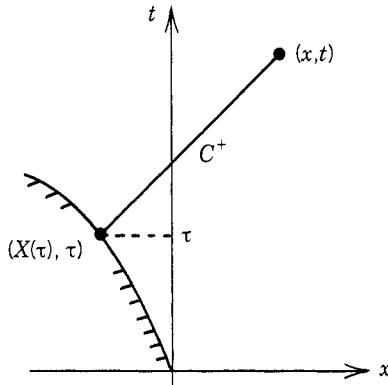


Figure 4.8 C^+ characteristic emanating from the retracting wall.

from $(X(\tau), \tau)$ is expressed by the equation

$$\begin{aligned} x - X(\tau) &= (u + \sqrt{h})(t - \tau) \\ &= \left[1 + \frac{3X'(\tau)}{2} \right] (t - \tau), \end{aligned} \quad (4.3.23)$$

where we used (4.3.21) and (4.3.17), and applied the fact that u and h are constant on the characteristic.

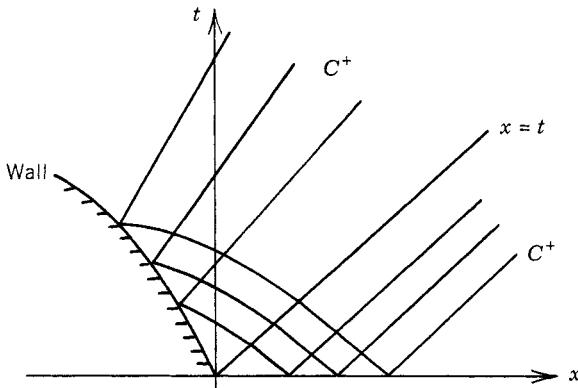


Figure 4.9 Characteristic diagram for the retracting wall problem.

Finally, we may determine the solution h and u at an arbitrary point (x, t) in the region $X(t) < x < t$. Again using the constancy of h and u on a C^+ , we have

$$u(x, t) = u(X(\tau), \tau) = X'(\tau) \quad (4.3.24)$$

and

$$h(x, t) = h(X(\tau), \tau) = \left[\frac{2 + X'(\tau)}{2} \right]^2, \quad (4.3.25)$$

where $\tau = \tau(x, t)$ is given implicitly by (4.3.23). The characteristic diagram is shown in Figure 4.9.

In the two preceding problems one of the Riemann invariants is constant throughout the region of interest. Whenever this occurs, we call the solution in the nonuniform region a *simple wave*. One can show that a simple wave solution always exists adjacent to a uniform state provided that the solution remains smooth. In the two examples we considered, the constancy of one of the Riemann invariants occurred because all of the characteristics originated from a constant state; if the initial conditions are not constant, the simple wave argument cannot be made.

4.3.4 Formation of a Bore

We can examine the preceding problem in the case where the wall is pushed forward instead of being retracted. If we draw a typical wall path in this case, it is easy to see that two C^+ characteristics will collide, indicating the formation of a shock wave or bore (see Figure 4.10). Assume that the wall path is $x = X(t)$, where $X'(t) > 0$ and $X''(t) > 0$, and let τ_1 and τ_2 be two values of t with $\tau_1 < \tau_2$. The C^+ characteristics emanating from $(X(\tau_1), \tau_1)$ and $(X(\tau_2), \tau_2)$ have speeds [see (4.3.23)] given by

$$1 + \frac{3X(\tau_1)}{2} \quad \text{and} \quad 1 + \frac{3X(\tau_2)}{2},$$

respectively. Because X' is increasing by assumption, $X'(\tau_2) > X'(\tau_1)$, and it follows that the characteristic at τ_2 is faster than the characteristic τ_1 . Thus the two characteristics must intersect.

Let us now consider a special case and take the wall path to be

$$x = X(t) = at^2, \quad t > 0,$$

where a is a positive constant. From (4.3.23) the C^+ characteristics coming from the wall have the equation

$$x - a\tau^2 = (1 + 3a\tau)(t - \tau). \quad (4.3.26)$$

Equation (4.3.26) defines τ implicitly as a function of x and t . Therefore, given x and t , we may ask when it is possible to solve (4.3.26) uniquely for τ [i.e., determine a unique C^+ characteristic passing through (x, t)]. Writing (4.3.26) as

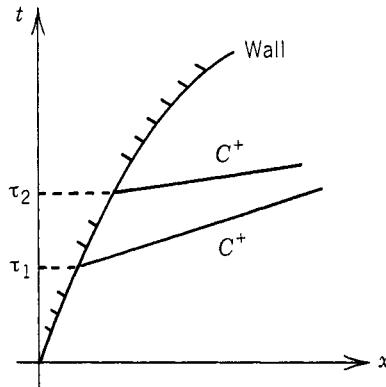


Figure 4.10 Two colliding C^+ characteristics emanating from an accelerating wall.

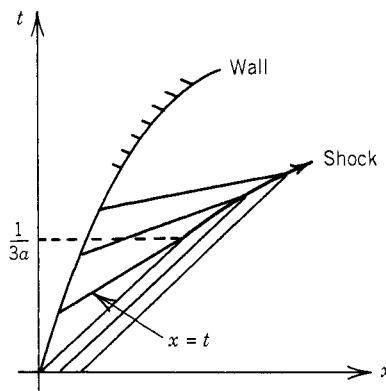


Figure 4.11 Characteristics intersecting to form a shock wave at time $t = 1/3a$.

$$F(x, t, \tau) = x - a\tau^2 + (1 + 3a\tau)(\tau - t) = 0, \quad (4.3.27)$$

we know from analysis that it is possible to locally solve $F(x, t, \tau) = 0$ for τ if $F_\tau(x, t, \tau) \neq 0$. In the present case

$$F_\tau = 1 + 4a\tau - 3at.$$

Therefore, the formation of the bore will occur at the first time t_b when $F_\tau = 0$, or

$$t_b = \min \left\{ \frac{1 + 4a\tau}{3a} : \tau > 0 \right\} = \frac{1}{3a}.$$

The breaking time is $1/(3a)$ and occurs along the $\tau = 0$ characteristic, or $x = t$. Figure 4.11 shows the characteristic diagram for this problem. The bore, after it forms, moves into a constant state $u = 0, h = 1$. The speed of the bore is determined by the jump conditions (4.3.6) and the state u and h behind the bore. Like most nonlinear initial-boundary value problems, this problem is computationally complex and cannot be resolved analytically, and one must resort to numerical methods.

We consider a simpler problem, where the wall moves into the motionless water with depth 1 at a *constant*, positive speed V ; that is, the path of the wall is $x = X(t) = Vt, t > 0$. We expect a bore to form immediately and move with constant velocity into the uniform state $u = 0, h = 1$. Behind the bore we expect that there is a uniform state $u = V$, with h still to be determined (note that the shallow-water equations can have constant solutions). The speed of the bore and the depth of water h_- behind the bore can be determined by the jump conditions (4.3.6). From those conditions, we obtain

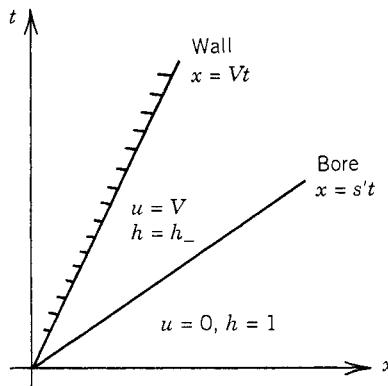


Figure 4.12 Bore with speed s' caused by a constant-velocity wall.

$$s' = \frac{h_- V}{h_- - 1}, \quad s' = \frac{V^2 h_- + \frac{1}{2} h_-^2 - \frac{1}{2}}{V h_-}. \quad (4.3.28)$$

The bore speed s' can be eliminated to obtain

$$h_-^3 - h_-^2 - (1 + 2V^2)h_- + 1 = 0, \quad (4.3.29)$$

which is a cubic equation for h_- . Using calculus, it is easy to see that there is a single negative root and two positive roots, one less than 1 and one greater than 1. We reject the negative root because the depth is nonnegative, and we reject the smaller positive root because the bore speed, from (4.3.28), would

be negative. Therefore, h_- is the largest positive root of (4.3.29). The solution is shown on the diagram in Figure 4.12.

4.3.5 Gas Dynamics

The equations governing one-dimensional gas dynamics, derived in Section 4.1, are given in conservation form by

$$\rho_t + (\rho u)_x = 0, \quad (\rho u)_t + (\rho u^2 + p)_x = 0, \quad (4.3.30)$$

where ρ is the density, u the velocity, and p the pressure given by the equation of state $p = k\rho^\gamma$, where $k > 0$ and $\gamma > 1$. It was shown in Exercise 2 of Section 4.2 that the characteristic form of (4.3.30) is

$$\frac{2c}{\gamma - 1} \pm u = \text{const} \quad \text{on} \quad \frac{dx}{dt} = u \pm c, \quad (4.3.31)$$

where $c^2 = p'(\rho)$. The form of the PDEs (4.3.30) and the Riemann invariants in (4.3.31) bear a strong resemblance to the equations from shallow-water theory. We chose to analyze the shallow-water equations, believing that the reader will have a better feeling for the underlying physical problem. We could just as well have studied the *piston problem* in gas dynamics. In this problem we imagine that the gas in a tube (see Figure 4.13) is set into motion by a piston at one end (the piston plays the role of the wall or wavemaker in the shallow-water equations). A device of a piston is not as unrealistic as it may first appear. For example, in aerodynamics the piston may model a blunt object moving into a gas; or, the piston may represent the fluid on one side of a valve after it is opened, or it may represent a detonator in an explosion process. The piston problem, that is, the problem of determining the motion of the gas for a given piston movement, is fundamental in nonlinear PDEs and gas dynamics. We leave some of these problems to the Exercises.

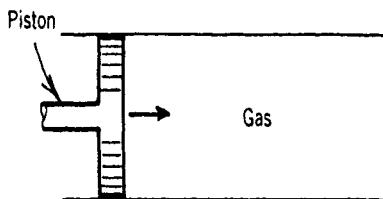


Figure 4.13 Piston moving into a gas.

EXERCISES

1. The shallow-water equations (4.3.4)–(4.3.5) are conservation laws in dimensioned variables. Select new dimensionless independent and dependent variables

$$\bar{x} = \frac{x}{L}, \quad \bar{t} = \frac{t}{T}, \quad \bar{h} = \frac{h}{H}, \quad \bar{u} = \frac{u}{\sqrt{gH}},$$

where $T = L/\sqrt{gH}$, and where H is the undisturbed water height and L is a typical wavelength. Rewrite the shallow-water equations in terms of dimensionless variables.

2. Solve the receding wall problem in the case that the wall path is given by $x = -Vt$, where V is a positive constant.
3. Verify that the cubic equation (4.3.29) always has one negative root and two positive roots.
4. Consider a gas at rest ($u = 0, \rho = \rho_0$) in a cylindrical tube $x > 0$, and suppose that the governing equations are given by (4.3.30) or (4.3.31). At time $t = 0$ a piston located at $x = 0$ begins to move along the spacetime path $x = X(t)$, $t > 0$. Discuss the resulting motion of the gas in the following cases:
 - (a) $X(t)$ satisfies the conditions $X(0) = 0$, $X'(t) < 0$, and $X''(t) < 0$.
 - (b) $X(t) = Vt$, where $V > 0$.
 - (c) If $X(t) = at^2$, where $a > 0$, at what time does a shock wave form?
5. A metallic rod of constant cross section initially at rest and occupying $x > 0$ undergoes longitudinal vibrations. The governing equations are

$$v_h - e_t = 0, \quad v_t - \frac{2Y}{\rho_0} ee_h = 0.$$

Here, t is time, h is a spatial coordinate attached to a fixed cross section with $h = x$ at $t = 0$; $v = v(h, t)$ is the velocity of the section h ; $e = e(h, t)$ is the strain, or the lowest-order approximation of the distortion at h ; $Y > 0$ is the stiffness (Young's modulus); and ρ_0 is the initial constant density. Beginning at $t = 0$, the back boundary ($h = 0$) is moved with velocity $-t$ [i.e., $v(0, t) = -t$, $t > 0$]. Write this problem in characteristic form, identifying the Riemann invariants and characteristics. Can the initial-boundary value problem be solved analytically by the Riemann method?

4.4 Hodographs and Wavefronts

Again using the shallow-water equations as the model, we illustrate some other standard procedures to analyze and understand nonlinear hyperbolic systems.

4.4.1 Hodograph Transformation

The hodograph transformation is a nonlinear transformation that changes a system of quasilinear equations to a linear one. The idea, attributed to Riemann, is to interchange the dependent and independent variables. To illustrate the method we consider the shallow-water equations, which we write in the dimensionless form (Exercise 1, Section 4.3)

$$h_t + uh_x + hu_x = 0, \quad (4.4.1)$$

$$u_t + uu_x + h_x = 0. \quad (4.4.2)$$

Both the depth h and the velocity u are functions of x and t :

$$h = h(x, t), \quad u = u(x, t). \quad (4.4.3)$$

Regarding (4.4.3) as a transformation from xt space to hu space, let us assume that the Jacobian of the transformation, namely

$$J = \begin{vmatrix} h_x & h_t \\ u_x & u_t \end{vmatrix} = h_x u_t - u_x h_t,$$

is never zero. Then the transformation (4.4.3) is invertible, and we can solve for x and t in terms of h and u to obtain $x = x(h, u)$, $t = t(h, u)$. The derivatives of the various quantities are related by the chain rule. We have

$$x_t = x_h h_t + x_u u_t = 0, \quad x_x = x_h h_x + x_u u_x = 1,$$

and similarly for t_t and t_x . From these equations it follows that

$$x_u = -\frac{h_t}{J}, \quad x_h = \frac{u_t}{J}, \quad t_u = \frac{h_x}{J}, \quad t_h = -\frac{u_x}{J}.$$

The shallow-water equations (4.4.1) and (4.4.2) can then be written

$$-x_u + ut_u - ht_h = 0, \quad (4.4.4)$$

$$x_h - ut_h + t_u = 0, \quad (4.4.5)$$

which is a linear system for $x = x(h, u)$ and $t = t(h, u)$. We can eliminate x by cross-differentiation to obtain a single second-order equation for $t = t(h, u)$ given by

$$t_{uu} = h^{-1}(h^2 t_h)_h. \quad (4.4.6)$$

This linear hyperbolic equation can be reduced to canonical form by the standard method of introducing characteristic coordinates ξ and η defined by

$$\xi = 2\sqrt{h} - u, \quad \eta = 2\sqrt{h} + u.$$

In terms of these coordinates, we obtain

$$-4t_{\xi\eta} = 0,$$

which has the general solution

$$t = f(\xi) + g(\eta),$$

where f and g are arbitrary functions. Therefore, equation (4.4.6) for t has general solution

$$t(h, u) = f(2\sqrt{h} - u) + g(2\sqrt{h} + u). \quad (4.4.7)$$

In principle, (4.4.4) can be solved for x (by integrating with respect to h) to obtain $x = x(h, u)$.

To obtain the solution to an initial-boundary value problem, one would have to determine the arbitrary functions from boundary or initial data and then invert the equations $t = t(h, u)$, $x = x(h, u)$ to obtain h and u . For example, suppose that initial conditions of the form

$$u(x, 0) = u_0(x), \quad h(x, 0) = h_0(x), \quad x \in \mathbb{R} \quad (4.4.8)$$

are appended to the shallow-water equations (4.4.1) and (4.4.2). Then, in principle, the variable x can be eliminated from (4.4.8) to obtain a locus in the hu plane along which $t = 0$ and x are given. Then, the initial value problem (4.4.1), (4.4.2), and (4.4.8) transforms to a Cauchy problem for (4.4.4) and (4.4.5). For all practical purposes this procedure is usually impossible to carry out, and it is better to perform numerical calculations on the original nonlinear problem; thus, as noted by Whitham (1974, p. 184), the hodograph analysis seems to be mainly of academic interest.

4.4.2 Wavefront Expansions

In Section 2.4 we introduced the idea of an expansion near a wavefront for a single first-order PDE. The same procedure works for systems as well. The problem is to determine how a continuous waveform with a derivative discontinuity propagates in time. In particular, along what path in spacetime do derivative discontinuities propagate, and how do the magnitudes of the jumps in the derivatives evolve? We expect, of course, that discontinuities in the derivatives will propagate along characteristics. We can validate this expectation by an

argument similar to that presented in Section 2.4. The main interest here is to determine how the magnitude of the jumps propagate in the special case that a wavefront is moving into a constant state. In this case an expansion in a neighborhood of the wavefront proves to be a convenient procedure, and we illustrate it with a specific example.

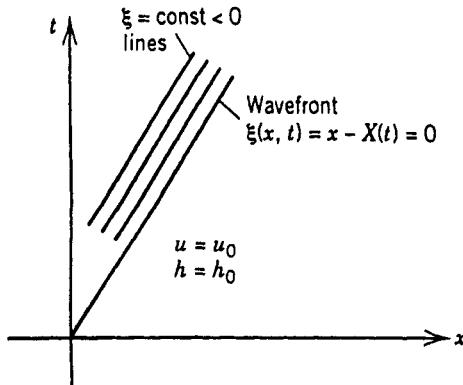


Figure 4.14 Wavefront moving into a constant medium.

At $t = 0$ we imagine a water wave, governed by the shallow-water theory, whose depth profile $h(x, 0)$ and accompanying velocity profile $u(x, 0)$ are continuous functions with a simple jump discontinuities in their spatial derivatives at $x = 0$, and smooth otherwise. Ahead of the wave ($x > 0$) there is a uniform state $h = h_0$, $u = u_0$. For $t > 0$ we assume that the discontinuity propagates along a curve $x = X(t)$ in spacetime, called a *wavefront*, where both h and u are smooth on each side of the curve. We introduce a coordinate ξ whose coordinate lines are parallel to the curve $x = X(t)$; in particular, we define

$$\xi = x - X(t).$$

Then the wavefront is given by $\xi = 0$; ahead of the wavefront $\xi > 0$, and $\xi < 0$ behind the wavefront (see Figure 4.14). The goal is to determine the behavior of the wave in a neighborhood of the wavefront, that is, for ξ small and negative.

Ahead of the wave we assume a uniform state, and behind the wave we assume that h and u have expansions as power series in ξ . Hence, we make the Ansatz

$$h = h_0, \quad u = u_0 \quad \text{for } \xi > 0, \tag{4.4.9}$$

$$h = h_0 + h_1(t)\xi + \frac{1}{2}h_2(t)\xi^2 + \dots, \tag{4.4.10}$$

$$u = u_0 + u_1(t)\xi + \frac{1}{2}u_2(t)\xi^2 + \dots, \quad \text{for } \xi < 0.$$

For $\xi < 0$ we then have

$$h_t = h'_1 \xi - h_1 X' - h_2 X' \xi + O(\xi^2), \quad h_x = h_1 + h_2 x + O(\xi^2), \quad (4.4.11)$$

and similarly for u_t and u_x . Substituting all these expressions into the governing PDEs (4.4.1) and (4.4.2), and then setting the coefficients of powers of ξ equal to zero, gives, to leading order

$$(u_0 - X')h_1 + h_0 u_1 = 0, \quad (4.4.12)$$

$$h_1 + (u_0 - X')u_1 = 0, \quad (4.4.13)$$

and at order $O(\xi)$

$$h'_1 - X'h_2 + u_0 h_2 + 2u_1 h_1 + h_0 u_2 = 0, \quad (4.4.14)$$

$$u'_1 - X'u_2 + u_1^2 + u_0 u_2 + h_2 = 0. \quad (4.4.15)$$

The leading order equations (4.4.12)–(4.4.13) are equations for u_1 and h_1 , that, according to (4.4.10), define corrections to u_0 and h_0 in the expansions for u and h . By (4.4.11) we see that h_1 gives the leading order approximation for h_x behind the wavefront, and because h is constant ahead of the front, h_1 therefore provides the leading order approximation for the jump in h_x across the front. A similar remark holds true for u_1 . The system (4.4.12)–(4.4.13) is homogeneous and will have a nontrivial solution u_1 and h_1 if, and only if, the determinant of the coefficient matrix vanishes, or

$$X' = u_0 \pm \sqrt{h_0}. \quad (4.4.16)$$

Equation (4.4.16) defines the wavefront speed, which is constant. Consequently, to leading order the wavefront is a straight line. There are two possibilities, depending on whether the plus or minus sign is chosen. If the plus sign is selected, we refer to the wavefront as a *downstream* wave, and if the minus sign is chosen (giving a slower speed), we refer to the wavefront as an *upstream* wave. In either case the solution to the homogeneous system (4.4.12)–(4.4.13) can be written

$$u_1 = \frac{(X' - u_0)h_1}{h_0}. \quad (4.4.17)$$

Therefore, u_1 and h_1 are linearly related, and so we need determine only one of the quantities.

We first examine downstream waves where $X' = u_0 + h_0^{1/2}$. Substituting into the $O(\xi)$ equations (4.4.14) and (4.4.15) yields

$$h'_1 - h_0^{1/2} h_2 + 2u_1 h_1 + h_0 u_2 = 0, \quad (4.4.18)$$

$$u'_1 - h_0^{1/2} u_2 + u_1^2 + h_2 = 0. \quad (4.4.19)$$

The goal is to obtain an equation that determines h_1 or u_1 . On first observation we have some difficulty because (4.4.18) and (4.4.19) contain both h_2 and u_2 , neither of which is known. However, further observation indicates that, remarkably, h_2 and u_2 may be eliminated from (4.4.18) and (4.4.19) by multiplying (4.4.19) by $\sqrt{h_0}$ and then adding the result to (4.4.18). This gives

$$h'_1 + 2u_1 h_1 + h_0^{1/2} u'_1 + h_0^{1/2} u_1^2 = 0. \quad (4.4.20)$$

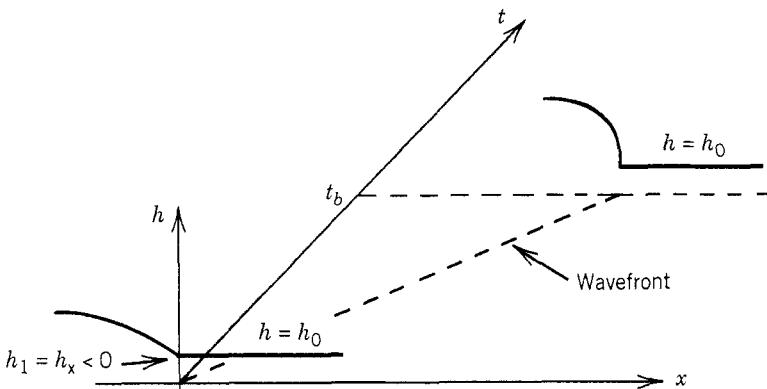


Figure 4.15 Wave with a negative gradient breaking downstream.

Now we can use (4.4.17) to eliminate u_1 from (4.4.20) and obtain

$$h'_1 + \frac{3}{2\sqrt{h_0}} h_1^2 = 0, \quad (4.4.21)$$

which is a single ordinary differential equation for h_1 . The variables in (4.4.21) separate and it is easy to find its solution

$$h_1 = \frac{2\sqrt{h_0}}{3t + 2\sqrt{h_0}/h_1(0)}, \quad (4.4.22)$$

where $h_1(0)$ denotes the initial value of h_1 (i.e., the initial value of the jump in the derivative h_x at $x = t = 0$).

Now, using (4.4.22), we can make some interesting observations regarding the evolution of a jump discontinuity. If $h_1(0)$ is negative, as shown in Figure 4.15, the wave will eventually break downstream. This is because the denominator of (4.4.22) will vanish at finite time, giving an infinite gradient just behind the wavefront. Waves with positive gradients will not break downstream.

In the case of *upstream waves*, where the minus sign is chosen in (4.4.16), a calculation similar to the one presented above leads to the following equation for h_1 :

$$h'_1 - \frac{3}{2\sqrt{h_0}} h_1^2 = 0.$$

Easily, this equation has solution

$$h_1 = \frac{2\sqrt{h_0}}{2\sqrt{h_0}/h_1(0) - 3t}.$$

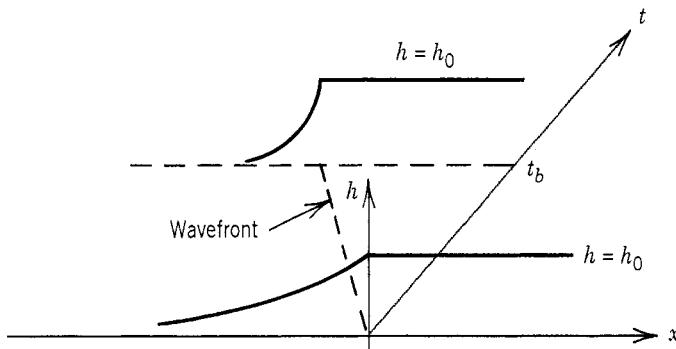


Figure 4.16 Tidal bore breaking upstream.

Consequently, if $h_1(0)$, the initial jump in h_x across the wavefront, is positive, the wave will break in finite time as shown in Figure 4.16. Such waves are called *tidal bores*.

We refer to Whitham (1974) for a full discussion of wavefront expansions for general hyperbolic systems and how those expansions relate to wavefronts in geometric optics. In the next section we devise a procedure, called *weakly nonlinear analysis*, similar to the one discussed above to determine how special types of short-wavelength signals are propagated in a gas dynamic medium.

EXERCISES

1. Consider the gas dynamic equations

$$\rho_t + (\rho u)_x = 0, \quad \rho(u_t + uu_x) + p_x = 0, \quad p = k\rho^\gamma.$$

Define the sound speed c by the equation $c^2 = p'(\rho)$.

- (a) Show that the equations may be written in terms of c and the particle velocity u as

$$c_t + uc_x + \frac{\gamma - 1}{2}cu_x = 0,$$

$$u_t + uu_x + \frac{2}{\gamma - 1}cc_x = 0.$$

- (b) Perform the hodograph transformation on the equations in part (a) to obtain the system

$$x_u - ut_u + \frac{\gamma - 1}{2}ct_c = 0, \quad x_c - ut_c + \frac{2}{\gamma - 1}ct_u = 0,$$

for $x = x(c, u)$ and $t = t(c, u)$.

- (c) Show that t satisfies the second-order wave equation

$$t_{bb} + \frac{2n}{b}t_b = t_{uu}, \quad \text{where } b = \frac{2c}{\gamma - 1},$$

where $n = (\gamma + 1)/[2(\gamma - 1)]$, and obtain the general solution in the case $n = 1$.

2. This exercise deals with the hodograph transformation and the *Born–Infeld equation*

$$(1 - \phi_t^2)\phi_{xx} + 2\phi_x\phi_t\phi_{xt} - (1 + \phi_x^2)\phi_{tt} = 0.$$

- (a) Introduce new independent and dependent variables via

$$\xi = x - t, \quad \eta = x + t,$$

$$u = \phi_\xi, \quad v = \phi_\eta,$$

and then apply the hodograph transformation to obtain the system

$$\xi_v - \eta_u = 0, \quad v^2\eta_v + (1 - 2uv)\xi_v + u^2\xi_u = 0.$$

- (b) Reduce the system in part (a) to the single linear hyperbolic equation

$$u^2\xi_{uu} + (1 + 2uv)\xi_{uv} + v^2\xi_{vv} + 2u\xi_u + 2v\xi_v = 0,$$

and then introduce characteristic coordinates

$$r = \frac{(1 + 4uv)^{1/2} - 1}{2v}, \quad s = \frac{(1 + 4uv)^{1/2} - 1}{2u},$$

to obtain $\xi_{rs} = 0$.

(c) Show that one may take

$$\xi = F(r) - \int s^2 G'(s) ds, \quad \eta = G(s) - \int r^2 F'(r) dr,$$

where F and G are arbitrary functions, and therefore derive the solution ϕ given by

$$\phi = \int r F'(r) dr + \int s G'(s) ds.$$

(d) Finally, introduce $\rho = F(r)$ and $\sigma = G(s)$, with inverses $r = \Phi_1'(\rho)$ and $s = \Phi_2'(\sigma)$, and show that the solution may be written in the form

$$\phi = \Phi_1(\rho) + \Phi_2(\sigma),$$

where

$$\rho = x - t + \int_{-\infty}^{\sigma} \Phi_2'(\sigma) d\sigma, \quad \sigma = x + t - \int_{\rho}^{\infty} \Phi_1'(\rho) d\rho.$$

(Whitham 1974, p. 617).

3. The shallow water equations governing flood waves on a flat river of incline α are

$$h_t + u h_x + h u_x = 0, \quad u_t + u u_x + g^* h_x = g^* S - \frac{C u^2}{h},$$

where $S = \tan \alpha$, $g^* = g \cos \alpha$, and C is the coefficient of friction. Here h is the depth of the water measured perpendicular from the bottom and u is the velocity in the x direction along the bed (Kevorkian 1990). Consider a wavefront $x = X(t)$ propagating into a uniform state $u = u_0$, $h = h_0$, along which h and u are continuous but have discontinuous derivatives. Perform a wavefront expansion analysis to show that a wave will break if

$$u_0 > 2\sqrt{g^* h_0},$$

regardless of the sign of $h_x(0^+, 0)$. Given $u_0 < 2\sqrt{g^* h_0}$, show that the jump in h_x will decay exponentially in the case $h_x(0^+, 0) > 0$, but in the case

$$h_x(0^+, 0) < 2S(g^* h_0)^{1/2} \frac{1 - u_0/(2(g^* h_0)^{1/2})}{3u_0},$$

the wave will break in finite time. (This means that a sufficiently strong downstream flood wave will be headed by a bore.)

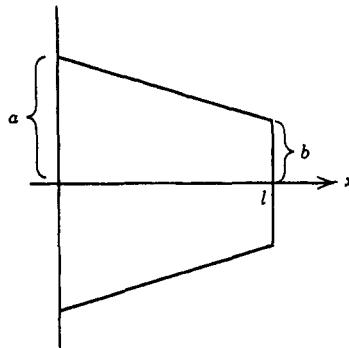


Figure 4.17 Wedge of unit thickness.

4. The small, longitudinal vibrations of a bar of density $\rho = \rho(x)$, stiffness $E = E(x)$, and cross-sectional area $\sigma = \sigma(x)$ are governed by the wave equation $\rho\sigma u_{tt} = (\sigma Eu_x)_x$, where $u = u(x, t)$ is the displacement of a cross section [for a derivation see, e.g., Lin & Segel 1974]. This equation can be transformed to a system of two first-order equations by introducing the variables $v = u_t$ (the velocity) and $w = u_x$ (the stress). The purpose of this exercise is to determine how a wavefront (caused, say, by striking one end of the bar) propagates down the length of the bar.

(a) Assume a wavefront expansion

$$v = v_0 + v_1(x)\xi + \dots, \quad w = w_0 + w_1(x)\xi + \dots \quad \text{for } \xi < 0,$$

where $\xi = t - \psi(x) = 0$ is the wavefront and where $v = v_0 = \text{const}$, $w = w_0 = 0$ for $\xi > 0$. Show that to leading order, the wavefront is described by the differential equation

$$\psi'(x) = \pm \sqrt{\frac{\rho(x)}{E(x)}}.$$

(b) Show that

$$v'_1(x) = -\frac{(\sigma E \psi')' v_1(x)}{2\sigma E \psi'},$$

and therefore

$$v_1(x) = v_1(0) \sqrt{\frac{\sigma(0)E(0)\psi'(0)}{\sigma(x)E(x)\psi'(x)}}$$

along the wavefront.

- (c) Consider a uniform (ρ and E constant) bar in the shape of a wedge as shown in Figure 4.17. Show that jumps in v_x and w_x both increase as the area decreases, and interpret the result physically.
5. Consider the nonlinear system

$$u_t - v_x = 0, \quad v_t - uu_x = 0.$$

Assume that a wavefront is moving into a constant state. To leading order, determine the speed of the front and determine how a jump in u_x evolves along the front.

4.5 Weakly Nonlinear Approximations

It is evident that general systems of nonlinear partial differential equations are complex and that no procedure, simple or otherwise, can be developed to analytically solve most of these problems. Therefore, much effort has gone into developing techniques to reduce the complexity of these systems. Often, this means examining a problem in some special limit whereby the system can be reduced to a simpler one, and progress can be made in elucidating some of the qualitative features of the original problem. One such method, which arose out of weakly nonlinear geometric optics, is an asymptotic method that yields a problem in what is termed a *weakly nonlinear limit*. In geometric optics one attempts to construct formal solutions that are high-frequency asymptotic solutions for bounded time intervals. The idea extends to other systems and can be explained as follows. Suppose that a medium exists in a rest state and that in some localized region of space a small-amplitude, small-wavelength disturbance occurs. This event is then propagated to nearby regions by the equations of motion, which describe the dynamics of the problem. In this process we try to identify a small parameter, and then we assume that the states can be represented as an expansion in that small parameter. The reader will recall that the acoustic approximation in gas dynamics was obtained in a similar manner, and the linear wave equation resulted in the special limit of waves of small amplitude. But now we wish to carry the analysis one step further and retain some of the nonlinear interactions that are present in the original system.

We illustrate the weakly nonlinear analysis by considering the gas dynamic equations. The central idea is to consider small-amplitude waves whose wavelength is small compared to the overall size of the spatial region where the problem is defined. We show that such disturbances are propagated by Burgers'

equation, thereby giving a derivation of one of the important model equations of applied analysis.

4.5.1 Derivation of Burgers' Equation

Consider the one-dimensional flow of a barotropic fluid that is governed by the gas dynamic equations

$$\rho_t + u\rho_x + \rho u_x = 0, \quad (4.5.1)$$

$$\rho u_t + \rho u u_x + p_x = \frac{4}{3}\mu u_{xx}, \quad (4.5.2)$$

$$p = k\rho^\gamma, \quad (4.5.3)$$

where ρ is the density, u the particle velocity, p the pressure, μ the viscosity, and k and γ positive constants. Equation (4.5.1) is mass conservation and equation (4.5.2) is momentum balance, where the only two acting forces are the pressure and the viscous force. Equation (4.5.3) is the barotropic equation of state. In this model flow there is no heat diffusion and no heat generated or dissipated by mechanical or viscous forces. Our goal is to reduce the complexity of (4.5.1)–(4.5.3) by restricting the range of applicability, that is, by investigating a wave propagation problem in a special limit.

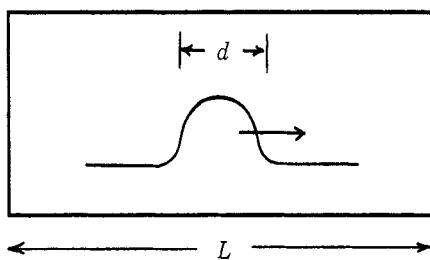


Figure 4.18 Wavelet of thickness d propagating in a container of characteristic length L .

To this end, assume that a wave with representative thickness d is propagating to the right into a constant, undisturbed state $u = 0, p = p_0, \rho = \rho_0$, and let $x = X(t)$ represent the location of some reference point on the wave as it moves forward in time; moreover, let L be a length that is representative of the size of the region where the wave is propagating (e.g., the length of a tube or vessel) (see Figure 4.18). The assumption is that d is small compared to L ,

and therefore the parameter ε defined by

$$\varepsilon = \frac{d}{L}$$

is a small quantity that can serve as an expansion parameter in a perturbation series.

In this problem it is essential to introduce normalized variables and reformulate the equations in dimensionless terms; only then can we be certain of the order of magnitude of the various quantities in the equations. The sound speed in the region ahead of the wave is given by $c_0 = (\gamma p_0 / \rho_0)^{1/2}$; therefore, the acoustic time t_a defined by $t_a = L/c_0$ is the time required for a signal traveling at speed c_0 to traverse a distance comparable to the magnitude of the length of the vessel. We take the acoustic time as a characteristic time for the problem and introduce the dimensionless time variable

$$\tau = \frac{t}{t_a}.$$

For the characteristic length we select the representative wave width d and introduce the dimensionless spatial coordinate

$$\xi = \frac{x - X(t)}{d},$$

which measures the distance to the location of the wave on the scale of the wave. We choose ρ_0 , p_0 , and c_0 to be the density, pressure, and velocity scales, respectively, and we define the dimensionless dependent variables $\bar{\rho}$, \bar{p} , and \bar{u} by

$$\bar{\rho} = \frac{\rho}{\rho_0}, \quad \bar{p} = \frac{p}{p_0}, \quad \bar{u} = \frac{u}{c_0}.$$

Letting $D(t) = X'(t)$ denote the velocity of the reference point on the wave, we also introduce $\bar{D}(t)$, defined by

$$\bar{D} = \frac{D}{c_0}.$$

In terms of the dimensionless variables, equations (4.5.1)–(4.5.3) become

$$\varepsilon \bar{\rho}_\tau - \bar{D} \bar{\rho}_\xi + \bar{u} \bar{\rho}_\xi + \bar{\rho} \bar{u}_\xi = 0, \quad (4.5.4)$$

$$\varepsilon \bar{\rho} \bar{u}_\tau - \bar{\rho} \bar{D} \bar{u}_\xi + \bar{\rho} \bar{u} \bar{u}_\xi + \gamma^{-1} \bar{\rho}_\xi = 2\varepsilon \nu \bar{u}_{\xi\xi}, \quad (4.5.5)$$

$$\bar{p} = \bar{\rho}^\gamma, \quad (4.5.6)$$

where

$$\nu = \frac{2\mu}{3\rho_0 c_0 \varepsilon d} \quad (4.5.7)$$

is a dimensionless constant and where $\varepsilon \ll 1$. The reader is asked to verify these equations in Exercise 1.

Next comes a crucial step, namely, the assumption of the form of the expansion of the dependent variables in powers of the small parameter ε . We take

$$\bar{\rho}(\xi, \tau) = 1 + \varepsilon \rho_0(\xi, \tau) + \varepsilon^2 \rho_1(\xi, \tau) + O(\varepsilon^3), \quad (4.5.8)$$

$$\bar{u}(\xi, \tau) = \varepsilon u_0(\xi, \tau) + \varepsilon^2 u_1(\xi, \tau) + O(\varepsilon^3), \quad (4.5.9)$$

$$\bar{D}(\tau) = D_0(\tau) + \varepsilon D_1(\tau) + O(\varepsilon^2). \quad (4.5.10)$$

Because of the equation of state (4.5.6), we automatically have, by substituting (4.5.8) into (4.5.6), the expansion

$$\bar{\rho}(\xi, \tau) = 1 + \varepsilon \gamma \rho_0(\xi, \tau) + \varepsilon^2 \frac{\gamma(\gamma - 1)}{2} \rho_0(\xi, \tau)^2 + O(\varepsilon^3). \quad (4.5.11)$$

This is just the Taylor series expansion of $\bar{\rho}^\gamma$ about $\varepsilon = 0$. Thus, the assumption is that the disturbance, whose density is $\bar{\rho}$, is of order 1 with small variations; the velocity D of the wave is on the same order as the speed of sound in the undisturbed medium ahead of the wave, and the particle velocity \bar{u} within the wave is small in comparison. Such assumptions describe what is termed a *weakly nonlinear model*.

The next step is to substitute the expansions (4.5.8)–(4.5.11) into the governing PDEs (4.5.4)–(4.5.5). It is straightforward to show that (4.5.4) becomes

$$\begin{aligned} \varepsilon^2 \rho_{0\tau} - (D_0 + \varepsilon D_1)(\varepsilon \rho_{0\xi} + \varepsilon^2 \rho_{1\xi}) + \varepsilon^2 u_0 \rho_{0\xi} \\ + (1 + \varepsilon \rho_0)(\varepsilon u_{0\xi} + \varepsilon^2 u_{1\xi}) = O(\varepsilon^3) \end{aligned} \quad (4.5.12)$$

and (4.5.5) becomes

$$\begin{aligned} \varepsilon^2 u_{0\tau} - (1 + \varepsilon \rho_0)(D_0 + \varepsilon D_1)(\varepsilon u_{0\xi} + \varepsilon^2 u_{1\xi}) + \varepsilon^2 u_0 u_{0\xi} \\ + \varepsilon \rho_{0\xi} + \varepsilon^2 \rho_{1\xi} + \varepsilon^2 (\gamma - 1) \rho_0 \rho_{0\xi} = 2\varepsilon^2 \nu u_{0\xi\xi} + O(\varepsilon^3). \end{aligned} \quad (4.5.13)$$

Because these equations should hold for any ε , we are free to set the coefficients of the various powers of ε equal to zero. The leading order contributions from equations (4.5.12)–(4.5.13) are

$$-D_0 \rho_{0\xi} + u_{0\xi} = 0, \quad (4.5.14)$$

$$-D_0 u_{0\xi} + \rho_{0\xi} = 0, \quad (4.5.15)$$

and the next-order contributions, which are the correction terms, are

$$-D_0 \rho_{1\xi} + u_{1\xi} = -\rho_{0\tau} + D_1 \rho_{0\xi} - u_0 \rho_{0\xi} - \rho_0 u_{0\xi}, \quad (4.5.16)$$

$$\begin{aligned} -D_0 u_{1\xi} + \rho_{1\xi} = -u_{0\tau} + D_1 u_{0\xi} + D_0 \rho_0 u_{0\xi} \\ - u_0 u_{0\xi} - (\gamma - 1) \rho_0 \rho_{0\xi} + 2\nu u_{0\xi\xi}. \end{aligned} \quad (4.5.17)$$

Equations (4.5.14) and (4.5.15) are two homogeneous equations for u_0 and ρ_0 ; there will be nontrivial solution provided that the determinant of the coefficient matrix vanishes:

$$D_0 = 1. \quad (4.5.18)$$

In this case both (4.5.14) and (4.5.15) become $\rho_{0\xi} = u_{0\xi}$. Because both u_0 and ρ_0 are zero ahead of the wave, we must have $u_0 = \rho_0$. We have therefore determined the leading order wave velocity D_0 in (4.5.10), and we have a relation between ρ_0 and u_0 .

To find a single equation for ρ_0 we analyze the correction equations (4.5.16) and (4.5.17). Substituting (4.5.18) into (4.5.16) and (4.5.17) and using $u_0 = \rho_0$, we obtain

$$-\rho_{1\xi} + u_{1\xi} = -\rho_{0\tau} + D_1\rho_{0\xi} - 2\rho_0\rho_{0\xi}, \quad (4.5.19)$$

$$\rho_{1\xi} - u_{1\xi} = -\rho_{0\tau} + D_1\rho_{0\xi} - (\gamma - 1)\rho_0\rho_{0\xi} + 2v\rho_0\rho_{0\xi\xi}. \quad (4.5.20)$$

Adding these two equations gives a single equation for ρ_0 . It is now convenient to introduce the quantity $U(\xi, \tau)$ defined by

$$U = \frac{(\gamma + 1)\rho_0}{2} = \frac{(\gamma + 1)u_0}{2}. \quad (4.5.21)$$

Then the sum of the equations gives

$$U_\tau - D_1 U_\xi + U U_\xi = \nu U_{\xi\xi}, \quad (4.5.22)$$

which is similar to Burgers' equation. We can eliminate the term in (4.5.22) involving D_1 by defining a new spatial coordinate η by

$$\eta = \xi + \int_0^\tau D_1(y) dy. \quad (4.5.23)$$

With this transformation, (4.5.22) becomes

$$u_\tau + uu_\eta = \nu u_{\eta\eta}, \quad (4.5.24)$$

which is *Burgers' equation*. Here, $u = u(\eta, \tau) = U(\xi(\eta, \tau), \tau)$; the lowercase letter u should not be confused with the velocity in the original equations (4.5.1)–(4.5.2).

To summarize, have obtained Burgers' equation in a weakly nonlinear limit of the gas dynamic equations (4.5.1)–(4.5.3). The procedure was to look for equations that govern small deviations from a uniform state in acoustic time, on the scale of the width of the disturbance. The reader should contrast the situation in acoustics, where linear equations (in fact, the wave equation) govern the small perturbations; there, acoustic signals are small-amplitude disturbances on the timescale of acoustic time and on the lengthscale of the vessel.

Burgers' equation, which followed from the weakly nonlinear limit, governs small-amplitude disturbances of short wavelength, compared to the size of the vessel, over the acoustic timescale.

It is common in applied mathematics to study complex problems on different timescales or lengthscales; generally, one obtains different governing equations in each limit, accompanied by an overall simplification. The solution of the problem in a specific limit often clarifies some aspects of complex physical systems, giving information about the behavior of the system under special conditions. A good example of the utility of scaling and examining a problem in special limits occurs in combustion theory, where there are several time scales: for example, the acoustic timescale, the chemical timescale, the timescale for diffusion, and the hydrodynamic timescale. By examining a problem on a rapid chemical timescale, for example, one may be able to neglect diffusion terms in the equations since diffusion may occur in some regimes on a much slower scale. In this manner one reduces the complexity of problems. Another illustration is water waves. Some of the important equations of applied mathematics, such as the Korteweg–deVries equation and the Boussinesq equations, result from studying the full nonlinear equations of hydrodynamics in a special limit (e.g., when the wavelength of a surface wave is long in comparison to the depth of the channel).

Earlier we showed the existence of traveling wave solutions to Burgers' equation. Now we relate that calculation to the discussion presented above. We recall that the determination of traveling waves for Burgers' equation involves a constant, unknown wave speed. This wave speed is related to the correction term D_1 in (4.5.10) and (4.5.23) in the following manner. The disturbance is centered along the coordinate $\xi = 0$ for all time. But in the $\eta\tau$ -coordinate system of equation (4.5.24) the wave is on the path $\eta = \int_0^\tau D_1(y)dy$, and its speed is $D_1(\tau)$, which is unknown. So we are free to search for constant values D_1 for which Burgers' equation (4.5.24) admits traveling wave solutions of speed D_1 . Generally, D_1 is not constant for wave profiles that change in time.

EXERCISES

1. Verify equations (4.5.4)–(4.5.6).
2. The purpose of this exercise is to perform a weakly nonlinear analysis on a simple nonlinear equation. Consider the scalar advection–diffusion equation

$$u_t + \phi(u)_x = \mu u_{xx},$$

where t , x , and u are dimensionless variables of order 1, $\mu = \mu_1\varepsilon^2 + O(\varepsilon^3)$, $\varepsilon \ll 1$, and $\phi(u)$ can be expanded in a Taylor series about $u = u_0 =$

constant. Let $x = X(t)$ be a representative location on a wave that is propagating into the uniform state u_0 . Introduce the variable $\xi = (x - X(t))/\varepsilon$ and assume that

$$\begin{aligned} u(\xi, t) &= u_0 + \varepsilon u_1(\xi, t) + \varepsilon^2 u_2(\xi, t) + O(\varepsilon^3), \\ D(t) &= D_0(t) + \varepsilon D_1(t) + O(\varepsilon^2), \end{aligned}$$

where $D = X'$. To leading order, determine the speed of the wave, and show that u_1 satisfies the Burgers'-like equation

$$u_{1t} + \left(\frac{\alpha}{2} u_1^2 \right)_\xi = \mu_1 u_{1\xi\xi},$$

where α is a constant.

3. Consider the PDE in Exercise 2 with $\mu = \mu(\varepsilon)$. To study the long-time behavior of the system, assumes the expansion

$$u(\eta, \tau) = u_0 + \varepsilon u_1(\eta, \tau) + \varepsilon^2 u_2(\eta, \tau) + O(\varepsilon^3),$$

where

$$\eta = x - X(t), \quad \tau = \varepsilon t,$$

and where $x = X(t)$ is a reference location on a wave propagating into the constant state $u = u_0$. To leading order, determine the speed of the wave and find an equation for the first correction $u_1(\eta, \tau)$. Discuss the two cases $\mu = O(\varepsilon)$ and $\mu = O(\varepsilon^2)$.

Reference Notes. Parallel introductory treatments of gas dynamics and the shallow water equations can be found in Kevorkian (1990) and Logan (2006a). Two definitive works on gas dynamics are the classic treatises by Courant & Friedrichs (1948) and Whitham (1974). Smoller (1994) addresses some of the mathematical issues in these areas. Another complete reference is Rozdestvenskii & Janenko (1983). An introduction to the theory of hyperbolic conservation laws is found in Majda (1986).

5

Diffusion Processes

In this chapter we examine equations that model basic diffusion processes and introduce mathematical methods that are useful in studying the structure of their solutions. These methods include similarity methods, the asymptotic expansion of integrals, and phase plane methods. The latter is included in an appendix to the chapter.

First we review fundamentals and relate diffusion to the probability concept of a random walk. In Section 5.2 the idea of invariance of a PDE under a one-parameter family of stretching transformations is introduced, and we show how this type of invariance leads to a reduction of the PDE to an ODE. Solutions obtained in this way are called *similarity solutions*, and the method of determining solutions through invariance properties is one of the basic techniques of applied mathematics; it is applicable to many classes of PDEs. In Section 5.3 we apply the similarity method to nonlinear diffusion models. One of the most significant results is that, unlike their linear counterparts, nonlinear diffusion equations can propagate signals that resemble a wavefront traveling at finite speed. Sections 5.4 and 5.5 extend the study to reaction-diffusion processes and advection-diffusion processes, respectively. Fisher's equation is a prototype of reaction-diffusion processes, and we investigate the existence of traveling wave solutions and their stability. In the same way, Burgers' equation, examined in Section 5.5, is the basic model equation for nonlinear advection-diffusion processes. In Section 5.6 the asymptotic expansion of integrals containing a large parameter is discussed briefly and Laplace's theorem is applied to understand the behavior of the solution to the initial value problem for Burgers' equation. In particular, we investigate the solution in the limit of small values of the

diffusion constant, and the initial value problem with a point source is solved.

Finally, the Appendix at the end of the chapter reviews two-dimensional systems and phase plane analysis. Many problems in PDEs, especially those associated with the determination of traveling wave solutions, can be studied in the context of a two-dimensional dynamical system.

5.1 Diffusion and Random Motion

To set the stage, we briefly state a few basic results from linear diffusion theory, some of which were introduced in Chapter 1. The linear diffusion equation is the parabolic PDE

$$u_t = Du_{xx},$$

which comes from the conservation law

$$u_t + \phi_x = 0$$

and Fick's law,

$$\phi = -Du_x.$$

The diffusion constant D , measured in length-squared per time, is a measure of how fast the quantity measured by u (particles, chemicals, animals, heat energy, etc.) diffuses from high concentrations to low concentrations. Fick's law stipulates that the flow is *down the gradient*.

In Section 1.5 we showed that the initial value problem for the diffusion equation

$$u_t = Du_{xx}, \quad x \in \mathbb{R}, \quad t > 0, \tag{5.1.1}$$

$$u(x, 0) = u_0(x), \quad x \in \mathbb{R}, \tag{5.1.2}$$

is

$$u(x, t) = \int_{\mathbb{R}} u_0(\xi) K(x - \xi, t) d\xi, \tag{5.1.3}$$

where the *diffusion kernel* $K(y, t)$ is

$$K(y, t) = \frac{1}{\sqrt{4\pi Dt}} e^{-y^2/4Dt}.$$

For any ξ and any $t > 0$ the kernel $K(x - \xi, t)$ is itself a solution to the diffusion equation, called the *fundamental solution*. Time snapshots of the kernel function are shown in Figure 5.1. One can show that K has the following properties:

- (1.) $\lim_{t \rightarrow 0^+} K(x - \xi, t) = 0$ for each fixed $x \neq \xi$.

- (2.) $\lim_{t \rightarrow 0^+} K(x - \xi, t) = +\infty$ for $x = \xi$.
- (3.) $\lim_{|x| \rightarrow \infty} K(x - \xi, t) = 0$ for each fixed $t > 0$.
- (4.) $\int_{\mathbb{R}} K(x - \xi, t) dx = 1, \quad t > 0$.

These properties imply that the diffusion kernel $K(x - \xi, t)$ is a solution to the initial value problem

$$\begin{aligned} u_t &= Du_{xx}, \quad x \in \mathbb{R}, \quad t > 0, \\ u(x, 0) &= \delta(x - \xi), \quad x \in \mathbb{R}, \end{aligned}$$

where $\delta(x - \xi)$ is the delta function, that is, a distribution representing a point source of unit intensity at $x = \xi$. Figure 5.1 illustrates how a unit amount (of energy, say) applied at $x = \xi$ at $t = 0$ diffuses through a medium with diffusion constant D . The solution of the initial value problem can be regarded as the superposition of a continuum distribution of sources $u_0(\xi)$, $\xi \in \mathbb{R}$. Furthermore, linear diffusion is a process where initial signals are propagated with infinite speed; for example, the solution $K(x - \xi, t)$ is nonzero for any x , regardless of how large, for any time $t > 0$, and regardless of how small. Therefore, we must always be aware of the limitations imposed by the diffusion equation when modeling a physical process (e.g., heat flow).

Another important result associated with the diffusion equation is the maximum principle. We take up this principle in Chapters 6 and 7. For the present we remark that a solution to the diffusion equation must take on its maximum value on the boundary of the domain over which the problem is defined. Physically, this means that the density function u cannot clump in the interior of the domain. This result is physically plausible because the diffusion equation comes from Fick's law, an assumption that forces movement from high to low concentrations. The time snapshots of the fundamental solution of the diffusion

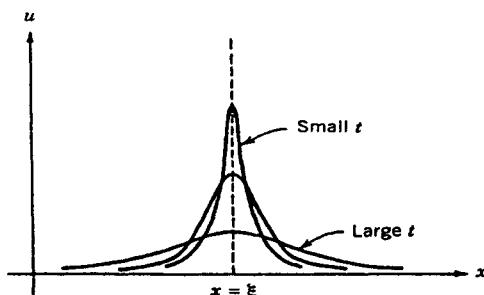


Figure 5.1 Time snapshots of the diffusion kernel $K(x - \xi, t)$ for various times t .

equation shown in Figure 5.1 suggest that the diffusion equation does indeed describe diffusion-like processes.

There are different ways to think about diffusion. In chemistry and physics it is easy to deduce, reasoning at an atomic or molecular level, how substances diffuse because of random motion. This microscopic description of diffusion is based on statistics and the fact that atoms or molecules collide randomly. These random collisions cause an assemblage of molecules to move from regions of high concentrations to regions of lower concentrations. For example, if a noxious gas is released at a point in a room, the colliding molecules of gas and the air lead to the eventual dispersion of the gas throughout the room. This stochastic, molecular model is the basis of the kinetic theory of gases, and it has a firm root in mathematics and in empirics.

In chemically based biological systems where we are tracking molecular concentrations, we observe the same phenomenon. For example, a chemical, toxic to fish, might be dumped locally in a reservoir; the spread of the chemical throughout the reservoir can be understood by the molecular collision model. In other biological systems, however, the same spreading effects are observed, but the cause is not random collisions on a microscopic scale, but interactions on a macroscopic scale. This effect is illustrated by the spread of an infectious disease. If infected individuals are introduced in a large population of susceptible individuals, then the infection can spread throughout the population by personal contact or respiration of airborne particulates. Even in clumped populations of organisms, there is a tendency for individuals to migrate from the high density regions to low density regions; this macroscopic dispersion is the result of population pressures such as competition for resources or might, in humans, even have some psychological basis. But the observation is the same as in the molecular collision model—movement from high to low concentrations.

We modeled diffusion by *Fick's law* (or Fourier's law in heat conduction), which states that the flux, or motion, is proportional to the steepness of the concentration curve: $\phi(x, t) = -Du_x(x, t)$. So the migration is “down the concentration gradient,” away from the most concentrated regions, and the greater the changes in concentration, the faster the motion. When this relation is substituted into the basic conservation law $u_t = -\phi_x$, we obtain the diffusion equation $u_t = Du_{xx}$.

Now we want to show how this diffusion model arises from a stochastic argument based on random motion. We might expect the two to be related because the fundamental solution to the diffusion equation generated by a point source at $x = \mu$, and given by the diffusion kernel

$$k(x - \mu, t) = \frac{1}{\sqrt{4\pi Dt}} e^{-(x-\mu)^2/4Dt},$$

looks very similar to the normal probability density. Recall that a normal ran-

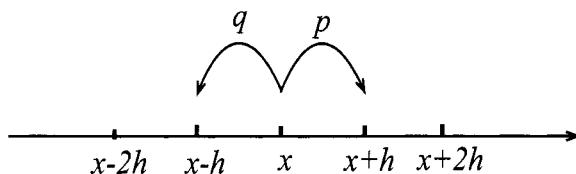


Figure 5.2 A lattice of points on the x axis on which a particle moves randomly. In each small time step τ it moves one lattice point to the right with probability p or one lattice point to the left with probability $q = 1 - p$.

dom variable X has probability density function (the “bell-shaped curve”)

$$N(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2},$$

where μ is the mean, or center of the density, and σ is the standard deviation, a measure of the spread about the mean. Amazingly enough, if we take the standard deviation to depend on time via $\sigma = \sqrt{2Dt}$, so that the spread of the data increases with time, then we obtain $K(x - \mu, t)$, which is the fundamental solution to the diffusion equation!

The close relationship between the diffusion and probability is not accidental. We now look in a different direction to affirm this relationship.

Let x be an arbitrary point on the x axis and divide the x axis into small equal segments of length h so that the entire axis is composed of a lattice of discrete points $\dots, x - 2h, x - h, x, x + h, x + 2h, \dots$. (See Figure 5.2.) We imagine a particle (an atom, cell, animal, etc.) that moves randomly on this lattice of points in such a way that if it resides at one of these lattice points at time t , then, during a small interval of time τ , it moves to the lattice point to the right with probability p or to the lattice point to the left with probability $q = 1 - p$. Such a process is called a *random walk*. Let $u = u(x, t)$ be the probability that the particle is at x at time t . We want an equation for u . Clearly, there are two ways to get to x at time $t + \tau$. The particle can be at $x - h$ at time t and jump to the right, or be at $x + h$ at time t and jump to the left. The law of total probability¹ therefore implies that

$$u(x, t + \tau) = u(x - h, t)p + u(x + h, t)q,$$

which is a partial difference equation in u . Now we expand the terms on the

¹ If an event E consists of two disjoint events A and B, the $\Pr(E) = \Pr(E|A)\Pr(A) + \Pr(E|B)\Pr(B)$.

left and the right in Taylor series about the point (x, t) to obtain

$$\begin{aligned} u(x, t) + u_t(x, t)\tau + O(\tau^2) &= p(u(x, t) - u_x(x, t)h + \frac{1}{2}u_{xx}(x, t)h^2 + O(h^3)) \\ &\quad + q(u(x, t) + u_x(x, t)h + \frac{1}{2}u_{xx}(x, t)h^2 + O(h^3)), \end{aligned}$$

where $O(\tau^2)$ denotes higher-order terms that have a factor of at least τ^2 and $O(h^3)$ denotes terms that have a factor of at least h^3 . Simplifying this equation gives

$$u_t(x, t) = -(p - q)\frac{h}{\tau}u_x(x, t) + \frac{h^2}{2\tau}u_{xx}(x, t) + O(\tau) + \frac{1}{\tau}O(h^3). \quad (5.1.4)$$

First, in the case $p = q = \frac{1}{2}$, this equation becomes

$$u_t(x, t) = \frac{h^2}{2\tau}u_{xx}(x, t) + O(\tau) + \frac{1}{\tau}O(h^3).$$

Now we take a special limit as $\tau \rightarrow 0$ and $h \rightarrow 0$ while maintaining the ratio $h^2/2\tau$ at a fixed constant D , then we obtain

$$u_t(x, t) = Du_{xx}(x, t),$$

which is the diffusion equation.

If $p \neq q$, we further assume that $p - q = O(h)$, which means that the probabilities of moving to the left or to the right are nearly the same. Then we take the limit in (5.1.4) as as $\tau \rightarrow 0$ and $h \rightarrow 0$ while maintaining

$$\lim \frac{h^2}{2\tau} = D, \quad \lim(p - q)\frac{h}{\tau} = c.$$

We obtain the advection-diffusion equation

$$u_t(x, t) = -cu_x(x, t) + Du_{xx}(x, t),$$

or the diffusion equation with drift.

Example. (Biological Invasion) Consider a very long canal where a nonindigenous alga species is accidentally released at $x = 0$. After a long period of time it is observed that the speed of the biological invasion outward in both directions from $x = 0$ is nearly constant. We show that this observation can be explained by a growth-diffusion model

$$u_t = Du_{xx} + \gamma u, \quad x \in \mathbb{R}, \quad t > 0.$$

Initially, $u(x, 0) = u_0\delta(x)$, which is a point source at $x = 0$ of magnitude u_0 . Let $x = x_f(t)$ denote the position of the wavefront, defined by the location

where $u = u_f$, where u_f is a small, given value of the density and $u_f < u_0$. Then, using the fundamental solution to the growth-diffusion equation, we get

$$\frac{u_0}{\sqrt{4\pi Dt}} e^{-x_f(t)^2/4Dt + \gamma t} = u_f.$$

Isolating the exponential term of the left side and then taking logarithms gives

$$\gamma t - \frac{x(t)^2}{4Dt} = \ln \left(\frac{u_f}{u_0} \sqrt{4\pi Dt} \right).$$

Now we make a clever approximation. For large times t , the linear term in t on the left side is much larger than the logarithm term on the right side because t grows much faster than does $\ln t$. Therefore, the only two terms in this equation that can balance for large t are the two terms on the left. Thus the long time, approximate position of the invasion front is

$$x_f(t) \approx \sqrt{4\gamma D}t.$$

Consequently, the speed of the front approaches a constant value $\sqrt{4\gamma D}$. \square

EXERCISES

1. Show that the advection-diffusion-decay equation

$$u_t = Du_{xx} - cu_x - ru$$

can be transformed into the linear diffusion equation with an appropriate change of independent and dependent variables.

2. As noted, the fundamental solution $K(x, t)$ of diffusion equation is nonzero for all $t > 0$, and therefore in a unit, concentrated point source spreads out infinitely fast, giving a nonphysical result. If we consider ultra-small amplitude signals, say of magnitude η , to be imperceptible, then the more physical *effective region of influence* of a point source is defined as

$$R = \{(x, t) : K(x, t) \geq \eta\}.$$

On a spacetime diagram sketch the region R .

3. If u_0 is a bounded, continuous function on \mathbb{R} , show that the solution $u(x, t)$ of the initial value problem (5.1.1)–(5.1.2) satisfies the condition

$$\inf u_0(z) \leq u(x, t) \leq \sup u_0(z),$$

where the sup and inf are taken over all of \mathbb{R} .

4. Find all solutions of the diffusion equation $u_t = u_{xx}$ having the form

$$u = \frac{1}{\sqrt{t}} f(z), \quad \text{where } z = \frac{x}{2\sqrt{t}}.$$

5. Find the solution of the initial value problem (5.1.1)–(5.1.2) when $D = 1$ and the initial condition is given by

$$u_0(x) = 0 \quad \text{if } x < 0, \quad u_0(x) = 1 \quad \text{if } x > 0.$$

The solution is

$$u = \frac{1 + \operatorname{erf}(x/\sqrt{4t})}{2}.$$

6. Consider the initial value problem (5.1.1)–(5.1.2) for the diffusion equation, and assume $u_0 > 0$, $D = 1$, and $\int_R u_0(x) dx = 1$. Let

$$\bar{x}(t) = \int_{\mathbb{R}} xu(x, t) dx \quad \sigma^2(t) = \int_{\mathbb{R}} (x - \bar{x})^2 u(x, t) dx$$

denote the mean and variance of the distribution $u(x, t)$, respectively. Prove that $\bar{x}(t) = \bar{x}(0)$ and $\sigma^2(t) = 2t + \sigma^2(0)$.

7. A model of nonlocal diffusion is the equation

$$u_t = (k * u)(x, t), \quad x \in \mathbb{R}, \quad t > 0,$$

where $k * u$ is the *convolution* of $k = k(x)$ and $u = u(x, t)$ defined by $(k * u)(x, t) = \int_{\mathbb{R}} k(x - y)u(y, t) dy$. The convolution averages u against the translates of the given continuous function k . Assume that k is an even function with the property that $k(x)x^n$ is absolutely integrable on \mathbb{R} for each $n \geq 0$, and assume that u is sufficiently smooth. Show that the nonlocal equation may be approximated by the local equation

$$u_t = m_0 u + m_2 u_{xx} + m_4 u_{xxxx} + \dots,$$

where m_n is the n th moment of k defined by

$$m_n = \frac{1}{n!} \int_{-\infty}^{\infty} k(y)y^n dy, \quad n = 0, 1, 2, \dots$$

Take $k(x) = \exp(-x^2)$ and numerically compute the even moments m_0 through m_6 .

8. Consider the Barenblatt equation

$$u_t = \begin{cases} ku_{xx} & \text{if } u_t \geq 0 \\ mu_{xx} & \text{if } u_t < 0 \end{cases},$$

where k and m are distinct positive constants. Is this equation linear? Show that the Barenblatt equation can be written

$$au_t + b|u_t| = u_{xx}$$

for some constants a and b .

5.2 Similarity Methods

In this section we introduce a powerful method, called the *similarity method*, for finding transformations that reduce PDEs to ODEs. This method takes advantage of the natural symmetries in a PDE and permits us to define special variables that facilitate the reduction. Equations that model physical problems often inherit symmetries from the underlying physical system; for example, a physical system that is translationally invariant produces governing equations that are unchanged under a translation of coordinates. Equations with symmetries, or equivalently, equations that are invariant under a given transformation, have a simple structure that can be used to advantage to simplify the problem. The similarity method is applicable to all types of PDEs, and it is particularly applicable to diffusion equations.

We do not present the most general similarity method, but rather, refer the reader to one of the books listed in the references. [The first edition of the applied mathematics text by Logan (1987) has a general introduction.] In this section we focus on a method developed by G. D. Birkhoff in the 1930s and consider the special case of scale transformations, or transformations of the variables represented by simple multiples of those variables. These are called *stretching transformations*, and Birkhoff's method is called the *method of stretchings*.

At the outset we consider a first-order PDE of the form

$$G(x, t, u, p, q) = 0, \quad p = u_x, \quad q = u_t. \quad (5.2.1)$$

The formal definition is as follows.

Definition. A *one-parameter family of stretching transformations*, denoted by T_ε , is a transformation on xtu space of the form

$$\bar{x} = \varepsilon^a x, \quad \bar{t} = \varepsilon^b t, \quad \bar{u} = \varepsilon^c u, \quad (5.2.2)$$

where a , b , and c are constants and ε is a real parameter restricted to some open interval I containing $\varepsilon = 1$.

By the chain rule, the stretching transformation (5.2.2) automatically induces a transformation on the derivatives p and q via the formulas

$$\bar{p} = \varepsilon^{c-a} p, \quad \bar{q} = \varepsilon^{c-b} q. \quad (5.2.3)$$

To motivate the definition of what it means for the PDE (5.2.1) to be invariant under T_ε , we use an example.

Example. Consider the nonlinear advection equation

$$u_t + uu_x = 0,$$

or, equivalently

$$q + up = 0. \quad (5.2.4)$$

In the transformed coordinate system the operator defining (5.2.4) is $\bar{q} + \bar{u}\bar{p}$. According to (5.2.2) and (5.2.3), we obtain

$$\bar{q} + \bar{u}\bar{p} = \varepsilon^{c-b}q + \varepsilon^{2c-a}up. \quad (5.2.5)$$

If we impose $-b = c + a$, then (5.2.5) can be written

$$\bar{q} + \bar{u}\bar{p} = \varepsilon^{2c-a}(q + up). \quad (5.2.6)$$

Therefore, under the transformation

$$\bar{x} = \varepsilon^a x, \quad \bar{t} = \varepsilon^{a-c}t, \quad \bar{u} = \varepsilon^c u, \quad (5.2.7)$$

for any constants a and c , the expression defining the PDE in the transformed coordinate system is a multiple of the original expression defining the PDE. This is what is meant by invariance of (5.2.4) under (5.2.7). \square

Definition. The PDE (5.2.1) is *invariant* under the one-parameter family T_ε of stretching transformations (5.2.2) if, and only if, there exists a smooth function $f(\varepsilon)$ such that

$$G(\bar{x}, \bar{t}, \bar{u}, \bar{p}, \bar{q}) = f(\varepsilon)G(x, t, u, p, q) \quad (5.2.8)$$

for all ε in I , with $f(1) = 1$. If $f(\varepsilon) = 1$ for all ε in I the PDE is *absolutely invariant*.

Now we can state and prove the basic reduction theorem.

Theorem. If the PDE (5.2.1) is invariant under T_ε given by (5.2.2), then the transformation

$$u = t^{c/b}y(z), \quad z = \frac{x}{t^{a/b}} \quad (5.2.9)$$

reduces the PDE (5.2.1) to a first order ordinary differential equation in $y(z)$ of the form

$$g(z, y, y') = 0. \quad (5.2.10)$$

The new independent variable z defined in (5.2.9) is called a *similarity variable*, and (5.2.9) is called a *similarity transformation*. After solving (5.2.10) for the unknown function y , substitution into (5.2.9) yields the *self-similar* form of the solution u .

Before proving the theorem, we give another example.

Example. We observed that the nonlinear advection equation (5.2.4) is invariant under the stretching transformation (5.2.7). Then the similarity transformations is given by

$$u = t^{c/(a-c)}y(z), \quad z = \frac{x}{t^{a/(a-c)}}. \quad (5.2.11)$$

The constants a and c are arbitrary at this point, and they are usually chosen so that (5.2.11) can satisfy certain initial or boundary conditions given with the problem. Substituting (5.2.11) into the (5.2.4) requires the partial derivatives u_t and u_x . The chain rule gives

$$u_x = \frac{y'}{t}, \quad u_t = \frac{c}{a-c}t^{c/(a-c)-1}y + \frac{ax}{c-a}t^{c/(a-c)}t^{a/(c-a)-1}y'.$$

Consequently, (5.2.4) becomes

$$yy' + \frac{a}{a-c}zy' + \frac{c}{c-a}y = 0, \quad (5.2.12)$$

which is an ordinary differential equation for $y = y(z)$. \square

The proof of the basic reduction theorem is straightforward. By invariance of (5.2.2) under T_ε we infer that (5.2.8) holds; and, because (5.2.2) holds for all ε in some interval I containing $\varepsilon = 1$, we may differentiate (5.2.8) with respect to ε and afterward set $\varepsilon = 1$ to obtain

$$axG_x + btG_t + cuG_u + (c-a)pG_p + (c-b)qG_q = f'(1)G. \quad (5.2.13)$$

This first-order linear PDE for G is a consequence of the invariance assumption. As expected, not every PDE (5.2.2) will be invariant under T_ε , and (5.2.13) imposes a condition on the form of G . The characteristic system associated with (5.2.13) is

$$\frac{dx}{ax} = \frac{dt}{bt} = \frac{du}{cu} = \frac{dp}{c-a} = \frac{dq}{c-b} = \frac{dG}{f'(1)G},$$

and there are five independent first integrals given by

$$\frac{x}{t^{a/b}}, \quad ut^{-c/b}, \quad pt^{(a-c)/b}, \quad qt^{1-b/c}, \quad \text{and} \quad Gt^{-f'(1)/b}.$$

Therefore, the general solution of (5.2.13) is (Chapter 2)

$$G = t^{f'(1)/b}\Psi(z, u^{t-c/b}, pt^{(a-c)/b}, qt^{1-b/c}), \quad (5.2.14)$$

where Ψ is an arbitrary function and $z = x/t^{a/b}$. This equation fixes the form of G . Now u is given by (5.2.9), and the partial derivatives p and q may be calculated to give

$$p = u_x = t^{(c-a)/b}y', \quad q = u_t = \left(\frac{c}{b}\right)t^{-1+c/b}y - \frac{a}{b}t^{c/b-a/b-1}xy'.$$

Substituting these expressions into (5.2.14) yields the equation

$$\Psi\left(z, y, y', \frac{cy}{b} - \frac{azy'}{b}\right) = 0,$$

which is an ordinary differential equation of the form (5.2.10), completing the proof. \square

The method is easily extended to second-order PDEs. We state the basic reduction theorem in this case and leave the proof to the reader.

Theorem. If the second-order PDE

$$G(x, t, u, u_x, u_t, u_{xx}, u_{xt}, u_{tt}) = 0 \quad (5.2.15)$$

is invariant under the one-parameter family T_ε of stretching transformations (5.2.2), then the transformation (5.2.9) reduces the (5.2.15) to a second-order ordinary differential equation of the form

$$g(z, y, y', y'') = 0. \quad (5.2.16)$$

We observe that the one-parameter family of transformations (5.2.2) on xtu space induces a transformation on the second derivatives, just as it did on the first derivatives [see (5.2.3)]. Denoting second derivatives by

$$r = u_{xx}, \quad s = u_{xt}, \quad v = u_{tt},$$

it follows immediately that

$$\bar{r} = \varepsilon^{c-2a}r, \quad \bar{s} = \varepsilon^{c-a-b}s, \quad \bar{v} = \varepsilon^{c-2b}v.$$

Example. Consider the diffusion equation

$$u_t - Du_{xx} = 0, \quad (5.2.17)$$

or

$$q - Dr = 0. \quad (5.2.18)$$

We now determine a stretching transformation (5.2.2) under which (5.2.18) is invariant. We have

$$\bar{q} - D\bar{r} = \varepsilon^{c-b}q - D\varepsilon^{c-2a}r = \varepsilon^{c-b}(q - Dr),$$

provided that

$$b = 2a. \quad (5.2.19)$$

Therefore (5.2.18) is invariant under the stretching transformation

$$\bar{x} = \varepsilon^a x, \quad \bar{t} = \varepsilon^{2a} t, \quad \bar{u} = \varepsilon^c u \quad (5.2.20)$$

for any choice of the constants a and c . The similarity transformation is then given by

$$u = t^{c/2a}y(z), \quad z = \frac{x}{\sqrt{t}}. \quad (5.2.21)$$

Substitution into (5.2.17) gives

$$Dy'' + \frac{z}{2}y' - \frac{c}{2a}y = 0. \quad (5.2.22)$$

Next let us impose conditions on (5.2.17) and consider the problem on the domain $x > 0, t > 0$, subject to the initial condition

$$u(x, 0) = 0, \quad x > 0, \quad (5.2.23)$$

and the boundary conditions

$$u(0, t) = 1, \quad u(\infty, t) = 0, \quad t > 0. \quad (5.2.24)$$

Physically, this problem models diffusion into the region $x > 0$ where the concentration u is zero initially and a constant concentration $u = 1$ is imposed at $x = 0$. From (5.2.21) it follows that (5.2.23) and the second equation in (5.2.24) both translate into the condition

$$y(\infty) = 0. \quad (5.2.25)$$

On the other hand, the first condition in (5.2.24) forces

$$u(0, t) = t^{c/2a}y(0) = 1, \quad t > 0.$$

The only possibility that the left side can be independent of t is to force $c = 0$. Therefore, the ODE (5.2.22) becomes

$$y'' + \frac{z}{2D}y' = 0, \quad z > 0, \quad (5.2.26)$$

subject to the boundary conditions

$$y(0) = 1, \quad y(\infty) = 0. \quad (5.2.27)$$

It is straightforward to solve (5.2.26) to obtain

$$y(z) = c_1 + c_2 \int_0^z e^{-\eta^2/4D} d\eta, \quad (5.2.28)$$

where c_1 and c_2 are constants of integration. Equation (5.2.28) may be written in terms of the *error function* $\text{erf}(s)$

$$\text{erf}(s) = \frac{2}{\sqrt{\pi}} \int_0^s e^{-\eta^2} d\eta,$$

as

$$y(z) = c_1 + c_2 \sqrt{\pi D} \text{erf} \left(\frac{z}{\sqrt{4D}} \right).$$

The condition $y(0) = 1$ forces $c_1 = 1$, and subsequently the condition $y(\infty) = 0$ implies that

$$1 + c_2 \sqrt{\pi D} \text{erf}(\infty) = 0,$$

or $c_2 = -1/\sqrt{\pi D}$. The solution to (5.2.26)–(5.2.27) is therefore

$$y(z) = 1 - \text{erf} \left(\frac{z}{\sqrt{4D}} \right).$$

Consequently, the solution to the initial-boundary value problem (5.2.17), (5.2.23), (5.2.24) is

$$u(x, t) = 1 - \text{erf} \left(\frac{x}{\sqrt{4Dt}} \right). \quad \square$$

The preceding discussion centered on an algorithm for determining a group of stretching transformations under which a given PDE is invariant, and then using the similarity transformation to reduce the PDE to an ODE. There was no mention of the underlying reasons for the reduction other than to say that invariance or symmetry implies simplicity in the equations. The theory goes back to the late nineteenth century with the seminal work of S. Lie on the invariance of ODEs under one-parameter groups of transformations. Quite generally, invariance leads to simplifications in terms of new variables that are *invariants* of the family of transformations. Notice, for example, that the quantities defining the similarity transformation (5.2.9), namely, $ut^{-c/b}$ and $x/t^{a/b}$, are both invariants of (5.2.2):

$$\bar{u}\bar{t}^{-c/b} = ut^{-c/b}, \quad \frac{\bar{x}}{\bar{t}^{a/b}} = \frac{x}{t^{a/b}}.$$

As mentioned above, the stretching method can be extended to much more general transformations.

One should not conclude that the similarity method is a universal method that is applicable in all cases. Even though a given PDE admits a symmetry of the form (5.2.2), it may be impossible to find constants a , b , and c that force u , given by (5.2.9), to satisfy given initial or boundary conditions; that is, auxiliary conditions, especially on finite domains, sometimes break the symmetry. However, the method is applicable to many problems, and it is one of the basic techniques for solving PDEs on infinite domains.

EXERCISES

1. Verify equation (5.2.22).
2. Use the similarity method to solve the problem

$$\begin{aligned} u_t - u_{xx} &= 0, \quad x > 0, \quad t > 0, \\ u(x, 0) &= 0, \quad x > 0, \\ u_x(0, t) &= -1, \quad u(\infty, t) = 0, \quad t > 0. \end{aligned}$$

Give a physical interpretation of this problem.

3. Use the similarity method to derive the fundamental solution $K(x, t)$ of the diffusion equation $u_t = Du_{xx}$. *Hint:* Use the condition $\int_{\mathbb{R}} K(x, t) dx = 1$, $t > 0$.
4. A first-order ODE of the form $p - f(x, y) = 0$, where $p = y'(x)$, is said to be absolutely invariant under the stretching transformation $\bar{x} = x\varepsilon, \bar{y} = \varepsilon^a y$ if $\bar{p} - f(\bar{x}, \bar{y}) = p - f(x, y)$ for all ε in I , where I is an open interval containing $\varepsilon = 1$. In this case prove that the ODE can be reduced to a separable equation of the form $ds/s = dr/(F(r) - ar)$ for appropriately chosen r and s .
5. Associated with an Ornstein–Uhlenbeck process in stochastic dynamics is the PDE

$$u_t = \frac{1}{2}u_{xx} - xu_x,$$

where $u(x, t)$ is the probability that a particle starting at x in $(-a, a)$ at time $t = 0$ stays within $[-a, a]$ for all s with $0 \leq s \leq t$. Assume $u(-a, t) = u(a, t) = 0$ for all t . In the long-time limit show that the particle will escape the domain $[-a, a]$ with probability 1. *Hint:* Look at the steady state.

6. Consider the spherically symmetric diffusion equation

$$u_t = k \left(u_{rr} + \frac{2}{r} u_r \right).$$

Find a stretching transformation under which the PDE is invariant and show that

$$\frac{r}{\sqrt{4kt}} \quad \text{and} \quad ut^{3/2}$$

are invariants. Find all solutions of the form $u = At^{-3/2}U(s)$, where $s = r/\sqrt{4kt}$. Find the solution in the case that there is a point source initial condition at the origin at time $t = 0$. The answer is

$$u(r,t) = \frac{1}{(4\pi kt)^{3/2}} e^{-r^2/4kt},$$

which is the fundamental solution of the diffusion equation in three dimensions.

7. Consider the radially symmetric diffusion

$$u_t = k \left(u_{rr} + \frac{1}{r} u_r \right)$$

with a unit point source initial condition at the origin given at time $t = 0$. Use similarity methods to obtain the fundamental solution to the diffusion equation in two dimensions:

$$u(r,t) = \frac{1}{4\pi kt} e^{-r^2/4kt}.$$

5.3 Nonlinear Diffusion Models

It is clear that nonlinearities can arise in diffusion models from source terms. They also originate in other ways, as the next two examples show.

Example. (Insect Dispersal) The study of insect and animal dispersal leads to a natural nonlinearity in the flux term. If there is an increase in diffusion due to population pressure, it is reasonable to assume that the diffusion coefficient D is a function of the density u . Then Fick's law takes the form

$$\phi = -D(u)u_x. \tag{5.3.1}$$

Substituting into the conservation law $u_t + \phi_x = 0$ yields the nonlinear diffusion equation

$$u_t - (D(u)u_x)_x = 0. \tag{5.3.2}$$

Writing out the derivatives in (5.3.2) gives

$$u_t - D(u)_x u_x - D(u)u_{xx} = 0, \tag{5.3.3}$$

so the variable diffusion constant gives rise to a nonlinear advection term $-D(u)_x u_x$ or $-(D'(u)u_x)u_x$, which implies a propagation signal of speed $-D'(u)u_x$. \square

Example. (Heat Conduction) For linear heat transfer the governing equations are the conservation law for energy density

$$(\rho CT)_t + \phi_x = 0, \quad (5.3.4)$$

and Fourier's law for the heat flux:

$$\phi = -KT_x. \quad (5.3.5)$$

Here, $T = T(x, t)$ is the temperature, and the constants ρ , C , and K are the density, specific heat, and thermal conductivity of the medium, respectively (see Chapter 1). In many applications where the temperature range is limited, the specific heat may be regarded essentially as a constant. However, over wide temperature ranges the specific heat is not constant, but rather, is a function of the temperature itself; that is, $C = C(T)$. In this case (5.3.4) and (5.3.5) combine to give

$$(C(T)T)_t - \frac{K}{\rho} T_{xx} = 0, \quad (5.3.6)$$

which is a nonlinear diffusion equation for the temperature T . \square

Equations (5.3.2) and (5.3.6) are nonlinear diffusion models that have attracted considerable attention in the literature. The simplest assumptions are that the diffusion coefficient D in (5.3.2) and the specific heat C in (5.3.6) are power functions, or

$$D(u) = D_0 \left(\frac{u}{u_0} \right)^n, \quad D_0, u_0 \text{ constants}, \quad n > 0, \quad (5.3.7)$$

and

$$C(T) = C_0 \left(\frac{T}{T_0} \right)^n, \quad C_0, T_0 \text{ constants}, \quad n > 0. \quad (5.3.8)$$

In a different context, equation (5.3.2), along with the constitutive assumption (5.3.7), is called the *porous medium equation*, and it governs the motion of a fluid through a porous domain (Section 1.3). \square

Fourier's law (5.3.5) may be generalized in yet another direction, as the next example shows.

Example. (Non-Fickian Flux) In Fickian diffusion the flux is proportional to the gradient u_x , and therefore the flux becomes arbitrarily large when very steep

profiles are present. In some cases this unbounded response to large gradients seems nonphysical, and a bounded flux may be preferable. In a general case, where the flux depends on u_x , we suppose

$$\phi = -DF(u_x),$$

where F is a smooth, monotonically increasing function on \mathbb{R} with the property that it remains bounded at $\pm\infty$. Specifically, we suppose that $|F(w)| \leq M$, $F(0) = 0$, $F'(0) = 1$, $F'(w) > 0$, and $F'(w) \rightarrow 0$ as $w \rightarrow \pm\infty$. Typical functions having these properties are

$$F(w) = \arctan w, \quad F(w) = \tanh w, \quad F(w) = \frac{w}{\sqrt{1+w^2}}.$$

Notice that Fickian flux is then an approximation to the general nonlinear flux when the gradients are small. This is because, by Taylor's theorem

$$\phi = -D \left(F(0) + F'(0)u_x + \frac{1}{2}F''(0)u_x^2 + \dots \right) = -Du_x + \dots.$$

We refer to Logan (2001) for examples and references on bounded fluxes. Also see Kurganov & Rosenau (1997). \square

Example. In a low-temperature liquid phase of helium, heat transport is not governed by Fourier's law, but rather by the nonlinear Gorter–Mellink law,

$$\phi = -KT_x^{1/3}, \quad K \text{ const.} \quad (5.3.9)$$

Combined with (5.3.4), this gives the nonlinear heat conduction equation

$$T_t - k(T_x^{1/3})_x = 0,$$

where $k = K/C\rho$ and C and ρ are constants. \square

In the sequel, and in the exercises, properties of some of these nonlinear equations are examined using the similarity method.

Example. Consider the nonlinear diffusion equation

$$u_t - (uu_x)_x = 0, \quad (5.3.10)$$

which is a special case of (5.3.2) and (5.3.7). We assume that the constants have been scaled out of the problem. To compare the nonlinear equation (5.3.10) (*Boltzmann's problem*) with the linear diffusion equation, we consider (5.3.10) over the domain $x \in \mathbb{R}$, $t > 0$, subject to the initial condition of a unit point source applied at $x = 0$ at time $t = 0$; that is,

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

with

$$\int_{\mathbb{R}} u(x, t) dx = 1 \quad \text{for all } t > 0, \quad (5.3.11)$$

and

$$u(\pm\infty, t) = 0, \quad t > 0.$$

We know from Section 5.1 that the solution to the *linear* diffusion equation in this case is the fundamental solution. To determine a solution to the nonlinear problem, we proceed by the similarity method. It is easy to see that (5.3.10) is invariant under the stretching transformation

$$\bar{x} = \varepsilon^a x \quad \bar{t} = \varepsilon^b t \quad \bar{u} = \varepsilon^{2b-a} u.$$

Consequently, the similarity transformation is given by

$$u = t^{(2a-b)/b} y(z), \quad z = \frac{x}{t^{a/b}}. \quad (5.3.12)$$

Rather than immediately determine the ODE for y , we specialize a and b by imposing the condition (5.3.11). We have

$$t^{(2a-b)/b} \int_{\mathbb{R}} y\left(\frac{x}{t^{a/b}}\right) dx = t^{3(a/b)-1} \int_{\mathbb{R}} y(z) dz = 1. \quad (5.3.13)$$

This condition can be independent of t only provided that $a/b = \frac{1}{3}$. Then the similarity transformation (5.3.12) becomes

$$u = t^{-1/3} y(z), \quad z = xt^{-1/3}. \quad (5.3.14)$$

Substituting into the PDE (5.3.10) yields an ODE for $y = y(z)$:

$$3(yy')' + y + zy' = 0.$$

This equation may be integrated at once to give

$$3yy' + zy = \text{const.} \quad (5.3.15)$$

Because the solution must be symmetric about $z = 0$ [note $u_x(0, t) = 0$ and thus $y'(0) = 0$], we infer that the constant must be zero. Further, the boundary condition $u = 0$ at infinity must hold. Hence, from (5.3.15) it follows that we may take

$$y(z) = \begin{cases} \frac{A^2 - z^2}{6}, & \text{if } |z| < A \\ 0, & \text{if } |z| > A \end{cases}, \quad (5.3.16)$$

where A is a constant of integration, which can be determined from condition (5.3.13). To this end, we obtain

$$1 = \int_{\mathbb{R}} y(z) dz = \int_{-A}^A y(z) dz = \frac{2}{9} A^3,$$

which gives $A = (\frac{9}{2})^{1/3}$. Therefore, we constructed a piecewise smooth solution

$$u(x, t) = \begin{cases} \frac{1}{6}t^{-2/3}(A^2t^{2/3} - x^2), & \text{if } |x| < At^{1/3} \\ 0, & \text{if } |x| > At^{1/3} \end{cases}. \quad (5.3.17)$$

Time snapshots are shown in Figure 5.3. The solution (5.3.17) to the nonlinear

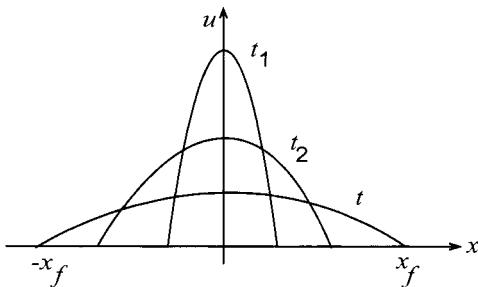


Figure 5.3 Time snapshots of the solution (5.3.17).

diffusion equation (5.3.10) is fundamentally different from the smooth fundamental solution $K(x, t)$ to the linear diffusion equation when a point source is present at $x = 0, t = 0$. The solution (5.3.17) represents a sharp wavefront $x_f = At^{1/3}$ propagating into the medium with speed

$$\frac{dx_f}{dt} = \frac{1}{3}At^{-2/3}.$$

Ahead of the wave the concentration u is zero, and at the front there is a jump discontinuity in the derivative of u . The wave slows down as t increases. Murray (2002) remarks, for example, that grasshoppers exhibit this type of dispersal behavior. In summary, nonlinear diffusion problems can behave quite differently from linear problems, even showing wave-like structure with propagating wavefronts. \square

Example. Consider the nonlinear diffusion model

$$uu_t - u_{xx} = 0, \quad x > 0, \quad t > 0, \quad (5.3.18)$$

subject to auxiliary conditions

$$u(x, 0) = 0, \quad x > 0, \quad (5.3.19)$$

$$u(\infty, t) = 0, \quad t > 0, \quad (5.3.20)$$

$$u_x(0, t) = -1, \quad t > 0. \quad (5.3.21)$$

This problem is a simplified version of the nonlinear heat conduction problem (5.3.6) and (5.3.8) subject to a heat flux condition (5.3.21) imposed at $x = 0$, and the condition that $u = 0$ initially, and at infinity. It is straightforward to see that (5.3.19) is invariant under the stretching transformation

$$\bar{x} = \varepsilon^a x, \quad \bar{t} = \varepsilon^b t, \quad \bar{u} = \varepsilon^{b-2a} u.$$

This invariance leads to the similarity transformation

$$u = t^{1-2a/b} y(z), \quad z = \frac{x}{t^{a/b}}. \quad (5.3.22)$$

Restrictions on the constants a and b are determined by the initial and boundary conditions. Computing $u_x(0, t)$ gives

$$u_x = t^{1-3a/b} y' \frac{x}{t^{a/b}}, \quad (5.3.23)$$

and therefore

$$u_x(0, t) = t^{1-3a/b} y'(0) = -1. \quad (5.3.24)$$

The left side of (5.3.24) cannot depend on t , so $a/b = \frac{1}{3}$. Consequently, the similarity transformation (5.3.22) becomes

$$u = t^{1/3} y(z), \quad z = \frac{xt^{-1/3}}{\sqrt{3}}. \quad (5.3.25)$$

(The inclusion of the factor $\sqrt{3}$ in the denominator of z makes the subsequent calculations more manageable.) Condition (5.3.20) implies that

$$y(\infty) = 0, \quad (5.3.26)$$

and the initial condition (5.3.19) can be written

$$\lim_{t \rightarrow 0} u(x, t) = \lim_{z \rightarrow \infty} t^{1/3} y(z) = \lim_{z \rightarrow \infty} \left(\frac{x}{\sqrt{3}z} \right) y(z) = 0,$$

because $z \rightarrow \infty$ as $t \rightarrow 0+$, for each fixed $x > 0$. Therefore, (5.3.19) yields the same condition as (5.3.20), namely (5.3.26). The flux condition (5.3.21) becomes

$$y'(0) = -\sqrt{3}. \quad (5.3.27)$$

Finally, substituting (5.3.25) into the PDE (5.3.18) yields the ODE

$$y'' - y(y - zy') = 0, \quad z > 0. \quad (5.3.28)$$

Therefore, the reduction is complete. The PDE (5.3.18) and three auxiliary conditions (5.3.19)–(5.3.21) have been transformed into a second-order ODE (5.3.28) subject to the two boundary conditions (5.3.26) and (5.3.27). Once the latter boundary value problem is solved for $y = y(z)$, the solution to

the original problem is given by (5.3.25). We point out that the three original boundary conditions coalesced into two conditions on y ; this coalescence had to occur because only two side conditions are required for the second-order equation (5.3.28). Had it not occurred, the boundary value problem for y would be overdetermined and the similarity method would have failed.

The solution to the nonautonomous equation (5.3.28) is by no means simple, and we must resort to numerical methods to determine $y(z)$. This retreat from analytic calculations is common in applied mathematics; it is more typical than not that analytic, closed-form solutions cannot be found, and hence numerical calculations are required in most problems. In the present case, (5.3.28) together with the conditions (5.3.26) and (5.3.27), form a boundary value problem on the semi-infinite interval $0 < z < \infty$, and we can use a *shooting method* to determine a numerical solution. The shooting method can be described briefly as follows. Most software packages for second-order differential equations require initial conditions, that is, conditions on both $y(0)$ and $y'(0)$. Here the initial condition $y(0)$ is not known. Therefore, we compute the numerical solution for several values of $y(0)$, along with the given value of $y'(0)$, until the numerical solution matches the right boundary condition $y(\infty) = 0$. Of course, the shooting method is more delicate on an infinite interval, and one must have confidence that a unique solution to the boundary value problem exists in the first place. See the Exercises for verification. Following the procedure described, we found that $y(0) = 1.5111$; the solution to the boundary value problem for y is plotted in Figure 5.4.

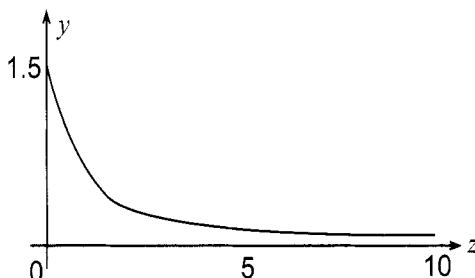


Figure 5.4 Plot of the numerical solution of the nonlinear equation (5.3.28) subject to the boundary conditions (5.3.26) and (5.3.27).

The solution to the original PDE (5.3.18) subject to (5.3.19)–(5.3.21) is

$u = u(x, t)$ given by (5.3.25). For each fixed time t_0 , the solution profile is

$$u(x, t_0) = t_0^{1/3} y \left(\frac{x}{\sqrt{3} t_0^{1/3}} \right),$$

which is the profile y in Figure 5.4 scaled horizontally by the factor $1/\sqrt{3}t_0^{1/3}$ and then amplified by the factor $t_0^{1/3}$. Hence, the time snapshots of the solution are geometrically similar, from whence the terminology *similarity method* can be justified. We remark that the class of self-similar solutions to a given problem is invariant itself; that is, solutions are mapped to solutions under the stretching transformation. \square

We present one additional example, leaving many of the detailed calculations to the reader.

Example. (The Porous Medium Equation) Consider

$$u_t = (u^m)_{xx}, \quad (m \geq 2).$$

Assuming a transformation

$$\bar{t} = \varepsilon t, \quad \bar{x} = \varepsilon^a x, \quad \bar{u} = \varepsilon^b u,$$

we find invariance when $b - 1 = mb - 2a$, and invariants u/t^b and x/t^a . Taking

$$u = t^b U(z), \quad z = xt^{-a},$$

we find

$$-azU' + bU = (U^m)''.$$

Choosing $b = -a$, we can integrate this equation to get

$$(U^m)' = -azU,$$

where we have set the constant of integration equal to zero. Solving this last equation then gives the one-parameter family of solutions

$$U(z) = \left(K - \frac{m-1}{2m} \frac{a}{2} z^2 \right)^{1/(m-1)}.$$

If we pick $m = 3$ and $a = 1$, then we obtain

$$u(x, t) = \frac{1}{t} \sqrt{K - \frac{x^2}{3t^2}}.$$

Under a point source initial condition, $u(x, 0) = \delta(x)$, the signal propagates outward in spacetime with support between the straight lines $x = \pm 3Kt$ with $u = 0$ for $x < -3Kt$ and $x > 3Kt$. The reader is invited to draw several

solution profiles, which are spreading arcs of circles with decreasing amplitude as time increases, in the region where the support lies. \square

Finally, the similarity method discussed in this and preceding sections can be extended to more general transformations. For example, traveling wave solutions of the form

$$u = y(z), \quad z = x - ct \quad (5.3.29)$$

are actually self-similar solutions and originate from an invariance property of the defining PDE under certain one-parameter families of transformations. We already observed (see, e.g., the Korteweg–deVries equation in Section 1.5) that a transformation of the form (5.3.29) reduces some PDEs to an ODE for $y(z)$.

Example. It is easily checked that the PDE

$$u_t = u_{xx} + uu_x$$

is invariant under the one-parameter family of transformations

$$\bar{x} = x - \varepsilon t, \quad \bar{t} = t, \quad \bar{u} = u + \varepsilon,$$

and that the invariants are

$$t, \quad u + \frac{x}{t}.$$

Taking

$$u = F(z) - \frac{x}{z}, \quad z = t,$$

the PDE reduces to

$$F'(z) + \frac{1}{z}F(z) = 0,$$

which is easily solved to obtain similarity solutions $F(z) = C/z$. Thus

$$u(x, t) = \frac{C - x}{t}. \quad \square$$

EXERCISES

1. Consider the differential equation (5.3.28) subject to the boundary conditions (5.3.26) and (5.3.27). Show that the transformation

$$w = z^2y, \quad v = z^2(y - zy')$$

reduces (5.3.28) to a first-order autonomous equation

$$\frac{dv}{dw} = \frac{v(2 - w)}{3w - v},$$

and show that the solution trajectory in the wv plane is along a separatrix connecting the origin to a saddle point at $w = 2, v = 6$. Thus show that y behaves like $y \sim 2/z^2$ as $z \rightarrow \infty$.

2. Use the similarity method to analyze the nonlinear diffusion problem

$$\begin{aligned} u_t - (u_x^{1/3})_x &= 0, \quad x > 0, \quad t > 0, \\ u_x(0, t) &= -1, \quad t > 0, \\ u(x, 0) &= 0, \quad x > 0, \quad u(\infty, t) = 0, \quad t > 0. \end{aligned}$$

Show that a similarity transformation is given by $u = \sqrt{t}y(z)$, $z = x/\sqrt{t}$ and that the problem reduces to

$$\begin{aligned} \frac{2}{3}(y')^{-2/3}y'' + zy' - y &= 0, \quad z > 0, \\ y'(0) &= 1, \quad y(\infty) = 0. \end{aligned}$$

Use the transformation $w = zy^{1/2}, v = z(y')^{1/3}$ to reduce the problem to a wv phase plane, and argue that the unique trajectory is along a separatrix connecting two critical points. Compute $y(0)$ and show that $y(z) \sim Cz^{-2}$ as $z \rightarrow \infty$, for some constant C , and hence for fixed t , that

$$u(x, t) \sim \frac{At^{3/2}}{x^2} \quad \text{as } x \rightarrow \infty.$$

Sketch a graph of $y = y(z)$.

3. Use the similarity method to find an analytic solution to the problem

$$\begin{aligned} xu_t - u_{xx} &= 0, \quad x > 0, \quad t > 0, \\ u(x, 0) &= 0, \quad x > 0, \\ u(\infty, t) &= 0, \quad t > 0, \\ u_x(0, t) &= -1, \quad t > 0. \end{aligned}$$

4. The second-order ordinary differential equation

$$y'' - G(x, y, y') = 0$$

is *invariant* under the one-parameter family of stretching transformations

$$\bar{x} = \varepsilon x, \quad \bar{y} = \varepsilon^b y$$

if there exists a smooth function $f(\varepsilon), f(1) = 1$, for which

$$\bar{r} - G(\bar{x}, \bar{y}, \bar{p}) = f(\varepsilon)(r - G(x, y, p)),$$

where $r = y''$ and $p = y'$. Prove that the second-order equation for y can be reduced to a first-order equation of the form

$$\frac{dv}{dw} = \frac{g(v, w) - (b-1)v}{v - bw},$$

where $w = y/x^b$ and $v = y'/x^{b-1}$, and g is a fixed function of v and w . (The variables w and v are called *Lie variables*, and the wv plane is called the *Lie plane*, after S. Lie.)

5. Consider the nonlinear PDE

$$uu_t + u_x^2 = 0.$$

Find the simplest, nontrivial stretching transformation under which the PDE is invariant, determine the similarity transformation, reduce the PDE to an ordinary differential equation, and solve to find the self-similar solutions. *Solution:* $u(x, t) = C \exp(4x^2/t)$.

6. Consider the porous medium equation

$$u_t = (u^m)_{xx}, \quad m > 2.$$

Discuss the behavior of similarity solutions of the form $u = t^{1/(m-1)}y(z)$, where $z = x/t$.

7. (Chapter 4 is required for this exercise.) A metallic rod of constant cross section initially at rest and occupying $x > 0$ undergoes longitudinal vibrations. The governing equations are

$$v_h - e_t = 0, \quad v_t - \frac{2Y}{\rho_0} ee_h = 0.$$

Here, t is time, h is a spatial coordinate attached to a fixed cross section with $h = x$ at $t = 0$, $v = v(h, t)$ is the velocity of the section h , $e = e(h, t)$ is the strain, or the lowest order approximation of the distortion at h ; $Y > 0$ is the stiffness (Young's modulus), and ρ_0 is the initial constant density. Beginning at $t = 0$ the back boundary ($h = 0$) is moved with velocity $-t$ [i.e., $v(0, t) = -t$, $t > 0$]. Use the method of stretching to determine the solution and the shock path.

8. The acoustic approximation equations are given by

$$u_t - v_x = 0, \quad v_t + uu_x = 0.$$

What is the stretching transformation under which the equations are invariant? Find all solutions of the form $u = U(z)$, $v = V(z)$, where $z = x/t$.

5.4 Reaction–Diffusion; Fisher’s Equation

Many natural processes inherently involve the mechanisms of both diffusion and reaction, and such problems are often modeled by *reaction–diffusion (R–D)* equations of the form

$$u_t - Du_{xx} = f(u), \tag{5.4.1}$$

where f is a given nonlinear function of u . In Section 1.3 we introduced, for example, the *Fisher equation*

$$u_t - Du_{xx} = ru \left(1 - \frac{u}{K}\right) \quad (5.4.2)$$

to model the diffusion of a species (e.g., an insect density u) when the reaction or growth term is the logistics law. Here, D is the diffusion constant, and r and K are the growth rate and carrying capacity, respectively. R-D equations have become one of the most important classes of nonlinear equations because of their occurrence in many biological and chemical (e.g., combustion) processes.

The Fisher equation is a prototype of R-D equations in one dimension, and we use it as a vehicle to introduce three mathematical techniques often employed to study R-D equations. The first is the question of existence of wavefront solutions, the second is a singular perturbation technique to determine the form of the wavefront, and the third is the stability of wavefront solutions under small perturbations of the waveform.

Let us rewrite the Fisher equation in dimensionless form by selecting dimensionless variables

$$\bar{t} = \frac{t}{r^{-1}}, \quad \bar{x} = \frac{x}{\sqrt{D/r}}, \quad \bar{u} = \frac{u}{K}.$$

Therefore, \bar{u} measures the population relative to the carrying capacity K , \bar{t} measures time relative to the growth rate r , and \bar{x} measures distances relative to $\sqrt{D/r}$, which is a diffusion lengthscale. In these variables Fisher's equation becomes

$$\bar{u}_{\bar{t}} - \bar{u}_{\bar{x}\bar{x}} = \bar{u}(1 - \bar{u}).$$

The constants scaled out of the problem. Now, to ease the notation we drop the overbars on the variables and study Fisher's equation in the form

$$u_t - u_{xx} = u(1 - u). \quad (5.4.3)$$

5.4.1 Traveling Wave Solutions

We investigate the question of whether equation (5.4.3) admits traveling wave solutions (abbreviated TWS) of wavefront type. To review, these are solutions of the form

$$u(x, t) = U(z), \quad \text{where } z = x - ct, \quad (5.4.4)$$

where c is a positive constant and $U(z)$ has the property that it approaches constant values at $z = \pm\infty$. The function U representing the waveform, and to be determined, should be twice continuously differentiable on \mathbb{R} . A priori, the

wave speed c is unknown, but must be determined as part of the solution to the problem. Substituting (5.4.4) into (5.4.3) yields the following second-order ordinary differential equation for $U(z)$

$$-cU' - U'' = U(1 - U), \quad -\infty < z < \infty, \quad (5.4.5)$$

where prime denotes the derivative d/dz . This equation cannot be solved in closed form, and the best approach to analyze the problem is in a two-dimensional phase plane. (See the Appendix at the end of the chapter.) In the standard way, this nonlinear differential equation can be reduced to a pair of first-order equations by introducing a new dependent variable V defined by $V = U'$. Then we obtain the autonomous system

$$\begin{aligned} U' &= V, \\ V' &= -cV - U(1 - U). \end{aligned} \quad (5.4.6)$$

The critical points of this system in the UV phase plane are $P:(0,0)$ and $Q:(1,0)$. The Jacobian matrix of the linearized system is

$$J(U, V) = \begin{pmatrix} 0 & 1 \\ 2U - 1 & -c \end{pmatrix}.$$

It is easy to check that the eigenvalues of $J(1, 0)$ are

$$\lambda_{\pm} = \frac{-c \pm \sqrt{c^2 + 4}}{2},$$

which are real and of opposite sign; therefore, $(1, 0)$ is a saddle point. The eigenvalues of $J(0, 0)$ are

$$\lambda_{\pm} = \frac{-c \pm \sqrt{c^2 - 4}}{2}.$$

Therefore $(0, 0)$ is a stable node if $c^2 \geq 4$ (the eigenvalues are real and both negative), and $(0, 0)$ is a stable spiral if $c^2 < 4$ (the eigenvalues are complex with negative real part). In phase space the parameter z must tend to $+\infty$ or $-\infty$ as the path enters or exits a critical point. Therefore, for the solution to approach constant states at infinity, we must show that there is a path connecting P and Q . Because P is stable and Q is unstable, the path can only connect Q to P . In summary, the path connecting Q to P , if it exists, is described by functions $U = U(z), V = V(z)$, with boundary conditions $U \rightarrow 1$ as $z \rightarrow -\infty$, and $U \rightarrow 0$ as $z \rightarrow +\infty$. The problem of showing that there is a TWS has been put in the context of finding a path connecting two critical points in the phase plane.

When $(0, 0)$ is a spiral, paths oscillate and U becomes negative. Because we want only physically meaningful solutions (U is a population density), we reject this case. Therefore we seek a unique path connecting the saddle point Q

at $z = -\infty$ to the stable node P at $+\infty$. Further, because Q is a saddle point, such a path must be one of the unstable manifolds exiting Q .

To draw a phase diagram we note that the nullclines consist of the U axis, where the vector field points directly downward, and the parabola $V = -(1/c)U(1 - U)$, where the vector field points to the left. The diagram is shown in Figure 5.5. Actually, the fact that the separatrix, or unstable manifold, leaving Q must enter the origin requires some proof, which we leave as an exercise. The separatrix connecting Q to P represents a monotone decreasing TWS whose profile is shown in Figure 5.6 for the case $c \geq 2$.

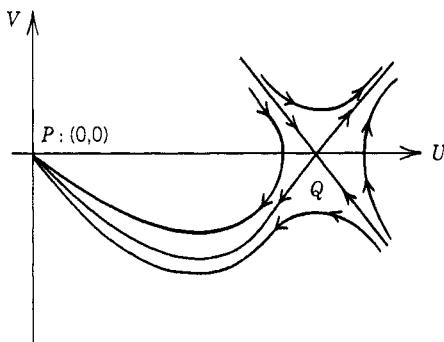


Figure 5.5 Phase portrait for (5.4.6) when $c \geq 2$.

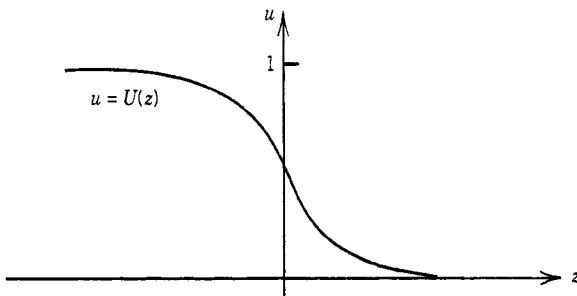


Figure 5.6 Traveling wave solution to Fisher's equation when $c \geq 2$. This corresponds to the separatrix connecting Q to P in Figure 5.5.

Theorem. For each $c \geq 2$ there exists a unique TWS $u(x, t) = U(x - ct)$ solution to equation (5.4.3) with the properties that U is monotonically decreasing on \mathbb{R} , $U(-\infty) = 1$, $U(\infty) = 0$, and $U'(\pm\infty) = 0$.

5.4.2 Perturbation Solution

We showed the existence of a TWS by appealing to a geometric argument in a phase plane, but we do not have a formula for the solution. Using a perturbation method we can obtain an approximation of the solution in the case $c \geq 2$. A perturbation method is an attempt to find an expansion for the solution of a problem in terms of a power series in some small parameter.

We seek an approximate solution to the boundary value problem

$$U'' + cU' + U(1 - U) = 0, \quad \infty < z < +\infty, \quad (5.4.7)$$

$$U(-\infty) = 1, \quad U(+\infty) = 0, \quad (5.4.8)$$

where $c \geq 2$. Because equation (5.4.7) is autonomous (independent of explicit dependence on z), if $U(z)$ is a solution, then so is $U(z + z_0)$ for any fixed constant z_0 . In other words, the solution curve may be translated to the left or right, and we still obtain a solution. Consequently, we can take

$$U(0) = \frac{1}{2}, \quad (5.4.9)$$

because the value at $z = 0$ can be chosen to be any number in the range of U . Now we identify a small parameter in the problem by taking

$$\varepsilon = \frac{1}{c^2} \leq 0.25.$$

Then the differential equation (5.4.7) becomes

$$\sqrt{\varepsilon}U'' + U' + \sqrt{\varepsilon}U(1 - U) = 0. \quad (5.4.10)$$

In (5.4.10) the dominant term is U' when ε is small, and the equation is, approximately, $U' = 0$, which has constant solutions; thus this approximation holds for large z since $U = 0$ for z large and positive, and $U = 1$ for z large and negative. In the large interval where U is changing from 1 to 0, a different dominant balance must occur. Therefore, we shrink this large interval to an order 1 interval by introducing the change of variables

$$s = \sqrt{\varepsilon}z = \frac{z}{c}, \quad g(s) = U\left(\frac{s}{\sqrt{\varepsilon}}\right). \quad (5.4.11)$$

The differential equation (5.4.10) then becomes

$$\varepsilon g'' + g' + g(1 - g) = 0, \quad (5.4.12)$$

where prime denotes d/ds . Thus, in terms of the order 1 variable s , the last two terms in (5.4.12) dominate the second derivative term. The auxiliary conditions (5.4.8) and (5.4.9) translate to

$$g(-\infty) = 1, \quad g(0) = \frac{1}{2}, \quad g(+\infty) = 0. \quad (5.4.13)$$

To find an approximate solution of (5.4.12)–(5.4.13) we assume a perturbation series

$$g(s) = g_0(s) + \varepsilon g_1(s) + \varepsilon^2 g_2(s) + \dots, \quad (5.4.14)$$

which is an expansion in powers of the small parameter ε , and where the coefficients g_0 , g_1 , g_2 , and so on, are to be determined. Substituting (5.4.14) into the differential equation (5.4.12) and auxiliary conditions (5.4.13), and then setting the coefficients of the powers of ε equal to zero gives the following sequence of problems for g_0 , g_1 , and so on:

$$g'_0 = -g_0(1 - g_0), \quad g_0(0) = \frac{1}{2}, \quad g_0(-\infty) = 1, \quad g_0(+\infty) = 0, \quad (5.4.15)$$

$$g'_1 = -g_1(1 - g_0) + g''_0, \quad g_1(-\infty) = g_1(0) = g_1(+\infty) = 0, \dots \quad (5.4.16)$$

Equation (5.4.15) can be solved to obtain

$$g_0(s) = (1 + e^s)^{-1}, \quad (5.4.17)$$

which meets all three boundary conditions in (5.4.15). Then (5.4.16) can be solved to get

$$g_1(s) = e^s (1 + e^s)^{-2} \ln \frac{4e^s}{(1 + e^s)^2}. \quad (5.4.18)$$

Consequently, in terms of the original variables U and z , and the wave speed c , the expansion (5.4.14) becomes

$$U(z) = \frac{1}{1 + e^{z/c}} + \frac{1}{c^2} e^{z/c} (1 + e^{z/c})^{-2} \ln \frac{4e^{z/c}}{(1 + e^{z/c})^2} + O\left(\frac{1}{c^4}\right). \quad (5.4.19)$$

Therefore, we have obtained an approximate, asymptotic form of the traveling wave $U(z)$ when $c \geq 2$. The approximation is most accurate for large c and least accurate for $c = 2$. It can be shown that the $O(1/c^4)$ term is uniform for all $z \in \mathbb{R}$ (i.e., the first two terms in the expansion give a uniform approximation that is valid for all z). Actually, the first term, $(1 + e^{z/c})^{-1}$, is remarkably close (within a few percent) to a numerically computed solution, even in the case $c = 2$.

Those familiar with perturbation methods recognize (5.4.10) as a standard singular perturbation problem because the small parameter ε multiplies the highest derivative. However, because the boundary conditions at infinity are automatically satisfied by the scaled problem (5.4.12), equation (5.4.10) is really just a regular perturbation problem, and the leading order term in the expansion gives a uniformly valid approximation on the entire domain.

5.4.3 Stability of Traveling Waves

One of the most important and interesting topics in applied mathematics is the stability of a given state. For example, the traveling wave solution for Fisher's equation represents a state, realized by a surface $u = U(x - ct)$ defined over spacetime. We say that a state is *asymptotically stable* if a small perturbation imposed on the state at some time (say, $t = 0$) eventually decays away and the system returns to its original state. We now show that the TWS to Fisher's equation is asymptotically stable to small perturbations in the moving coordinate frame of the wave, subject to the condition that the perturbation vanish outside some closed interval.

We first write Fisher's equation (5.4.3) in a moving coordinate frame by changing variables to

$$t = t, \quad z = x - ct,$$

where $c \geq 2$. Then (5.4.3) becomes

$$u_t - u_{zz} - cu_z = u(1 - u). \quad (5.4.20)$$

As is common practice, we are reusing the symbol u to denote the dependent variable as a function of t and z . We obtained a wavefront solution $U(z)$ to (5.4.20), so we consider solutions of the form

$$u = U(z) + V(z, t), \quad (5.4.21)$$

where V is a small perturbation of the state $U(z)$. We assume that

$$V(z, t) = 0 \quad \text{for } |z| \geq L, \quad \text{for some } L > 0,$$

which means that the perturbation vanishes outside some bounded interval in the moving frame. Figure 5.7 shows a typical perturbation V on the wave $U(z)$. We further suppose that $V(z, 0)$, the perturbation at time $t = 0$, is given. Substituting (5.4.21) into (5.4.20) gives a PDE for the perturbation, namely,

$$V_t - V_{zz} - cV_z = (1 - 2U)V - V^2. \quad (5.4.22)$$

Equation (5.4.22) is called the *nonlinear perturbation equation*, and it governs the small deviations V . However, because V is small, we may neglect the V -squared term in favor of the remaining lower-order terms. Then we obtain the *linearized perturbation equation*

$$V_t - V_{zz} - cV_z = (1 - 2U)V. \quad (5.4.23)$$

We look for solutions of the form

$$V(z, t) = v(z)e^{-\lambda t}. \quad (5.4.24)$$

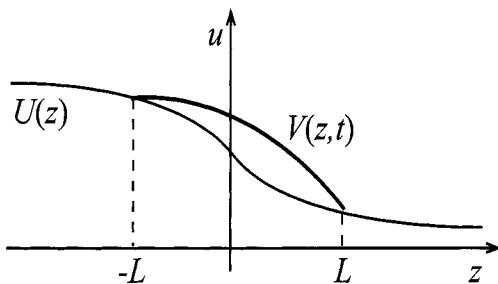


Figure 5.7 Traveling wave solution to Fisher's equation with an imposed small perturbation $V(z, t)$ in the moving coordinate frame.

Substituting (5.4.24) into (5.4.23) yields a second-order equation for the spatial part $v(z)$

$$v'' + cv' + [\lambda + 1 - 2U(z)] = 0, \quad (5.4.25)$$

and the boundary condition on V implies

$$v(-L) = v(L) = 0. \quad (5.4.26)$$

Equations (5.4.25) and (5.4.26) form an eigenvalue problem for $v(z)$, where the growth factor λ is interpreted as an eigenvalue. If there exist negative eigenvalues, (5.4.24) implies that the perturbation V will grow in time and the TWS $U(z)$ is unstable. However, if the only eigenvalues to (5.4.25)–(5.4.26) are positive, V will decay (exponentially) to zero as $t \rightarrow \infty$, and $U(z)$ is asymptotically stable.

To solve the eigenvalue problem (5.4.25)–(5.4.26) we apply the Liouville–Green transformation to eliminate the first derivative term and put the problem in normal form. In the present case the transformation is

$$v(z) = w(z)e^{-cz/2}. \quad (5.4.27)$$

Then

$$w'' + \left[\lambda - \left(2U(z) + \frac{c^2}{4} - 1 \right) \right] w = 0, \quad (5.4.28)$$

$$w(-L) = w(L) = 0. \quad (5.4.29)$$

A lot is known about boundary value problems in this form. For example, if $q(z) > 0$ and q is continuous, the eigenvalues λ of the boundary value problem $w'' + (\lambda - q(z))w = 0$, $w(-L) = w(L) = 0$, on the interval $-L < z < L$, are all positive. In our case

$$2U(z) + \frac{c^2}{4} - 1 \geq 2U(z) > 0 \quad \text{for } c \geq 2,$$

and consequently, the eigenvalue problem (5.4.28)–(5.4.29), and hence (5.4.25)–(5.4.26), has positive eigenvalues, giving the asymptotic stability result.

In summary, we showed that a small perturbation of finite extent, imposed in the traveling wave frame on the waveform $U(z)$, decays as $t \rightarrow \infty$, and the system returns asymptotically to its original state. This type of small perturbation argument is common in applied mathematics, and the stability calculation for Fisher's equation is representative of many of them.

5.4.4 Nagumo's Equation

Another important equation in mathematical biology is *Nagumo's equation*, which arises in the study of electrical impulses along nerves. But we can think of this equation ecologically as a model of growth–diffusion where the growth rate is given by the Allee model, $f(u) = u(u - a)(1 - u)$ (quantities are scaled), where $a < \frac{1}{2}$ and 1 is the carrying capacity. Nagumo's equation is

$$u_t = u_{xx} + u(u - a)(1 - u). \quad (5.4.30)$$

The Allee effect is a generalization of the logistic law with negative growth at small population densities below a value a . An interpretation is that at small densities animals are less able to find mates, resulting in negative growth.

We look for wavefront solutions of the form $u = U(z)$, where $z = x - ct$. Substituting into (5.4.30) and then reducing the resulting second-order ODE for U to a system, we obtain

$$U' = V, \quad (5.4.31)$$

$$V' = -cV - U(U - a)(1 - U). \quad (5.4.32)$$

There are three equilibria lying on the U axis in phase space: $(0, 0)$, $(a, 0)$, and $(1, 0)$. The Jacobian matrix is

$$J(U, V) = \begin{pmatrix} 0 & 1 \\ -([U - a](1 - U) + U(1 - U) - U(U - a)) & -c \end{pmatrix}.$$

It is easy to show that $J(0, 0)$ and $J(1, 0)$ have real eigenvalues of opposite sign, and therefore $(0, 0)$ and $(1, 0)$ are saddle points. The matrix $J(a, 0)$ has negative trace and positive determinant, and so $(a, 0)$ is asymptotically stable. The eigenvalues of $J(a, 0)$ are

$$\lambda = \frac{1}{2} \left(-c \pm \sqrt{c^2 - 4a(1 - a)} \right),$$

and so we conclude that $(a, 0)$ is a spiral if $c^2 < 4a(1 - a)$ and a node if $c^2 > 4a(1 - a)$. A phase diagram is drawn in Figure 5.8 in the spiral case. For

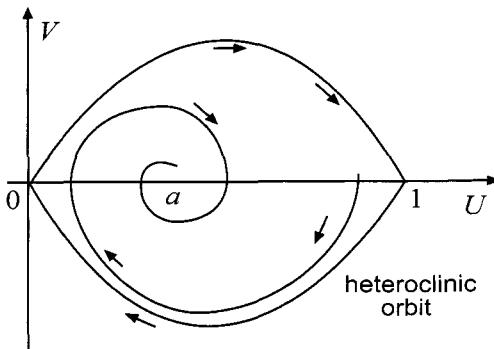


Figure 5.8 Phase plane diagram for Nagumo's equation.

a TWS to exist with constant states at infinity, there must be an orbit that connects two critical points.

We can find orbits by dividing the two equations in (5.4.31)–(5.4.32) to obtain

$$\frac{dV}{dU} = \frac{cV - U(U - a)(1 - U)}{V}.$$

Careful examination reveals that \$V\$ has the form \$V = bU(1 - U)\$, where \$b\$ is a constant. Substituting gives

$$b = \frac{1}{\sqrt{2}}, \quad c = \sqrt{2} \left(\frac{1}{2} - a \right).$$

Therefore we know the wave speed. We can solve the equation

$$V = U' = bU(U - 1)$$

using separation of variables to obtain

$$U(z) = \frac{1}{2} \left(1 - \tanh \left(\frac{z}{2\sqrt{2}} \right) \right), \quad (5.4.33)$$

which is the traveling wave corresponding to the heteroclinic orbit in the lower half-plane connecting \$(1, 0)\$ to \$(0, 0)\$.

EXERCISES

1. Show that the traveling wave solution \$U(z)\$ obtained for Fisher's equation in the case \$c \geq 2\$ has the property that the slower the wave moves (i.e., the smaller \$c\$ is), the steeper the wavefront.

2. Find an exact solution to Fisher's equation (5.4.3) of the form

$$U(z) = \frac{1}{(1 + ae^{bz})^d}, \quad z = x - ct.$$

To what wave speed c does this solution correspond?

3. Consider the nonlinear R-D equation

$$u_t - (uu_x)_x = u(1 - u).$$

Investigate the existence of wavefront-type TWS of speed $c = 1/\sqrt{2}$. Show that solutions exist of the form

$$u = U(z) = 1 - \exp\left(-\frac{z}{\sqrt{2}}\right), \quad z < 0; \quad u = 0, \quad z > 0 \quad (z = x - ct).$$

4. Equilibrium solutions $u = u(x)$ to Fisher's equation on the domain $0 < x < a$ with Dirichlet boundary conditions satisfy the nonlinear boundary value problem

$$u'' + u(1 - u) = 0, \quad 0 < x < a; \quad u(0) = u(a) = 0.$$

- (a) Show that nontrivial, positive, steady solutions exist provided that $a > \pi$. *Hint:* In the phase plane show that

$$a = 2 \int_0^{u_m} \frac{du}{(u_m^2 - u^2 + \frac{2}{3}(u^3 - u_m^3))^{1/2}},$$

where $u_m = u(a/2) = \max u$.

- (b) For $0 < a - \pi \ll 1$ show that the nontrivial solution takes the form $u = \frac{3}{4}(a - \pi) \sin(\pi x/a) + O((a - \pi)^2)$. *Hint:* Rescale the problem to obtain

$$\frac{d^2u}{dy^2} + a^2 u(1 - u) = 0, \quad 0 < y < 1; \quad u(0) = u(1) = 0,$$

and then assume $u(y) = u_0(y) + \varepsilon u_1(y) + \varepsilon^2 u_2(y) + \dots$, where $\varepsilon = a - \pi$.

- (c) Using the hint in part (a), show that $u_m \sim 1 - e^{-a/2}$ as $a \rightarrow \infty$.

5. Derive equation (5.4.33), which is the TWS to Nagumo's equation.
6. Determine the existence of smooth TWS of the problem

$$u_t = u_{xx} + f(u),$$

where

$$f(u) = u \text{ if } 0 \leq u \leq \frac{1}{2}, \quad f(u) = 1 - u \text{ if } \frac{1}{2} < u \leq 1,$$

and $u \rightarrow 1$ as $z \rightarrow -\infty$ and $u \rightarrow 0$ as $z \rightarrow +\infty$.

7. Find all traveling wave solutions $u = U(z)$, $z = x - ct$, to the problem

$$(u + \frac{1}{2}u^2)_t = u_{xx} - u_x,$$

with the property $U(-\infty) = 1$, $U(\infty) = 0$, $U(0) = \frac{1}{2}$.

8. Find all traveling wavefront solutions to the reaction–advection equation

$$u_t = -u_x + u(1 - u).$$

9. Consider the nonlinear advection–diffusion equation

$$u_t = Du_{xx} - g(u)_x.$$

- (a) Given $g(u) = u$, show that there are no nonconstant wavefront solutions.
- (b) Assuming that g is strictly convex ($g''(u) > 0$, or $g''(u) < 0$), show there exists a wavefront solution with constant states at infinity, and find an implicit formula for the solution.

10. Consider the system

$$u_t = Du_{xx} - u_x - \beta s_t, \quad s_t = -s + f(u),$$

where $f(0) = 0$, $f' > 0$, $f'' < 0$, and $D, \beta > 0$. Show that a wavefront solution exists with $s, u \rightarrow 0$ as $x \rightarrow +\infty$, and $u \rightarrow 1$ as $x \rightarrow -\infty$. What is $s(-\infty)$?

5.5 Advection–Diffusion; Burgers' Equation

The prototype for nonlinear advection–diffusion processes is *Burgers' equation*:

$$u_t + uu_x - Du_{xx} = 0. \tag{5.5.1}$$

The term uu_x represents a nonlinear advection or transport term, and Du_{xx} is a Fickian diffusion term. On one hand, the advection term has a *shocking-up* effect on an initial waveform, while the diffusion term attempts to smear out the solution. Thus (5.5.1) is a balance between these two effects. Equation (5.5.1) is often taken as the analog equation of compressible, viscous fluid flow; in that case the diffusion term is interpreted as a model viscosity term, which is also a dissipative term that tends to smear out signals. In Section 4.5.1 we derived Burgers' equation in a weakly nonlinear asymptotic limit of governing equations of viscous flow.

5.5.1 Traveling Wave Solution

The competition between the advection term and the dissipative term is best observed by deriving traveling wave solutions. Therefore, we look for twice continuously differentiable solutions of (5.5.1) of the form

$$u(x, t) = U(z) \quad z = x - ct, \quad (5.5.2)$$

where the wave speed c is to be determined and U approaches positive constant values u_1 at $z = +\infty$, and u_2 at $z = -\infty$, respectively; we assume $u_1 < u_2$. Substituting (5.5.2) into (5.5.1) gives an ODE for $u = U(z)$,

$$-cU' + UU' - DU'' = 0,$$

where prime denotes d/dz . This equation can be integrated immediately to obtain

$$-cU + \frac{1}{2}U^2 - DU' = A, \quad (5.5.3)$$

where A is a constant of integration. Writing (5.5.3) in standard form gives

$$U' = D^{-1}(\frac{1}{2}U^2 - cU - A). \quad (5.5.4)$$

To evaluate A and c we take the limit as $z \rightarrow -\infty$ and $z \rightarrow \infty$ in (5.5.4) to get

$$A = \frac{1}{2}u_1^2 - cu_1 = \frac{1}{2}u_2^2 - cu_2.$$

Hence

$$c = \frac{u_1 + u_2}{2}, \quad (5.5.5)$$

which gives the wave speed as the average of the two known states at infinity. Therefore, the constant of integration A is given by

$$A = -\frac{u_1 u_2}{2}.$$

The differential equation (5.5.4) therefore becomes

$$-2DU' = (U - u_1)(u_2 - U). \quad (5.5.6)$$

Separating variables and integrating gives

$$\frac{z}{D} = \frac{2}{u_2 - u_1} \ln \frac{u_2 - U}{U - u_1},$$

where the constant of integration is chosen so that $U(0) = c$. Note that c is the average value of u_1 and u_2 , and because (5.5.6) is autonomous, U can be chosen to be any value between u_1 and u_2 (solutions are translationally invariant). Solving for U then gives the wave profile

$$U(z) = u_1 + \frac{u_2 - u_1}{1 + \exp [(u_2 - u_1)z/2D]}, \quad (5.5.7)$$

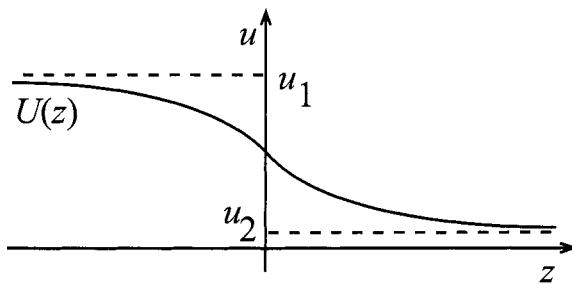


Figure 5.9 Traveling wave solution (5.5.7) to Burgers' equation.

where $z = x - ct$ is a moving coordinate. A plot of (5.5.7) is shown in Figure 5.9.

In summary, we constructed a positive monotonically decreasing TWS that approaches constant states at infinity. The speed c of the wave is the average value of the states u_1 and u_2 at infinity. A waveform of this shape would break and form a shock if the diffusion term were absent; its presence prevents the deformation, and the advection term and diffusion term are exactly balanced. The diffusion coefficient D affects the shape of the waveform; if D is large, there is a greater diffusive effect and the wave has a shallow gradient. On the other hand, if D is small, the gradient is steep. Analytically, we can see this as follows. The thickness of the wave (5.5.7) is defined to be the quantity $(u_2 - u_1) / \max |u'(z)|$, which one can easily compute to be $8D/(u_2 - u_1)$, because the maximum derivative occurs at $z = 0$. Because the profile in Figure 5.9 resembles the actual profile in a real shock wave when D is small, the TWS (5.5.7) is usually called the *shock structure* solution; in an actual, physical shock wave (not the mathematical shock wave defined by a discontinuity as in Chapter 3), the tendency of a wave to break is balanced by viscous effects in the region where the steep gradients occur. We discuss the limiting behavior as $D \rightarrow 0$ in the sequel.

5.5.2 Initial Value Problem

Now we change gears and address the initial value problem for Burgers' equation:

$$u_t + uu_x - Du_{xx} = 0, \quad x \in \mathbb{R}, \quad t > 0, \quad (5.5.8)$$

$$u(x, 0) = u_0(x), \quad x \in \mathbb{R}. \quad (5.5.9)$$

Amazingly enough, this problem can be reduced to the initial value problem for the linear diffusion equation and can be solved analytically. The reduction is accomplished by the Cole–Hopf transformation, a change of variables discovered independently by J. Cole and E. Hopf in the early 1950s. First, we introduce a (potential) function w defined by the equation

$$u = w_x. \quad (5.5.10)$$

Then (5.5.8) becomes

$$w_{xt} + w_x w_{xx} - D w_{xxx} = 0,$$

which can be integrated immediately with respect to x to obtain

$$w_t + \frac{1}{2} w_x^2 - D w_{xx} = 0. \quad (5.5.11)$$

Next introduce the dependent function v defined by

$$w = -2D \ln v. \quad (5.5.12)$$

It is easy to calculate the derivatives of w and substitute them into (5.5.11) to obtain

$$0 = w_t + \frac{1}{2} w_x^2 - D w_{xx} = -2Dv^{-1}(v_t - Dv_{xx}).$$

Therefore, the transformation

$$u = -\frac{2Dv_x}{v}, \quad (5.5.13)$$

which is a combination of (5.5.10) and (5.5.12), reduces (5.5.8) to

$$v_t - Dv_{xx} = 0, \quad (5.5.14)$$

which is the diffusion equation. Equation (5.5.13) is the *Cole–Hopf transformation*. Also, via (5.5.13), the initial condition (5.5.9) on u transforms into an initial condition on v . We have

$$u(x, 0) = u_0(x) = -\frac{2Dv_x(x, 0)}{v(x, 0)}.$$

Integrating both sides of this equation yields

$$v(x, 0) = v_0(x) = \exp\left(-\frac{1}{2D} \int_0^x u_0(y) dy\right). \quad (5.5.15)$$

Now, the solution to the initial value problem (5.5.14)–(5.5.15) for the diffusion equation is

$$v(x, t) = \frac{1}{\sqrt{4\pi Dt}} \int_{\mathbb{R}} v_0(\xi) e^{-(x-\xi)^2/4Dt} d\xi.$$

Therefore

$$v_x(x, t) = -\frac{1}{\sqrt{4\pi Dt}} \int_{\mathbb{R}} v_0(\xi) \frac{x - \xi}{2Dt} e^{-(x - \xi)^2/4Dt} d\xi.$$

Consequently, from (5.5.13), we obtain

$$u = -\frac{2Dv_x}{v} = \frac{\int_{\mathbb{R}} [(x - \xi)/t] v_0(\xi) \exp[-(x - \xi)^2/4Dt] dx}{\int_{\mathbb{R}} v_0(\xi) \exp[-(x - \xi)^2/4Dt] d\xi}, \quad (5.5.16)$$

where $v_0(\xi)$ is given by (5.5.15). It is now straightforward to see that the solution (5.5.16) of the initial value problem can be written as

$$u(x, t) = \frac{\int_{\mathbb{R}} [(x - \xi)/t] e^{-G(\xi, x, t)/2D} d\xi}{\int_{\mathbb{R}} e^{-G(\xi, x, t)/2D} d\xi}, \quad (5.5.17)$$

where

$$G(\xi, x, t) = \frac{(x - \xi)^2}{2t} + \int_0^{\xi} u_0(y) dy. \quad (5.5.18)$$

In summary, we found an analytic expression for the solution of the initial value problem associated with Burgers' equation.

EXERCISES

1. Show that the solution to the initial value problem

$$\begin{aligned} u_t + uu_x - Du_{xx} &= 0, \quad x \in \mathbb{R}, \quad t > 0, \\ u(x, 0) &= U \quad \text{if } x < 0; \quad u(x, 0) = 0 \quad \text{if } x > 0, \end{aligned}$$

can be written

$$Uu^{-1} = 1 + \frac{\exp[U(x - Ut/2)/2D] \operatorname{erfc}[-x/(2\sqrt{Dt})]}{\operatorname{erfc}[(x - Ut)/(2\sqrt{Dt})]},$$

where $\operatorname{erfc} = 1 - \operatorname{erf}$ is the complementary error function. Discuss the behavior of the solution for small D . Sketch time snapshots of the solution.

2. (a) Find an equilibrium solution $u = u^*(x)$ to Burgers' equation $u_t + uu_x = u_{xx}$ on the interval $0 < x < a$ subject to the boundary conditions $u = 0$ at $x = 0$ and $u = 1$ at $x = a$.
- (b) Let $u(x, t) = u^*(x) + U(x, t)$, where U is a small perturbation of the equilibrium solution, and where $U = 0$ at $x = 0, a$. Show that the linearized perturbation equation for U is given by

$$U_t = U_{xx} - (u^*U)_x, \quad 0 < x < a, \quad t > 0.$$

- (c) By assuming solutions of the form $U = \phi(x)e^{\lambda t}$, $\phi(0) = \phi(a) = 0$, show that λ must be negative and therefore u^* is asymptotically stable to small perturbations. Hint: Multiply the ϕ -equation by ϕ and integrate from $x = 0$ to $x = a$.

5.6 Asymptotic Solution to Burgers' Equation

In the preceding section we obtained a formula [equations (5.5.17)–(5.5.18)] for the solution to the initial value problem for Burgers' equation. We want to understand certain aspects of this solution. It is common practice in nonlinear partial differential equations to examine problems, or solutions to problems, in various limits. For example, if a parameter is present, one can inquire about the behavior in the limit of small or large values of the parameter. Or, one may ask about the long-time behavior of the solution. The first problem we consider is the behavior of the solution to the initial value problem for Burgers' equation in the limit of small values D of the diffusion constant. Intuitively, we expect the solution to approach the solution to the inviscid Burgers' equation in the limit of small D , and we show that our intuition is correct. We then investigate how Burgers' equation propagates a delta function (i.e., a point source), both in the limit of small D and in the limit of large D .

To determine the solution in specific cases it is necessary to approximate the integrals in formula (5.5.17) by asymptotic formulas [see, e.g., Murray 1984 or Bender & Orszag 1978]. We state one of the key results, Laplace's theorem, and give a heuristic proof.

Theorem. (Laplace) Consider the integral

$$f(s) = \int_{-\infty}^{\infty} g(\xi) e^{sh(\xi)} d\xi, \quad (5.6.1)$$

where g is continuous and h is in class C^2 . Let a be a single stationary point of h , or $h'(a) = 0$, and assume that $h''(a) < 0$. Then

$$f(s) \sim g(a) e^{sh(a)} \sqrt{-\frac{2\pi}{sh''(a)}} \quad \text{as } s \rightarrow \infty. \quad (5.6.2)$$

Formula (5.6.2) is an estimate giving the dominant behavior of the integral (5.6.1) for large values of the parameter s . We can observe instantly that the integrals in equation (5.5.17) have the form (5.6.1) if we make the identification $s = 1/D$, $h(\xi) = -G(\xi, x, t)/2$, and $g(\xi) = (x - \xi)/t$ in one case [the numerator of (5.5.17)], and $g(\xi) = 1$ in the other case (the denominator of (5.5.17)). Then (5.6.2) gives the behavior as $D \rightarrow 0$.

To demonstrate the validity of (5.6.2) in a special case we expand g and h about $\xi = a$, using Taylor's formula to get

$$\begin{aligned} g(\xi) &= g(a) + g'(a)(\xi - a) + \dots, \\ h(\xi) &= h(a) + h'(a)(\xi - a) + \frac{1}{2}h''(a)(\xi - a)^2 + \dots \\ &= h(a) + \frac{1}{2}h''(a)(\xi - a)^2 + \dots \end{aligned}$$

(We are assuming a higher degree of smoothness than stated in the theorem; as remarked, our argument contains a degree of hand-waving.) Then, keeping only leading order terms, we obtain

$$f(s) \sim g(a)e^{sh(a)} \int_{-\infty}^{\infty} e^{sh''(a)(\xi-a)^2/2} d\xi.$$

Making the change of variables

$$-z^2 = \frac{sh''(a)(\xi-a)^2}{2}, \quad dz = \sqrt{-\frac{sh''(a)}{2}} d\xi$$

gives

$$\begin{aligned} f(s) &\sim g(a)e^{sh(a)} \sqrt{-\frac{2}{sh''(a)}} \int_{-\infty}^{\infty} e^{-z^2} dz \\ &= g(a)e^{sh(a)} \sqrt{-\frac{2}{sh''(a)}}, \end{aligned}$$

which gives (5.6.2). \square

Using the result of Laplace's theorem we can estimate the integrals in equation (5.5.17). Making the identification mentioned above, we obtain

$$\int_{-\infty}^{\infty} e^{-G(\xi,x,t)/2D} d\xi \sim e^{-G(a,x,t)/2D} \sqrt{\frac{4D\pi}{|G''(a,x,t)|}} \text{ as } D \rightarrow 0,$$

where $\xi = a$ is a single stationary point (a maximum) of G , regarded as a function of ξ with parameters x and t . Later we deduce the consequences of the assumption that G has a single stationary point. From (5.5.18) we get

$$G'(\xi, x, t) = u_0(\xi) - \frac{(x-\xi)}{t}, \quad G''(\xi, x, t) = u'_0(\xi) + \frac{1}{t}.$$

Furthermore

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{x-\xi}{t} e^{-G(\xi,x,t)/2D} d\xi &\sim \frac{x-a}{t} e^{-G(a,x,t)/2D} \\ &\times \sqrt{\frac{4D\pi}{|G''(a,x,t)|}} \quad \text{as } D \rightarrow 0. \end{aligned}$$

Therefore, from (5.5.17) in Section 5.5, we conclude that

$$u(x, t) \sim \frac{x-a}{t} \quad \text{as } D \rightarrow 0,$$

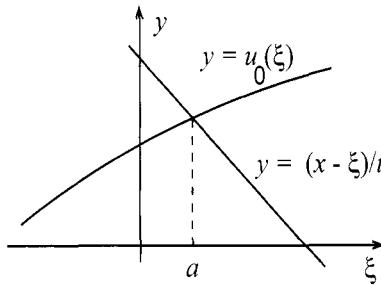


Figure 5.10

where a is the single stationary point of G and is given by $u_0(a) = (x - a)/t$. Therefore, in this case we can write the solution to the initial value problem to Burgers' equation as

$$u(x, t) \sim u_0(a) \quad \text{as } D \rightarrow 0,$$

where a is the single root of $x = a + tu_0(a)$. Notice that this is the same solution that we obtained for the initial value problem for the inviscid Burgers' equation (with $D = 0$); therefore, we have recovered the inviscid solution in the limit as $D \rightarrow 0$.

If the assumption that $\xi = a$ is a single stationary point of G holds true, then the graphs of $u_0(\xi)$ and $(x - \xi)/t$ must have a single intersection point. Figure 5.10 shows how the graph of the initial data u_0 may appear in this case, namely, as an increasing function of ξ . We recall that shocks will not develop for such initial data when propagated by the inviscid Burgers' equation.

5.6.1 Evolution of a Point Source

For the linear diffusion equation we showed that a point source (a delta function) at $x = 0, t = 0$ evolves according to the fundamental solution $K(x, t)$. Further, in Section 5.3 we calculated the evolution of such a source for a nonlinear diffusion equation. It is interesting to ask how such a point source is propagated by Burgers' equation, which is a nonlinear advection-diffusion equation. Consequently, we consider the initial value problem

$$u_t + uu_x = Du_{xx}, \quad x \in \mathbb{R}, \quad t > 0, \quad (5.6.3)$$

$$u(x, 0) = \delta(x), \quad x \in \mathbb{R}. \quad (5.6.4)$$

We assume that the source is located at $x = 0-$, that is, just to the left of the origin. We apply the solution (5.5.17)–(5.5.18) of Section 5.5. In the present

case the quantity G in (5.5.18) is

$$G(\xi, x, t) = \frac{(x - \xi)^2}{2t} + \int_0^\xi \delta(\hat{\xi}) d\hat{\xi}.$$

Consequently

$$G(\xi, x, t) = \frac{(x - \xi)^2}{2t} \quad \text{if } \xi > 0; \quad G(\xi, x, t) = \frac{(x - \xi)^2}{2t} - 1 \quad \text{if } \xi < 0.$$

Letting N and Δ denote the numerator and denominator in (5.5.17), we have

$$\Delta = e^{1/2D} \int_{-\infty}^0 e^{-(x-\xi)^2/4Dt} d\xi + \int_0^\infty e^{-(x-\xi)^2/4Dt} d\xi.$$

Now make the substitution $z = (x - \xi)/\sqrt{4Dt}$ to obtain, after some calculus and algebra

$$\Delta = (4Dt)^{1/2} \left[\pi + (e^{1/2D} - 1) \int_{x/\sqrt{4Dt}}^\infty e^{-z^2} dz \right]. \quad (5.6.5)$$

The numerator can be calculated in nearly the same manner to obtain

$$N = 2De^{-x^2/4Dt}(e^{1/2D} - 1). \quad (5.6.6)$$

Therefore, the solution to the initial value problem (5.6.3)–(5.6.4) can be written

$$u(x, t) = \frac{N}{\Delta} = \sqrt{\frac{D}{t}} (e^r - 1) \frac{e^{-x^2/4Dt}}{\sqrt{\pi} + (e^r - 1) \int_{x/\sqrt{4Dt}}^\infty e^{-z^2} dz}, \quad (5.6.7)$$

where $r = 1/2D$.

Observe that the solution in (5.6.7) has the form of a similarity solution $u = \sqrt{D/t}f(r, z)$, where $z = x/\sqrt{4Dt}$ (see Sections 5.2 and 5.3).

There are two different limits in which (5.6.7) can be examined. In the limit of small r (or large D) we expect diffusion to dominate, and in the limit of large r (or small D) we expect advection to dominate.

1. *Limit of Large D .* In this case we consider the ratio $u(x, t)/K(x, t)$, where u is given by (5.6.7) and K is the fundamental solution to the linear diffusion equation. We show that this ratio tends to unity as D increases, therefore proving the expected result that u behaves asymptotically like the solution to the linear diffusion equation. To this end, it is straightforward to see that we may write u/K in terms of r as

$$\frac{u}{K} = \frac{r^{-1}(e^r - 1)}{1 + \pi^{-1/2}(e^r - 1) \int_x^\infty \sqrt{r/2t} e^{-z^2} dz}. \quad (5.6.8)$$

The limit of the right side of (5.6.8) as $r \rightarrow 0$ is easily calculated to be 1 (Exercise 3), thereby proving

$$u(x, t) \sim K(x, t) \quad \text{as } D \rightarrow \infty. \quad (5.6.9)$$

2. *Limit of Small D .* This case is more difficult, and also more interesting. We first write the solution (5.6.7) of the initial value problem (5.6.3)–(5.6.4) as

$$u(x, t) = \sqrt{\frac{2}{t}} F(z, r), \quad (5.6.10)$$

where $z = x/\sqrt{2t}$ and

$$F(z, r) = \frac{1}{2\sqrt{r}} \frac{(e^r - 1)e^{-rz^2}}{\sqrt{\pi} + (e^r - 1) \int_{z\sqrt{r}}^{\infty} e^{-\zeta^2} d\zeta}.$$

Clearly, we may replace $e^r - 1$ by e^r for large r and write

$$F(z, r) \sim \frac{1}{2\sqrt{r}} \frac{e^{r(1-z^2)}}{\sqrt{\pi} + e^r \int_{z\sqrt{r}}^{\infty} e^{-\zeta^2} d\zeta} \quad \text{as } r \rightarrow \infty. \quad (5.6.11)$$

We can now determine the behavior of $F(z, r)$ for large r over different ranges of z .

If $z < 0$, the integral in the denominator of (5.6.11) approaches $\sqrt{\pi}$, and it is routine to show

$$F(z, r) \sim \frac{e^{-rz^2}}{2\sqrt{\pi r}} \quad \text{as } r \rightarrow \infty.$$

Therefore

$$F(z, r) \sim 0 \quad \text{as } r \rightarrow \infty \quad \text{for } z < 0. \quad (5.6.12)$$

If $z > 0$, we require the asymptotic approximation $\int_{\eta}^{\infty} e^{-\zeta} d\zeta \sim e^{-\eta^2}/2\eta$ as $\eta \rightarrow \infty$, which is the leading order approximation of the complementary error function erfc for large η (Abramowitz & Stegun 1964). Then (5.6.11) becomes

$$F(z, r) \sim \frac{z}{1 + 2\sqrt{\pi r}ze^{r(z^2-1)}} \quad \text{as } r \rightarrow \infty \quad \text{for } z > 0. \quad (5.6.13)$$

Now, if $z > 1$, then $F(z, r) \sim 0$ as $r \rightarrow \infty$, and if $0 < z < 1$, then $F(z, r) \sim z$ as $r \rightarrow \infty$.

In summary, we showed that $F(z, r) \sim 0$ as $r \rightarrow \infty$ for $z < 0$ and $z > 1$. In the interval $0 < z < 1$ we have $F(z, r) \sim z$. Translating this information back to the solution u given by (5.6.10), we have, as $D \rightarrow 0$,

$$u(x, t) \sim \frac{x}{t} \quad \text{for } 0 < x < \sqrt{2t}; \quad u(x, t) \sim 0 \quad \text{otherwise.} \quad (5.6.14)$$

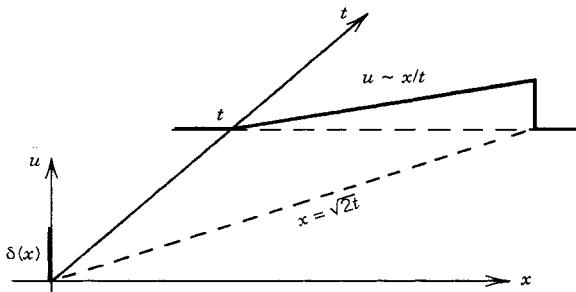


Figure 5.11 Evolution of a delta function initial condition, propagated by Burgers' equation, in the case of a small diffusion constant D .

Figure 5.11 shows the limiting solution (5.6.14) for a small diffusion constant D ; as expected, we obtain a shock-like structure where diffusion plays a small role compared to nonlinear advection. Near $z = 1$, that is, near the wavefront $x = \sqrt{2t}$, it can be shown that there is a steep transition region of order $O(D)$ (Whitham 1974, p. 105).

The preceding analysis for Burgers' equation can be carried out because a closed-form solution to the initial value problem is available. For other nonlinear equations, like Fisher's equation, the analysis is often more difficult because analytic solutions are not known.

EXERCISES

1. Verify (5.6.5) and (5.6.6).
2. In Laplace's theorem let $h(\xi) = 1 - \xi^2$ and sketch the graph of $e^{s(1-\xi^2)}$ for various values of s . Note that for large s the main contribution to the integral $f(s)$ comes from the neighborhood of $\xi = 0$, where h has a local maximum. Estimate the integral

$$f(s) = \int_{-\infty}^{\infty} \cos^2 \xi e^{s(1-\xi^2)} d\xi$$

for large s .

3. Show that the limit as $r \rightarrow 0$ of the right side of (5.6.8) is unity. Hint: Expand e^r in a Taylor series and note that the integral $\int_{x\sqrt{r/2t}}^{\infty} e^{-z^2} dz$ tends to a constant.
4. Verify (5.6.12) and (5.6.13).
5. Show that near the wavefront ($z = 1$), formula (5.6.10) for $F(z, r)$ becomes

$$F(z, r) \sim \frac{1}{1 + 2\sqrt{\pi r} e^{2r(z-1)}} \quad \text{as } r \rightarrow \infty.$$

Hint: Note that $z^2 - 1 \sim 2(z - 1)$.

6. Derive the solution (5.6.7) using similarity methods.

Appendix: Dynamical Systems

This appendix is a brief presentation of phase plane analysis in two-dimensional nonlinear dynamical systems. It is meant to be a tool that allows quick reference when reading certain parts of the book.

A two-dimensional nonlinear autonomous system has the form

$$\frac{dx}{dt} = p(x, y), \quad \frac{dy}{dt} = q(x, y), \quad (5.7.1)$$

where p and q are given functions that are assumed to have continuous derivatives of all orders. By a *solution* of (5.7.1) we mean a pair of smooth functions $x = x(t), y = y(t)$ that satisfy the differential equations (5.7.1) for all t in some interval I . The interval I is often the whole real line. Graphically, we represent the solution as a curve in the xy plane, called the *phase plane*. A solution curve is called an *orbit*, *path*, or *trajectory* of (5.7.1). The independent variable t is regarded as a parameter along the curve and is interpreted as time. The orbits have a natural positive direction to them, namely, the direction in which they are traced out as the time parameter t increases; to indicate this direction an arrow is placed on a given orbit. Because the system is autonomous (t does not appear on the right sides), the solution is invariant under a time translation; therefore, the time t may be shifted along any orbit. A constant solution $x(t) = x_0, y(t) = y_0$ to (5.7.1) is called an *equilibrium solution*, and its orbit is a single point (x_0, y_0) in the phase plane. Clearly, such points satisfy the algebraic relations

$$p(x, y) = 0 \quad q(x, y) = 0. \quad (5.7.2)$$

Points that satisfy (5.7.2) are called *critical points* (also, *rest points* and *equilibrium points*), and each such point represents an equilibrium solution. It is evident that no other orbit can pass through a critical point at finite time t ; otherwise, uniqueness of the initial value problem is violated. For the same

reason, no orbits can cross. The totality of all the orbits of (5.7.1) and critical points, graphed in the phase plane, is called the *phase diagram*. The qualitative behavior of the phase diagram is determined to a large extent by the location of the critical points and the local behavior of orbits near those points. The *Poincaré—Bendixson theorem* in two dimensions characterizes the behavior of the possible orbits of (5.7.1):

- (a) An orbit cannot approach a critical point in finite time; that is, if an orbit approaches a critical point, then, necessarily, $t \rightarrow \pm\infty$.
- (b) As $t \rightarrow \pm\infty$, an orbit either approaches a critical point, moves on a closed path, approaches a closed path, or leaves every bounded set. A closed orbit is a *periodic solution*.

The Poincaré—Bendixson theorem does not hold in three or more dimensions.

In principle, orbits can be found by integrating the differential relationship

$$\frac{dy}{dx} = \frac{q(x, y)}{p(x, y)}, \quad (5.7.3)$$

which comes from dividing the two equations in (5.7.1). When this is done, information about how a given orbit depends on the time parameter t is lost; however, this is seldom crucial.

The right sides of (5.7.1) define a vector field $\langle p, q \rangle$ in the phase plane; the orbits are the curves that have this vector field as their tangents. Thus the orbits are the integral curves of (5.7.3). The loci $p(x, y) = 0$ and $q(x, y) = 0$, where the vector field is vertical and horizontal, respectively, are called the *nullclines*. Determining the phase portrait of (5.7.1) is generally facilitated by graphing the nullclines and determining the direction of the vector field $\langle p, q \rangle$ in the regions separated by the nullclines.

The following two results of Bendixson and Poincaré, respectively, are also helpful.

- (a) If $p_x + q_y$ is of one sign in a region of the phase plane, the system (5.7.1) cannot have a closed orbit in that region.
- (b) A closed orbit of (5.7.1) must surround at least one critical point.

Critical points are also classified as to their stability. A critical point is *stable* if each orbit sufficiently near the point at some time t_0 remains in a prescribed circle about the point for all $t > t_0$. If this is not the case, the critical point is *unstable*. A critical point is *asymptotically stable* if it is stable and every orbit sufficiently near the point at some time t_0 approaches the point as $t \rightarrow \infty$. An asymptotically stable critical point is called an *attractor*.

Linear Systems

A linear system has the form

$$x' = ax + by, \quad (5.7.4)$$

$$y' = cx + dy. \quad (5.7.5)$$

Hereafter, we use a prime to denote the time derivative. We assume up front that $ad - bc$ is nonzero; then the only critical point of (5.7.4)–(5.7.5) is the origin $x = 0, y = 0$. We write the system in matrix form as

$$\mathbf{x}' = A\mathbf{x}, \quad (5.7.6)$$

where $\mathbf{x} = (x, y)^T$ (the superscript T means transpose) is the vector of unknowns and A is the coefficient matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

By assumption, $\det A \neq 0$. Solutions of (5.7.6) are obtained by assuming

$$\mathbf{x}(t) = \mathbf{v}e^{\lambda t}, \quad (5.7.7)$$

where \mathbf{v} is a constant vector and λ is a constant, both to be determined. Substituting this form into (5.7.6) yields the algebraic eigenvalue problem

$$A\mathbf{v} = \lambda\mathbf{v}. \quad (5.7.8)$$

Any eigenpair (λ, \mathbf{v}) of (5.7.8) gives a solution of (5.7.6) of the form (5.7.7). Therefore, if $(\lambda_1, \mathbf{v}_1)$ and $(\lambda_2, \mathbf{v}_2)$ are two eigenpairs with λ_1 and λ_2 distinct, all solutions of (5.7.6) are given by the linear combination

$$\mathbf{x}(t) = c_1\mathbf{v}_1 \exp(\lambda_1 t) + c_2\mathbf{v}_2 \exp(\lambda_2 t), \quad (5.7.9)$$

where c_1 and c_2 are arbitrary constants. This includes the case when λ_1 and λ_2 are complex conjugates; then real solutions may be found by taking the real and imaginary parts of (5.7.9). If $\lambda_1 = \lambda_2 = \lambda$, there may not be two linear independent eigenvectors \mathbf{v}_1 and \mathbf{v}_2 ; if two independent eigenvectors exist, (5.7.9) remains valid; if not (there is a single eigenvector \mathbf{v}), the general solution to (5.7.6) is

$$\mathbf{x}(t) = c_1\mathbf{v} \exp(\lambda t) + c_2(\mathbf{w} + \mathbf{v}t) \exp(\lambda t), \quad (5.7.10)$$

for some constant vector \mathbf{w} that satisfies $(A - \lambda I)\mathbf{w} = \mathbf{v}$.

Therefore, we may catalog the different types of solutions of the linear system (5.7.6), depending on the eigenvalues and eigenvectors of the coefficient matrix A . The results in the following summary come directly from the forms of the general solution (5.7.9) or (5.7.10).

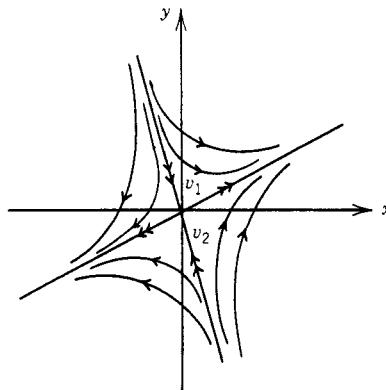


Figure 5.12 Saddle point.

Case 1. If the eigenvalues are real and have opposite signs, the critical point $(0,0)$ is a *saddle point*, and a generic phase portrait is shown in Figure 5.12. The two orbits entering the origin are the two *stable manifolds*, and the two orbits exiting the origin are the two *unstable manifolds*; the directions of these manifolds at the critical point are determined by the two eigenvectors. The stable manifolds correspond to the negative eigenvalue and the unstable manifolds correspond to the positive eigenvalue. These special manifolds are called *separatrices*. A saddle point is unstable.

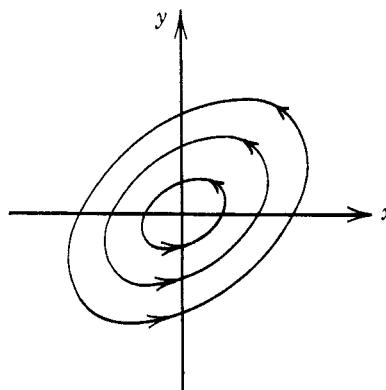


Figure 5.13 Center.

Case 2. If the eigenvalues are purely imaginary, the orbits form closed curves (ellipses) representing periodic solutions, and the origin is a *center* (see Fig-

ure 5.13). A center is stable.

Case 3. If the eigenvalues are complex (conjugates) and not purely imaginary, the orbits spiral into, or out of, the origin, depending on whether the real part of the eigenvalues is negative or positive, respectively. In this case the origin is called a *spiral point* or a focus (see Figure 5.14). If the real part of the eigenvalues is negative, the spiral point is an attractor, which is asymptotically stable; if the real part is positive, the point is unstable.

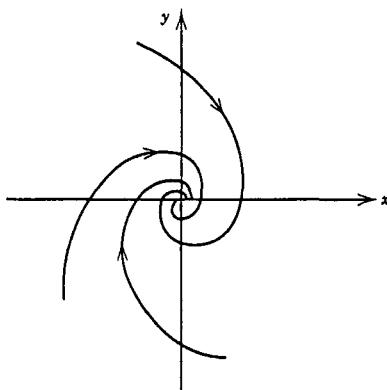


Figure 5.14 Asymptotically stable spiral.

Case 4. If the eigenvalues are real, are distinct, and have the same sign, the origin is classified as a *node*. In this case all the orbits enter the origin if the eigenvalues are negative (giving an attractor), and all the orbits exit the origin

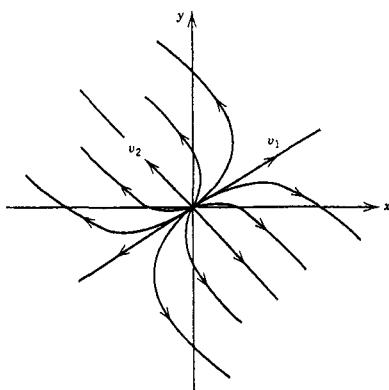


Figure 5.15 Unstable node.

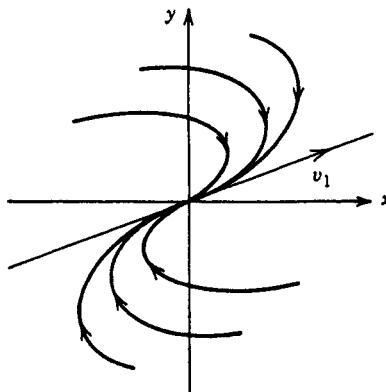


Figure 5.16 Degenerate stable node.

if the eigenvalues are positive (giving an unstable node). The detailed orbital structure near the origin is determined by the eigenvectors. One of them defines the direction in which orbits enter (or exit) the origin, and the other defines the direction approached by the orbits as t approaches minus, or plus, infinity (see Figure 5.15).

Case 5. If the eigenvalues are real and equal, the critical point at the origin is still classified as a node, but it is of different type than in case 4 because there may be only one eigenvector. In this case the node is called *degenerate* (see Figure 5.16). The node is an attractor if the eigenvalue is negative, and it is unstable if the eigenvalue is positive.

The eigenvalues of A are calculated as the roots of the characteristic equation

$$\lambda^2 - \text{tr } A\lambda + \det A = 0$$

where $\text{tr } A$ denotes the trace of A and $\det A$ is the determinant of A . The corresponding eigenvectors are the solution to the homogeneous equation

$$(A - \lambda I)\mathbf{v} = \mathbf{0}.$$

A useful result is that $(0, 0)$ is asymptotically stable if, and only if, $\text{tr } A < 0$ and $\det A > 0$.

Nonlinear Systems

Now we return to the nonlinear system (5.7.1). The key idea is that (5.7.1) can be studied by examining a linearization near the critical points. Let $x = x_0, y = y_0$ be an isolated critical point of (5.7.1), meaning that there

is a neighborhood of (x_0, y_0) that contains no other critical points. The local structure of the orbits near this critical point for the nonlinear system (5.7.1) can be determined, under fairly broad conditions, by examining the linearized system at that point. Let $\bar{x}(t)$ and $\bar{y}(t)$ denote small perturbations about the equilibrium (x_0, y_0) , with

$$x(t) = x_0 + \bar{x}(t), \quad y(t) = y_0 + \bar{y}(t).$$

Substituting into the nonlinear system gives

$$\bar{x}' = p(x_0 + \bar{x}, y_0 + \bar{y}), \quad \bar{y}' = q(x_0 + \bar{x}, y_0 + \bar{y}).$$

If the right sides are expanded in Taylor series, we obtain

$$\begin{aligned} \bar{x}' &= p(x_0, y_0) + p_x(x_0, y_0)\bar{x} + p_y(x_0, y_0)\bar{y} + \text{higher order terms,} \\ \bar{y}' &= q(x_0, y_0) + q_x(x_0, y_0)\bar{x} + q_y(x_0, y_0)\bar{y} + \text{higher order terms.} \end{aligned}$$

But $p(x_0, y_0) = q(x_0, y_0) = 0$, so to leading order the perturbations satisfy the linear system

$$\begin{pmatrix} \bar{x}' \\ \bar{y}' \end{pmatrix} = \begin{pmatrix} p_x(x_0, y_0) & p_y(x_0, y_0) \\ q_x(x_0, y_0) & q_y(x_0, y_0) \end{pmatrix} \begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix}. \quad (5.7.11)$$

We assume that the coefficient matrix

$$J(x_0, y_0) = \begin{pmatrix} p_x(x_0, y_0) & p_y(x_0, y_0) \\ q_x(x_0, y_0) & q_y(x_0, y_0) \end{pmatrix}.$$

which is called the *Jacobian matrix* at (x_0, y_0) , has a nonzero determinant. The linearized system for the small perturbations, whose behavior is determined by the eigenvalues and eigenvectors of J , dictate the local behavior of orbits for the nonlinear system near (x_0, y_0) .

The main result is as follows: Let (x_0, y_0) be an isolated critical point of the nonlinear system (5.7.1). Suppose that $\det J(x_0, y_0) \neq 0$ and that $J(x_0, y_0)$ does not have purely imaginary eigenvalues. Then (5.7.1) has the same qualitative orbital structure near (x_0, y_0) as the linearized system has near $(0, 0)$.

By *qualitative structure* we mean the same stability characteristics and the same nature of critical point (saddle, node, or focus). The case that does not extend to the nonlinear system is that of a center for the linear system. In this case one must examine the higher-order terms to determine the nature of the critical point of the nonlinear system. If the Jacobian matrix for the linearized system has a zero eigenvalue, then higher-order terms play a crucial role and the nature of the critical point of the nonlinear system may differ from a node, saddle, focus, or center. For example, it may have nodal structure on one side and a saddle structure on the other. A careful analysis is required in the case when $\det J = 0$.

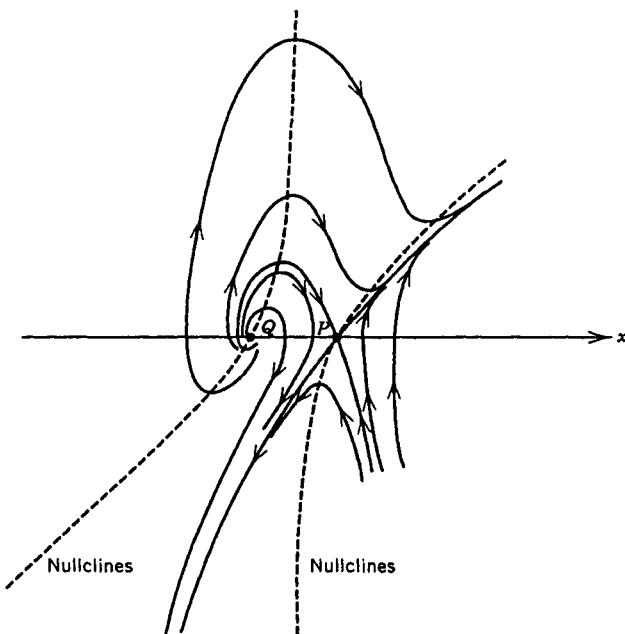


Figure 5.17 Phase portrait of the nonlinear system (5.7.12). $P : (0, 0)$ and $Q : (-1, 0)$ are the two critical points, a saddle and an unstable spiral, respectively. The nullclines $y' = 0$ are represented by the dashed lines.

The phase diagram of (5.7.1) is determined by finding all the critical points, analyzing their nature and stability, and then examining the global behavior and structure of the orbits. For example, it is an interesting problem to determine whether an orbit connects two critical points in the system; such connections are called *heteroclinic orbits*. An orbit connecting a critical point to itself is a *homoclinic orbit*.

Example. Consider the nonlinear system

$$x' = y, \quad y' = x - y + x^2 - 2xy. \quad (5.7.12)$$

There are two critical points $P : (0, 0)$ and $Q : (-1, 0)$. The Jacobi matrix $J(0, 0)$ at P is

$$\begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix},$$

which has eigenvalues $(-1 \pm \sqrt{5})/2$. These are real with opposite signs, and

therefore P is a saddle point. The Jacobi matrix $J(-1, 0)$ at Q is

$$\begin{pmatrix} 0 & 1 \\ 1 - 2x - 2y & -1 - 2x \end{pmatrix}_{(0,0)} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

The eigenvalues are $(1 \pm i\sqrt{3})/2$, which are complex with positive real part. Thus Q is an unstable spiral point. The nullclines are $y = 0$ (where $x' = 0$ and the vector field is vertical) and $y = x(1+x)/(1+2x)$ (where $y' = 0$ and the vector field is horizontal). A phase diagram is shown in Figure 5.17. The nullclines $y' = 0$ are shown dotted. There is a heteroclinic orbit exiting the spiral point Q and entering the saddle point P along a separatrix, or one of its stable manifolds. \square

All computer algebra systems (MATLAB, Mathematica, Maple, and others) have packages, or easily used programs, that plot phase diagrams in two dimensions. An easy to use, outstanding ODE solver and graphics package has been developed in MATLAB by John Polking (Polking 2004).

EXERCISES

- For the following linear systems, find the type of the critical point and its stability, find the general solution, and sketch a phase diagram.
 - $x' = -\frac{3}{2}x + \frac{1}{2}y, \quad y' = x - y.$
 - $x' = 4x - 3y, \quad y' = 6x - 7y.$
 - $x' = -2x - 3y, \quad y' = 3x - 2y.$
 - $x' = -3y, \quad y' = 6x.$
- For the following nonlinear systems, find all the equilibria, analyze their stability, draw the nullclines and sketch a phase diagram.
 - $x' = x - xy, \quad y' = y - xy.$
 - $x' = y, \quad y' = x^2 - 1 - y.$
 - $x' = y + (1-x)(2-x), \quad y' = y - ax^2, \quad a > 0.$
 - $x' = x - y, \quad y' = -y + (5x^2)/(4 + x^2).$
- State why the system $x' = x + x^3 - 2y, \quad y' = y^5 - 3x$ has no periodic solutions (closed cycles).

Reference Notes. A good introduction to the diffusion equation and random walks is given in the first chapter of Zauderer (2006). Also see Murray (2002, 2003) and Okubo & Levin (2001). An elementary discussion of the general similarity method is presented in Logan (1987). See also Dresner (1983) and Rogers

& Ames (1989); the latter contains an extensive bibliography as well as numerous, recent applications. The similarity method as it applies to problems in the calculus of variations is discussed in Logan (1977, 1987). For a treatment of Fisher's equation, see Murray (1977, 2002, 2003). Canosa (1973) is a classic paper on TWS to Fisher's paper. A good source for Burgers' equation is Whitham (1974); Kreiss and Lorenz (1989) discuss questions of uniqueness and existence of solutions to boundary value problems associated with Burgers' equation.

For a detailed discussion of phase plane phenomena, the reader may consult one of the many excellent books on the subject; see, for example, Strogatz (1994) or Hirsch et al. (2004). Most sophomore-level differential equation texts give introductory material (see, e.g., Logan 2006b), as do texts on mathematical biology, where phase plane methods are essential (Edelstein-Keshet 2005, Brauer & Castillo-Chavez 2001). In fact, mathematical biology texts are one of the best sources for elementary approaches and interesting applications.

6

Reaction–Diffusion Systems

During the last few decades there has been intense research devoted to the study and application of reaction–diffusion equations. One of the motivating factors was a celebrated paper by A. Turing in 1952. He proposed that reaction–diffusion could serve as a model for the chemical basis of morphogenesis, or the development of form and pattern in the embryological stage of various organisms. His observation was that spatially inhomogeneous patterns can be generated by diffusion-driven instabilities. It is not surprising, therefore, that tremendous interest was spawned in the biological sciences in the general subject of reaction–diffusion equations. A parallel development took place in combustion phenomena, where reaction–diffusion processes occur with equal importance. The interest in reaction–diffusion equations that these two areas, the biological sciences and combustion theory, generated in the mathematical community has been profound and a significant step has been taken in the understanding of nonlinear processes and nonlinear PDEs.

This chapter is an introduction to reaction–diffusion systems. In Chapter 5 we introduced Fisher’s equation, which is a model for a diffusing population with logistics growth. Now we wish to study some simple examples where systems of such equations arise. In Section 6.1 we formulate models for competing populations, interacting chemical species, chemotaxis (the interaction of a diffusing population and a chemical attractant), and the deposition of a pollutant in groundwater flows. In Section 6.2 we examine a basic problem for reaction–diffusion equations: the existence of traveling wave solutions. In Sections 6.3 and 6.4 we introduce some of the theoretical questions associated with reaction–diffusion equations, and nonlinear parabolic systems in general. In particular,

we ask about existence and uniqueness of solutions, and we formulate maximum principles and comparison theorems. In Section 6.5 we introduce the energy method, a technique for obtaining estimates and bounds on various quantities (e.g., the energy) associated with the solutions of reaction–diffusion or other evolution equations; we also study the blowup problem and the method of invariant sets. Finally, in the last section we examine Turing systems and pattern formation in biology.

6.1 Reaction–Diffusion Models

In Chapter 5 we introduced Fisher’s equation

$$u_t - Du_{xx} = ru \left(1 - \frac{u}{K}\right),$$

which is a prototype for a single reaction–diffusion equation. In this section we discuss the origin of *systems* of reaction–diffusion equations and we take up several mathematical models that illustrate applications of interest. Let $u_i(x, t)$, $i = 1, \dots, n$, be smooth functions representing n unknown concentrations or densities in a given one-dimensional system (say, in a tube of unit cross-sectional area and of infinite or finite extent). These quantities may be interacting population densities, chemical concentrations, energy densities, or whatever, distributed over the length of the tube. Let $\phi_i(x, t)$, $i = 1, \dots, n$, denote the flux of the quantity u_i , with the usual convention that ϕ_i is positive if the net flow of u_i is to the right, and let f_i be the rate of production of u_i . The dimensions of the quantities u_i , ϕ_i , and f_i are

$$[u_i] = \frac{\text{amount}}{\text{volume}}, \quad [\phi_i] = \frac{\text{amount}}{\text{time} \cdot \text{area}}, \quad [f_i] = \frac{\text{amount}}{\text{time} \cdot \text{volume}}.$$

Generally, the production rate f_i of u_i may depend on both space and time as well as all the densities:

$$f_i = f_i(x, t, u_1, \dots, u_n).$$

The basic conservation law developed in Chapter 1 for a single density now applies to each density u_i . That is, if $I = [a, b]$ is an interval representing an arbitrary section of the tube, then

$$\frac{d}{dt} \int_a^b u_i dx = -\phi_i \Big|_{x=a}^{x=b} + \int_a^b f_i dx, \quad i = 1, \dots, n. \quad (6.1.1)$$

These n equations express the fundamental principle that the rate of change of the total amount of the quantity u_i in I must be balanced by the net rate

that u_i flows into I plus the rate that u_i is produced in I . In the usual way, assuming smoothness, the system of integral balance laws can be reduced to a system of n partial differential equations

$$(u_i)_t + (\phi_i)_x = f_i, \quad i = 1, \dots, n, \quad (6.1.2)$$

where the densities u_i and the fluxes ϕ_i are unknown, and the production rates (source or reaction terms) f_i are assumed to be given. It is clear that (6.1.2) can be written in vector form as

$$\mathbf{u}_t + \boldsymbol{\phi}_x = \mathbf{f}(x, t, \mathbf{u}), \quad (6.1.3)$$

where

$$\mathbf{u} = (u_1, \dots, u_n)^T, \quad \boldsymbol{\phi} = (\phi_1, \dots, \phi_n)^T, \quad \mathbf{f} = (f_1, \dots, f_n)^T.$$

System (6.1.3) represents n equations with $2n$ unknowns (\mathbf{u} and $\boldsymbol{\phi}$). Additional equations come from constitutive relations that connect the flux to the density gradients. When appended to (6.1.3), the constitutive relations effectively eliminate the flux from the equations and we obtain a system of PDEs for the unknown densities alone.

The simplest constitutive assumption is to impose that the flux ϕ_i depend only on the gradient of the i th density u_i , not on the remaining densities. Symbolically this is expressed as

$$\phi_i = -D_i(u_i)_x, \quad i = 1, 2, \dots, n, \quad (6.1.4)$$

where the D_i are diffusion constants. If the diffusion constants are positive, then the flow is from high concentration regions to low concentration regions. But it may occur, for example, that organisms are attracted to their kind and movement is *up the concentration gradient*, that is, from low concentration to high concentration regions. In this case, a diffusion constant could be negative. Or, one density may represent an animal population, while another may represent a chemical concentration. In the case of chemotaxis, animals may be attracted to chemicals (e.g., pheromones) with a negative diffusion constant. In any case, substituting (6.1.4) into (6.1.2) gives a *system of reaction–diffusion equations*

$$(u_i)_t - D_i(u_i)_{xx} = f_i(x, t, u_1, \dots, u_n), \quad i = 1, 2, \dots, n, \quad (6.1.5)$$

for the n unknown densities u_1, \dots, u_n . In vector form this is

$$\mathbf{u}_t - D\mathbf{u}_{xx} = \mathbf{f}(x, t, \mathbf{u}), \quad (6.1.6)$$

where $D = \text{diag}(D_1, \dots, D_n)$ is an n by n diagonal matrix with the diffusion constants on the diagonal. In the system (6.1.6) the coupling is through the

reaction terms f_i ; the diffusion operators on the left side of the equations are decoupled because of the simple form of the constitutive relations (6.1.4).

A step up in difficulty is to assume that there is *cross-diffusion* between the densities u_i . This means that the flux ϕ_i of the density u_i depends on the gradients of all the densities, not just the gradient of u_i alone. Thus

$$\phi_i = - \sum D_{ij}(u_j)_x, \quad i = 1, 2, \dots, n, \quad (6.1.7)$$

or, in matrix form

$$\boldsymbol{\phi} = -D\mathbf{u}_x, \quad (6.1.8)$$

where $D = (D_{ij})$ is an $n \times n$ matrix of diffusion coefficients. D_{ij} is a measure of how u_i diffuses into u_j . Therefore, substitution of (6.1.8) into (6.1.3) gives a system the same form as (6.1.6), but with D a nondiagonal matrix. In the case of cross-diffusion, (6.1.6) represents a system of reaction–diffusion equations that is coupled through both the reaction terms and the differential operators. It is, of course, possible that the diffusion coefficients depend on the densities u_i . Examples of this type of phenomenon were given in Chapter 5 for one-dimensional heat flow. These models extend to several spatial variables as well.

We now present several situations where reaction–diffusion equations occur naturally. These examples come from a variety of physical and biological contexts, and they indicate the breadth of application of this important class of equations. In general, reaction–diffusion equations are of great importance in applied mathematics.

6.1.1 Predator–Prey Model

We noted that Fisher's equation

$$u_t - Du_{xx} = ru \left(1 - \frac{u}{K}\right)$$

is a simple model of a population u that diffuses while it grows according to the logistics growth law. Multiple populations lead to interacting populations. For example, suppose that one population is a predator whose population density is v and another is prey whose population density is u . Assume the prey grows at a rate proportional to its population (the Malthus model), and the predator, in the absence of prey, dies at a rate proportional to its population. The number of interactions between the predators and the prey can be simply modeled by mass action kinetics; this means it is proportional to the product uv ; this interaction increases the number of predators while it reduces the number of

prey. In summary

$$\text{Growth rate of prey} = f_1(u, v) = au - buv,$$

$$\text{Growth rate of predator} = f_2(u, v) = -cv + duv,$$

where a , b , and c are positive proportionality constants. If ϕ_1 and ϕ_2 denote the fluxes of the prey and the predators, respectively, and we assume Fick's law for each, that is

$$\phi_1 = -D_1 u_x, \quad \phi_2 = -D_2 v_x \quad (D_1 \text{ and } D_2 \text{ positive constants}),$$

then the balance laws are

$$u_t - D_1 u_{xx} = au - buv,$$

$$v_t - D_2 v_{xx} = -cv + duv.$$

The growth rates in these coupled reaction-diffusion equations come from laws originally proposed by A. Lotka and V. Volterra in 1926. The Lotka–Volterra model provides for an interaction model where one growth rate is increased by the interaction, while the other is decreased. This type of model, in general, is called a *predator-prey* model. Population models in ecology are a rich source of reaction-diffusion equations. The consideration of different interaction terms, population-dependent diffusion coefficients, and several species all lead to important models that are studied extensively in population dynamics.

6.1.2 Combustion

A combustion process involves mass, momentum, and energy transport in a chemically reacting fluid. We consider a simplified system where a substance in a tube undergoes an exothermic (heat releasing) chemical reaction $\mathbf{A} \rightarrow \mathbf{B}$, where \mathbf{A} is the reactant and \mathbf{B} is the product. As the reaction proceeds, temperature changes occur and both heat energy and matter diffuse through the medium. In this simple model, there is no advection. Let $a = a(x, t)$ be the concentration (in moles per volume) of the reactant \mathbf{A} , and let $T = T(x, t)$ be the absolute temperature of the medium, in degrees Kelvin. Mass balance takes the form

$$a_t + \phi_x = f(a, T), \tag{6.1.9}$$

where ϕ is the flux of \mathbf{A} and $f(a, T)$ is the consumption rate of \mathbf{A} , or the rate that \mathbf{A} is consumed by the reaction. Assume that the flux is given by Fick's law

$$\phi = -Da_x, \tag{6.1.10}$$

where D is the mass diffusion coefficient. Thus, the flux is caused by concentration gradients. From chemical kinetics theory it is known that the reaction rate has the form

$$f(a, T) = -sa^m e^{-E/RT}, \quad (6.1.11)$$

where $m > 0$ is the *order of the reaction*, E is the *activation energy*, R is the universal gas constant, and s is the *preexponential factor*. Equation (6.1.11) is called the *Arrhenius equation*, after its originator. Finally, (6.1.9)–(6.1.11) give the *mass balance law*

$$a_t - Da_{xx} = -sa^m e^{-E/RT}. \quad (6.1.12)$$

Now energy balance. If c is the specific heat per unit volume (cal/deg · vol), cT represents the energy density in calories per volume. We assume that the entire chemical mixture has the same specific heat. The energy flux is taken to $-KT_x$ (Fourier's law), where K is the thermal conductivity, and the energy (heat) production term is proportional to the reaction rate and is given by

$$Qsa^m e^{-E/RT},$$

where $Q > 0$ is the heat of reaction. The *energy balance* equation is therefore

$$(cT)_t - (KT_x)_x = Qsa^m e^{-E/RT}$$

or

$$T_t - kT_{xx} = c^{-1}Qsa^m e^{-E/RT}, \quad (6.1.13)$$

where $k = K/c$ is the diffusivity. Therefore, we obtained a reaction–diffusion model (6.1.12)–(6.1.13) for the temperature and concentration of the reactant.

Equations (6.1.12)–(6.1.13) are usually supplemented by initial and boundary conditions. If the tube has finite length ($0 \leq x \leq L$), at the ends of the tube we impose

$$a_x(0, t) = a_x(L, t) = 0, \quad t > 0, \quad (6.1.14)$$

which means that the flux of **A** is zero at the ends, or no chemical can escape the confines of the tube. If the ends of the tube are held at fixed temperatures, then

$$T(0, t) = T_1, \quad T(L, t) = T_2, \quad t > 0. \quad (6.1.15)$$

Initial conditions are

$$T(x, 0) = T_0(x), \quad a(x, 0) = a_0(x), \quad 0 < x < L. \quad (6.1.16)$$

Equations (6.1.12)–(6.1.16) define a well-posed, reaction–diffusion system called the *solid-fuel combustion* problem.

The solid-fuel problem can be non-dimensionalized by defining

$$\bar{t} = st, \quad \bar{x} = \frac{x}{\sqrt{D/s}}, \quad \bar{a} = \frac{a}{a_0}, \quad \bar{T} = \frac{T}{T_0},$$

where a_0 and T_0 are reference values for mass and temperature. Then, equations (6.1.12)–(6.1.13) become, on taking $m = 1$ and dropping the overbars,

$$a_t = a_{xx} - ae^{\theta/T}, \quad T_t = dT_{xx} + \beta ae^{-\theta/T},$$

with

$$\theta = \frac{E}{RT_0}, \quad \beta = \frac{Qa_0}{CT_0}.$$

Example. (Semenov Problem) If both heat and mass diffusion are slow compared to reaction kinetics, then diffusion can be neglected, then a and T are functions of time only and we obtain

$$a' = -ae^{\theta/T}, \quad T' = \beta ae^{\theta/T}, \quad (6.1.17)$$

which is the *Semenov problem*. The initial conditions are

$$a(0) = 1, \quad T(0) = 1,$$

which means that the system is at temperature T_0 at time $t = 0$, and a_0 is the mass of the reactant. The Semenov problem can actually be solved in analytic form using the exponential integral

$$\text{Ei}(z) = - \int_z^\infty \frac{1}{u} e^{-u} du.$$

See Logan (1994, p 318) for the formula. But here we take a qualitative approach. Multiplying the first equation in (6.1.17) by β and then adding gives

$$\beta a + T = \beta + 1,$$

where we have used the initial conditions to evaluate the constant of integration. We can now eliminate a and obtain the model equation

$$T' = (\beta + 1 - T)e^{-\theta/T}, \quad T(0) = 1.$$

This differential equation is an first-order autonomous equation, and it is easy to see that there is a single positive, equilibrium at the value $T^* = \beta + 1$. (Draw the phase line.) Therefore, the solution curve begins at $T = 0$ at time $t = 0$ and then increases up to the equilibrium value T^* . The mass of the reactant decreases from 1 to 0 as the reactant is depleted.

In real reactions the activation energy is large, and so θ is a large parameter. In this case the temperature increases slowly as the reaction proceeds; then

there is a threshold value of time when there is a sharp increase in temperature up to the equilibrium value. This event, where the reaction rapidly goes to completion, is called *ignition*. Using singular perturbation methods this phenomenon can be analyzed analytically [see, e.g., Kapila 1983, and Logan 1994, pp 318–322], and the length of the induction period can be approximated.

6.1.3 Chemotaxis

Another reaction–diffusion process is the motion of organisms under the influence of diffusion and chemotaxis; the latter is motion induced by variations in the concentration of chemicals produced by the organisms themselves. A classic example can be described as follows. Slime mold amoebae feed on bacteria in the soil and are uniformly spatially distributed when the food supply is plentiful. However, as the food supply is depleted, the organisms begin to secrete a chemical [cyclic adenosine monophosphate (cAMP)] that acts as an attractant. It is observed that the amoebae move *up the concentration gradient* toward the high concentrations of the chemical and interesting wave patterns and aggregates form. These types of chemotactic motion occur in a variety of biological phenomena. Another example is the release of pheromones by some organisms; these chemicals act may as sexual attractants or as communication devices to signal predators.

We analyze a one-dimensional model and let $a = a(x, t)$ denote the population density of an organism (say, amoebae) and $c = c(x, t)$ denote the concentration of a chemical attractant secreted by the amoeba. We assume that there is motion of the organism due to random movement (diffusion) and due to chemotaxis. The chemical produced by the organism diffuses through the medium. The reaction–diffusion system is, by previous accounts

$$a_t + \phi_{1x} = 0, \quad (6.1.18)$$

$$c_t + \phi_{2x} = F(a, c), \quad (6.1.19)$$

where ϕ_1 and ϕ_2 are the fluxes of the organism and chemical, respectively, and F is the rate of production of the chemical. Now we make some specific constitutive assumptions. The chemical is assumed to move only by diffusion, and therefore

$$\phi_2 = -\delta c_x,$$

where δ is the diffusion constant. The chemical decays at a constant rate, and its production rate is proportional to the population of the organism, or

$$F(a, c) = fa - kc,$$

where f is the rate of secretion and k is the decay rate. Finally, the flux ϕ_1 of the organism consists of two parts: one contribution from diffusion and the other, from chemotaxis. Then

$$\phi_1 = \phi_{\text{diff}} + \phi_{\text{chem}}.$$

We assume Fickian diffusion, or

$$\phi_{\text{diff}} = -\mu a_x,$$

where $\mu > 0$ is the motility. For the chemotactic contribution we argue that ϕ_{chem} is proportional to the gradient of the chemical concentration; in addition, if the number of organisms is doubled for a given concentration gradient, the flux should be twice as great, and therefore ϕ_{chem} should also be proportional to a . Consequently

$$\phi_{\text{chem}} = \nu a c_x,$$

where $\nu > 0$ is a proportionality constant that measures the strength of the chemotaxis. There is no minus sign in the last constitutive equation because the organism is assumed to move toward the attractant, or up the gradient.

Substitution of all these quantities into (6.1.18) and (6.1.19) gives a coupled system of reaction–diffusion equations

$$a_t = \mu a_{xx} - \nu (ac_x)_x = 0, \quad c_t = \delta c_{xx} + fa - kc, \quad (6.1.20)$$

which governs the concentration c of the chemical and population density a of the organism. This system is strongly coupled in that the differential operator in (6.1.20) involves derivatives of both c and a . Further, the nonlinear chemotactic flux assumption leads to a more complicated nonlinear diffusion term than in the preceding examples.

It is clear that different constitutive relations for F , ϕ_{chem} , and ϕ_{diff} can be considered, leading to yet more complicated models. For example, δ and μ may depend on the concentrations themselves, or the source term F for the secreted chemical may have a nonlinear functional dependence on c and a . Some of these models have been examined in the literature, and Murray (2002) can be consulted for references and applications.

Examples in this section show broad applicability of reaction–diffusion equations. We have done no analysis of these systems, but we infer a rich underlying theory that applies to these systems. Do solutions exist, and are they unique? What kinds of initial boundary value problems are well posed? Are there wavefront solutions? What is the long-time behavior of solutions, and can spatial patterns form? Are wavefront solutions stable to small perturbations? In the sequel we address some of these questions. In particular, in Section 7.6 we examine the stability of the uniform state for (6.1.20).

EXERCISES

1. In the Semenov problem numerically solve the differential equation

$$T' = (\beta + 1 - T)e^{-\theta/T}, \quad T(0) = 1,$$

when $\theta = 10$ and $\beta = 6$ and draw the solution curve in the interval $300 \leq t \leq 600$. Observe the ignition effect. Explain why this occurs by plotting the exponential factor on the right side of the equation when θ is large.

2. Consider the reaction–diffusion system

$$u_t - u_{xx} = u(1 - u - rv), \quad v_t - v_{xx} = -buu, \quad (r, b > 0).$$

In the special case $v = (1 - b)(1 - u)/r$ (b not equal to 1), show that the system reduces to Fisher's equation.

3. Let $u = u(x, t)$ denote the population of an organism and $n = n(x, t)$ denote the concentration of a nutrient to which the organisms are attracted. Give an interpretation of the model equations

$$u_t = \left(Du_x - \frac{au}{n} n_x \right)_x, \quad n_t = -ku,$$

where D , a , and k are positive constants. In the special case that $a = 2D$, find solutions of the form $u = U(z)$, $n = N(z)$, where $z = x - ct$, $U(\pm\infty) = 0$ and $N(-\infty) = 0$, $N(+\infty) = 1$. Sketch the solutions and interpret them biologically.

4. Consider the chemical reaction mechanism



where the concentrations a , b , d , and e of the chemical species **A**, **B**, **D**, and **E**, respectively, are held constant. If rate constants are unity, write out a system of two ODEs that govern the time evolution of the concentrations x and y of **X** and **Y**. For $a = 2$ and $b = 6$, describe how the concentrations x and y evolve. If the diffusion constants for **X** and **Y** are identical, what is the reaction–diffusion system that governs the spacetime evolution of the species?

5. (*Ohmic Heating*) Electrical devices often generate heat when current flows through them. Therefore we consider equations of heat conduction coupled to the equations of electromagnetism, through ohmic heating. Let D be a nice domain in \mathbb{R}^2 , and let $T = T(x, y, t)$ and $\phi = \phi(x, y, t)$ denote the temperature and electrical potential, respectively. The steady-state model

is given by the conservation law for heat energy and the equation for conservation of charge¹

$$\nabla \cdot (K(T)\nabla T) + \sigma(T) |\nabla \phi|^2 = 0, \quad \nabla \cdot (\sigma(T)\nabla \phi) = 0,$$

where $(x, y) \in D$, K is the temperature-dependent thermal conductivity, and σ is the temperature-dependent electrical conductivity. We expect to give boundary conditions on T and ϕ on ∂D .

- (a) Write out the model explicitly using partial derivatives in both rectangular and polar coordinates.
- (b) Let D be a rectangle $0 < x < L$, $0 < y < b$. Assume the boundaries $y = 0$ and $y = b$ to be thermally and electrically insulated. On the side $x = 0$ take $\phi = V_0$ and $T = T_0$, while on $x = L$ take $\phi = V_1$ and $T = T_0$. Write out the boundary value problem explicitly, and, assuming $T = T(x)$ and $\phi = \phi(x)$, show ϕ and $T = T(\phi)$ satisfy

$$\frac{1}{2}(\phi - V_0)(V_1 - \phi) = \int_{T_0}^{T(\phi)} \frac{K(\tau)}{\sigma(\tau)} d\tau.$$

In the special case the resistivity (the inverse of conductivity) is linear in temperature and K is constant, find the temperature distribution in the region D .

- (c) In the last part, with general conductivities K and σ , show that the maximum temperature depends only on the potentials at the endpoints and on the two conductivities, but not on the size of the region.
- (d) Next assume the boundaries $y = 0$ and $y = b$ are electrically insulated and $T = 0$ on those boundaries. On $x = 0$ take $\phi = 0$, while on $x = L$ take $\phi = V$; assume that the sides are thermally insulated. Find a formula for the potential ϕ , and find an equation that determines T implicitly.

6.2 Traveling Wave Solutions

A fundamental question regarding nonlinear PDEs is the existence of traveling waves. We introduced this concept in Chapters 1 and 5 for a single PDE. In this section we study the existence of such solutions for systems of partial differential equations by considering two examples, the geographic spread of a disease,

¹ In this exercise we use ∇ for the grad operator.

which is a basic reactive–diffusive system, and the flow and sedimentation of pollutants in groundwater, which is a reaction–advection–diffusion system.

It is difficult to formulate encompassing principles that apply to general classes of equations, and therefore the usual strategy is to examine each system separately. The procedure, however, is easy to follow: Assume a solution of the PDEs of the form of a traveling wave, $\mathbf{u} = \mathbf{u}(x - ct)$, where c is the wave speed, and reduce the problem to a system of ordinary differential equations for the waveforms \mathbf{u} . The latter dynamical system is a problem in a two- or higher-dimensional phase space, and the question of existence of TWS is often reduced to the existence of an orbit of the dynamical system connecting two critical points (a heteroclinic orbit) or an orbit connecting a single critical point (a homoclinic orbit). The critical points represent the boundary conditions at infinity. Interpreted in a different way, the unknown wave speed c acts as an unknown eigenvalue (i.e., a number for which there is a solution to a given boundary value problem). To show rigorously that such orbits exist in phase space is often a difficult task and involves determining the local structure of the vector field near the critical points, as well as the global structure. Two examples illustrate this procedure.

6.2.1 Model for the Spread of a Disease

We consider the geographic spread (in one dimension) of a rabies epidemic in a population of foxes. Some of the same assumptions apply to the spread of other epidemics among other organisms. We assume that the fox population is divided into two classes, the number of susceptibles $S = S(x, t)$, and the number of infectives $I = I(x, t)$, which includes those in the incubation stage. We assume the infected class increases at a rate βSI , where $\beta > 0$ is the infection rate, and therefore the susceptible population decreases at the same rate. Further, the death rate of infectives is proportional to the number of infectives $-aI$, where $a > 0$ is the mortality rate. We assume that susceptibles become infectives with a very short incubation period. Finally, we postulate that infectives diffuse, possibly because of disorientation, but that susceptibles do not diffuse. The model equations are therefore

$$S_t = -\beta SI, \quad (6.2.1)$$

$$I_t = DI_{xx} + \beta SI - \mu I, \quad (6.2.2)$$

where D is the diffusion constant of the infectives. It is straightforward to normalize this reaction-diffusion system and obtain dimensionless equations

$$S_t = -SI, \quad (6.2.3)$$

$$I_t = I_{xx} + SI - bI, \quad (6.2.4)$$

where b is a dimensionless parameter given by $b = \mu/\beta S_0$, representing the ratio of the death rate to the initial rate of infection. In the sequel we assume that $b < 1$. The quantity S_0 is the initial number of susceptibles. [To nondimensionalize, we scaled both S and I by S_0 , time by $(\beta S_0)^{-1}$, and x by $D/\sqrt{\beta S_0}$.]

We seek wavefront solutions to the reaction-diffusion system (6.2.3)–(6.2.4). These are solutions of the form

$$S = S(z), \quad I = I(z), \quad z = x - ct. \quad (6.2.5)$$

We anticipate that ahead of the wave we should have the boundary conditions

$$S(+\infty) = 1, \quad I(+\infty) = 0, \quad (6.2.6)$$

or that the wavefront is moving into a state where there are no infectives. After the wave passes, we expect that there are no infectives (rabies is almost always fatal). At the present time it is not possible to determine the number of susceptibles, if any, after passage of the epidemic wave. Therefore, at $z = -\infty$ we assume boundary conditions

$$S'(-\infty) = 0, \quad I(-\infty) = 0, \quad (6.2.7)$$

where it is required that the derivative of S vanish. Substituting (6.2.5) into (6.2.3) and (6.2.4) gives the nonlinear system

$$cS' = SI, \quad -cI' - I'' = SI - bI. \quad (6.2.8)$$

Applying the first equation in (6.2.8), we see that the second becomes

$$-cI' - I'' = cS' - \frac{cbS'}{S},$$

which integrates to

$$-cI + I' = cS - cb \ln S + A, \quad (6.2.9)$$

where A is a constant of integration. Evaluating (6.2.9) at $z = +\infty$ gives $A = -c$, and therefore

$$S' = c^{-1}SI, \quad (6.2.10)$$

$$I' = -c(S + I) + bc \ln S + c. \quad (6.2.11)$$

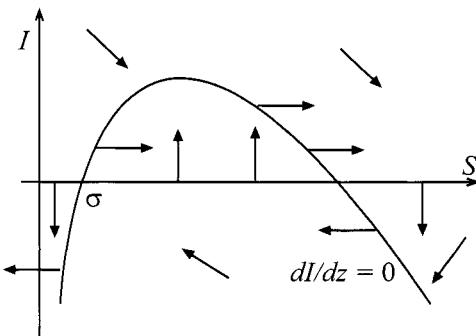


Figure 6.1 Nullclines and direction field associated with the system (6.2.10)–(6.2.11).

This system can be analyzed by phase plane techniques (see the Appendix to Chapter 5). The nullclines (where $S' = 0$ and where $I' = 0$) are easily determined. $S' = 0$ along $S = 0$ and $I = 0$, and $I' = 0$ along the locus

$$I = b \ln S - S + 1. \quad (6.2.12)$$

It is easily checked that this locus, when sketched in the SI plane, has a positive maximum at $S = b$ and crosses the S axis at $S = 1$ and $S = \sigma$, where σ satisfies the algebraic equation $b \ln \sigma - \sigma + 1 = 0$. Note that $\sigma < b < 1$. In the first quadrant we have $S' > 0$, while in the third quadrant $S' < 0$. Above the locus defined by (6.2.12) we have $I' < 0$, and below the locus $I' > 0$. The critical points are $(1, 0)$ and $(\sigma, 0)$, which are the intersections of the loci $S' = 0$ and $I' = 0$. A graph of the direction field is shown in Figure 6.1.

The type and stability of the two critical points can be determined by linearization. The Jacobian matrix is

$$J(S, I) = \begin{pmatrix} \frac{I}{c} & \frac{S}{c} \\ -c + \frac{bc}{S} & -c \end{pmatrix}.$$

At $S = \sigma$ and $I = 0$ the matrix $J(\sigma, 0)$ has real eigenvalues of opposite sign, so $(\sigma, 0)$ is a saddle point. At $S = 1$ and $I = 0$ the matrix $J(1, 0)$ has eigenvalues λ that satisfy the characteristic equation $\lambda^2 + c\lambda + (1 - b) = 0$. Therefore

$$\lambda = \frac{1}{2}\{-c \pm [c^2 - 4(1 - b)]^{1/2}\}.$$

If $c > 2(1 - b)^{1/2}$, the eigenvalues are both real and negative and $(1, 0)$ is a stable node. If $c < 2(1 - b)^{1/2}$, the eigenvalues are complex with negative real parts and $(1, 0)$ is a stable spiral. The phase diagram is shown in Figures 6.2 and 6.3. In both cases there is a unique separatrix connecting the critical point

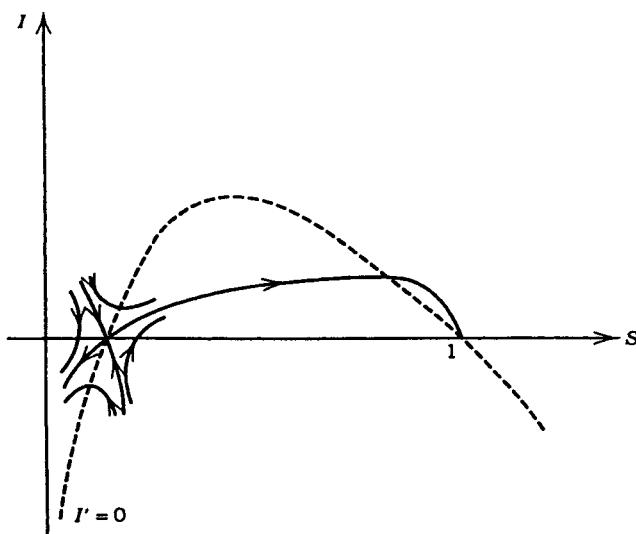


Figure 6.2 Phase portrait of (6.2.10)–(6.2.11) in the case $c > 2\sqrt{1 - b}$

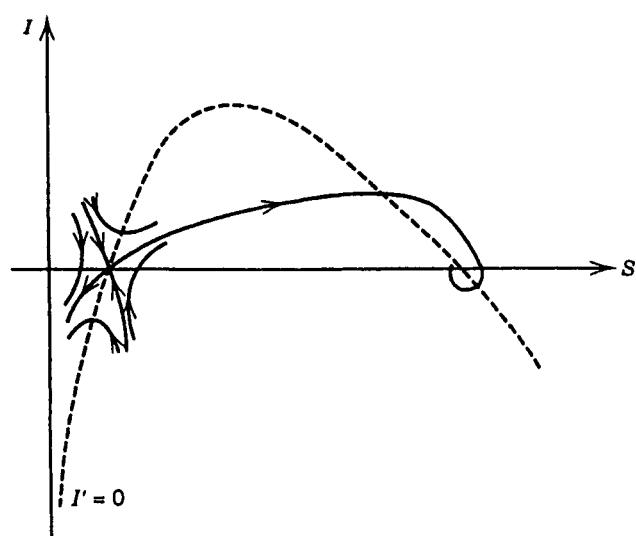


Figure 6.3 Phase portrait of (6.2.10)–(6.2.11) in the case $c < 2\sqrt{1 - b}$.

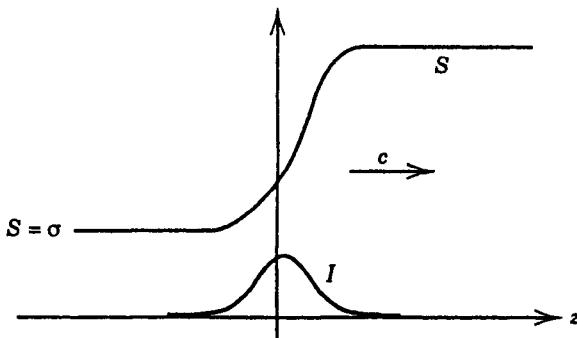


Figure 6.4 Traveling wave solution to the reaction–diffusion system (6.2.3)–(6.2.4).

$(\sigma, 0)$ at $z = -\infty$ to $(1, 0)$ at $z = +\infty$, and each corresponds to a traveling wave. On physical grounds we reject the case when $(1, 0)$ is a spiral because the number of infectives oscillates around $I = 0$ for large z , giving negative values. Consequently, if $c > 2(1 - b)^{1/2}$, there is traveling wave solution of the form shown in Figure 6.4 corresponding to the heteroclinic orbit (the separatrix) connecting the saddle point $(\sigma, 0)$ to the stable node $(1, 0)$. This wave travels at speed c .

Note that σ is the number of susceptibles that survive the epidemic, and it is an increasing function of the parameter b ; thus the smaller b , the fewer foxes survive the rabies epidemic. If $b = 0.5$, for example, σ is about 0.2, and so 20% of the foxes survive; the dimensionless wave speed is $c = 1.414$. The reader is referred to Murray (2002, 2003) for references to actual studies comparing the theoretical results to experimental observations.

Example. (Diffusion of Susceptibles) When susceptibles also diffuse, the dynamics is given by the system

$$\begin{aligned} S_t &= -\beta SI + DS_{xx}, \\ I_t &= \beta SI - \mu I + DI_{xx}, \end{aligned}$$

where D is the diffusion constant, assumed the same for both populations. As boundary conditions we assume that

$$\begin{aligned} I &\rightarrow 0 \text{ as } x \rightarrow \pm\infty, \\ S &\rightarrow N \text{ as } x \rightarrow +\infty, \quad \lim_{x \rightarrow -\infty} S \geq 0. \end{aligned}$$

Substituting $S = S(z)$, $I = I(z)$, $z = x - ct$, gives

$$-cS' = -\beta SI + DS'', \quad (6.2.13)$$

$$-cI' = \beta SI - \mu I + DI''. \quad (6.2.14)$$

Adding these equations gives

$$U'' + \frac{c}{D}U' = \frac{\mu}{D}I, \quad U \equiv S + I.$$

Multiplying by the integrating factor $\exp(cz/D)$ leads to

$$\frac{d}{dz} \left(e^{cz/D} U' \right) = \frac{\mu}{D} I e^{cz/D}.$$

Now integrate from $z = -\infty$ to z and use the boundary conditions to obtain

$$(S + I)' = \frac{\mu}{D} e^{-cz/D} \int_{-\infty}^z I(r) e^{cr/D} dr.$$

The right side is positive, and it follows that

$$S(z) + I(z) < S(\infty) + I(\infty) = N,$$

and therefore $S(z) < N$ for all z .

Even though we cannot solve the problem, we can still determine a necessary condition for a traveling wave to propagate. Note that

$$\int_{-\infty}^{\infty} I(z) dz > \frac{1}{N} \int_{-\infty}^{\infty} I(z) S(z) dz.$$

Also, if we integrate (6.2.14) over the real line, again using the boundary conditions, we obtain

$$\frac{\beta}{\mu} \int_{-\infty}^{\infty} S(z) I(z) dz = \int_{-\infty}^{\infty} I(z) dz.$$

Comparing this to the previous equation forces

$$\frac{\beta N}{\mu} > 1.$$

This condition is a necessary condition for a wave to propagate. Clearly, the condition is violated if the mortality rate μ is too large, or if the infection rate β is too small. Stated in a different way, there is a minimal population density $N = \mu/\beta$ necessary to propagate an epidemic wave. One can show (Noble 1974) that the epidemic wave propagates with speed

$$c \approx \sqrt{\beta ND}.$$

6.2.2 Contaminant Transport in Groundwater

In this section we formulate a nonlinear advection–reaction–diffusion model for the transport of solutes (e.g., contaminants) through a one-dimensional porous medium (e.g., soil). The model is based on mass balance and constitutive assumptions for the flux of the solute and the rate of adsorption of the solute to the fixed soil matrix. The assumptions leads to a coupled system of reaction–diffusion equations with a nonlinear advection term, and the question of existence of TWS is addressed.

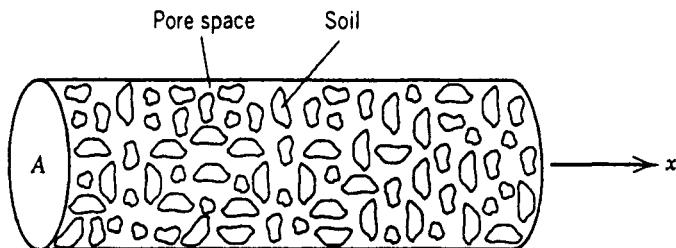


Figure 6.5 Porous medium of cross-sectional area A carrying a chemical dissolved in a liquid.

We consider a one-dimensional underground flow in the horizontal direction (denoted by x) occurring in a tube of cross-sectional area A (Figure 6.5). The tube is assumed to be a porous medium where the fluid (water) can occupy only a fraction ϕ of the total volume. The fraction ϕ is called the *porosity* of the medium, and we assume that it is constant. We let $C = C(x, t)$ denote the mass concentration of a chemical (the solute) dissolved in the water and $N = N(x, t)$ denote the mass concentration of the chemical that is adsorbed on the soil. By $q = q(x, t)$ we denote the flux of the solute. In the usual manner we can express mass balance in integral form as

$$\frac{d}{dt} \int_a^b \phi A C(x, t) dx = A q(a, t) - A q(b, t) - \int_a^b A N_t(x, t) dx, \quad (6.2.15)$$

where $[a, b]$ is an arbitrary section of the tube. Equation (6.2.15) states that the time rate of change of the total amount of the solute in the section $[a, b]$ is balanced by the net rate at which the solute flows into the section minus the rate of solute adsorption by the soil in that section. Assuming smoothness of the functions C , N , and q , we may express (6.2.15) as a partial differential equation

$$\phi C_t + q_x + N_t = 0. \quad (6.2.16)$$

To complete the formulation we require constitutive relations. We assume that the flux q has two contributions; the first is Fickian diffusion, and the second is a transport term caused by the bulk movement of the fluid through the medium. Thus

$$q = -DC_x + \frac{\phi BC^2}{2}, \quad (6.2.17)$$

where D is the diffusion constant and B is a bulk movement constant. An additional constitutive relation is needed to relate C and N . Such relations are usually determined empirically; the simplest is a linear equilibrium relationship $N = \alpha C$. Another possible assumption is a nonequilibrium expression $N_t = a(N_0 - N)$, where a is a positive constant and N_0 is the maximum concentration where the soil becomes saturated. This equation determines N as a function of time and it can be substituted, along with (6.2.17), into (6.2.16) to obtain a single equation for C . The growth rate a may also depend in some manner on the solute concentration C . Some of these assumptions are explored in the exercises. For the present discussion we assume a nonequilibrium constitutive equation for the rate of adsorption of the form

$$N_t = KR(N, C), \quad (6.2.18)$$

where K is a rate constant and where R is a rate function having the sigmoid form

$$R(N, C) = \frac{N_0 C^2}{C_0^2 + C^2} - N, \quad (6.2.19)$$

where N_0 is a maximum, constant concentration where the soil is saturated

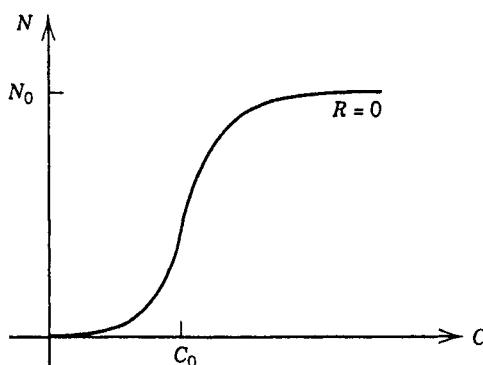


Figure 6.6 Graph of the locus $R(N, C) = 0$ defined by (6.2.19).

with the chemical, and C_0 is a threshold value of the solute concentration where the rate function switches on. Equations (6.2.18) and (6.2.19) imply that the

rate of adsorption increases with the concentration C up to some maximum rate, while it decreases with the concentration N ; in other words, the rate decreases as the soil becomes more adsorbed with the chemical pollutant. A graph of the equilibrium locus $R(N, C) = 0$ is shown in Figure 6.6. Above the equilibrium curve the adsorption rate is negative and the adsorbed chemical is returned to the water; below the equilibrium curve the adsorption rate is positive and the solute is deposited on the soil. Consequently, in this model there is a tendency toward an equilibrium state. This model is a generalization of some of the other models found in the literature; the reader can consult, for example, Guenther & Lee (1988) for an elementary discussion of linear models.

With assumptions (6.2.17) and (6.2.19), equations (6.2.16) and (6.2.18) become

$$\phi C_t + \phi BCC_x + N_t = DC_{xx}, \quad (6.2.20)$$

$$N_t + K \left(N - \frac{N_0 C^2}{C_0^2 + C^2} \right) = 0. \quad (6.2.21)$$

These equations can be nondimensionalized by introducing the dimensionless parameters

$$c = \frac{C}{C_0}, \quad n = \frac{N}{N_0}, \quad \tau = \frac{C_0^2 B^2 \phi}{D} t, \quad \xi = \frac{C_0 B \phi}{D} x.$$

Then we have

$$c_\tau + cc_\xi + \beta n_\tau = c\xi\xi, \quad (6.2.22)$$

$$n_\tau + k \left(n - \frac{c^2}{1 + c^2} \right) = 0, \quad (6.2.23)$$

where β and k are dimensionless constants defined by

$$\beta = \frac{N_0}{\phi C_0}, \quad k = \frac{KD}{C_0^2 B \phi}.$$

We seek traveling wave solutions of (6.2.22)–(6.2.23) of the form (we use the symbol v for the scaled wave speed)

$$c = c(\xi - v\tau), \quad n = n(\xi - v\tau). \quad (6.2.24)$$

Ahead of the wave we assume that $c = n = 0$, that is, that the solute and adsorbed concentrations are zero; after the wave passes, the conditions behind it will be determined depending on the velocity v . Thus the boundary conditions are

$$c(+\infty) = n(+\infty) = 0. \quad (6.2.25)$$

Substituting (6.2.24) into (6.2.22) and (6.2.23) gives

$$-vc' + cc' - \beta v n' = c'', \quad (6.2.26)$$

$$-vn' + k \left(n - \frac{c^2}{1+c^2} \right) = 0, \quad (6.2.27)$$

where a prime denotes d/dz , where $z = \xi - v\tau$. Noting that $cc' = (c^2/2)'$, equation (6.2.26) can be integrated with respect to z , and the constant of integration can be determined by the boundary conditions (6.2.25). After rearrangement, we obtain

$$c' = -vc - \beta v n + \frac{c^2}{2}, \quad (6.2.28)$$

$$n' = kv^{-1} \left(n - \frac{c^2}{1+c^2} \right). \quad (6.2.29)$$

We analyze this system in the phase plane. The nullclines are given by the parabola

$$n = 2(\beta v)^{-1} c(c - 2v), \quad (6.2.30)$$

where the vector field is vertical ($c' = 0$), and the equilibrium curve

$$n = \frac{c^2}{1+c^2}, \quad (6.2.31)$$

where the vector field is horizontal ($n' = 0$). These curves are shown in Figure 6.7. The origin $(0,0)$, representing the state at plus infinity, is a critical point; there is at least one additional critical point in the first quadrant. The c -coordinates of the critical points are the positive real roots of the cubic equation

$$c^3 - 2vc^2 + (1 - 2\beta v)c - 2v = 0, \quad (6.2.32)$$

which comes from equating (6.2.30) and (6.2.31). Figure 6.7 shows the case of a single positive root of (6.2.32). We argue that for any $v > 0$ there is an orbit connecting (c_-, n_-) to $(0,0)$, where c_- is the smallest positive root of (6.2.32) and $n_- = c_-^2/(1+c_-^2)$. This heteroclinic orbit represents a TWS to the system (6.2.22)–(6.2.23).

Figure 6.7 shows a generic case; the equilibrium curve is fixed, but the parabola depends on the parameters β and v . In any case, the parabola has a negative minimum at $c = v$, and it lies strictly below the equilibrium curve for $0 < c < c_-$. The linearized Jacobian matrix associated with the nonlinear system (6.2.28)–(6.2.29) is

$$J(c, n) = \begin{pmatrix} c - v & -\beta v \\ -\frac{2kcv^{-1}}{1+c^2} & \frac{k}{v} \end{pmatrix}.$$

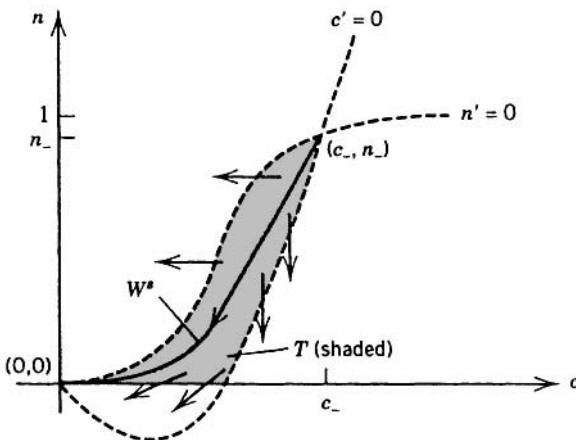


Figure 6.7 Nullclines and direction field associated with the system (6.2.28)–(6.2.29). W^s is the stable manifold entering the critical point $(0, 0)$.

The eigenvalues of $J(0, 0)$ are $-v$ and k/v , and therefore $(0, 0)$ is a saddle point. The eigenvectors are $(1, 0)$ and $(1, -(1 + k/v^2)/\beta)$. Now define the region T by

$$T = \left\{ (c, n) : n > 0, c < c_-, n > 2(\beta v)^{-1}c(c - 2v), \text{ and } n < \frac{c^2}{1 + c^2} \right\},$$

which is the region in the first quadrant below the equilibrium curve, above the parabola, and to the left of the first positive intersection point of the two curves. This region is shaded in Figure 6.7. Let W^s be the one-dimensional stable manifold (the separatrix) entering $(0, 0)$. This manifold must enter $(0, 0)$ from the region T . Now follow this manifold backward as $z \rightarrow -\infty$. Because the vector field along the boundaries of T , taken backward in time, point inward, we argue that W^s cannot leave the region T , and consequently it must enter the critical point (c_-, n_-) . The argument here can be stated differently—if time is reversed, the boundary of T consists of ingress points (the vector field is inward), so by the Wazewski retract theorem [see, e.g., Hartman 1964], the curve W^s must enter (c_-, n_-) . We can also note that on the contrary, if W^s did touch the boundary of T , uniqueness would be violated. Therefore, there exists an orbit (the separatrix) connecting (c_-, n_-) to $(0, 0)$, thus producing a traveling wave solution to the advection-reaction-diffusion system for any wave speed $v > 0$. Figure 6.8 shows the qualitative behavior of these concentration waves (further details may be found in Cohn & Logan 1995).

In the two preceding examples the problem of determining traveling waves reduced to a two-dimensional phase plane analysis. This is more than can usually be expected in a problem. More often than not, the dynamical system has

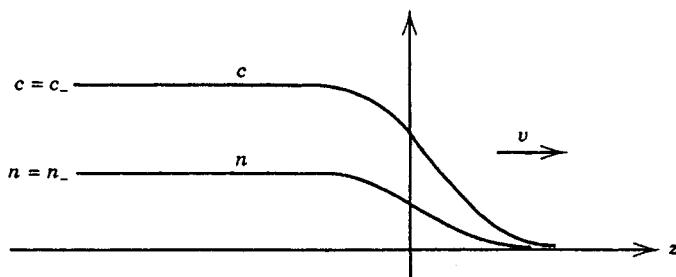


Figure 6.8 Traveling concentration waves.

dimension three (or more), and the analysis is considerably more difficult.

EXERCISES

1. In the rabies epidemic model, discuss the existence of traveling wave solutions in the case that the parameter b , the ratio of the death rate to the infection rate, is greater than unity.
2. Consider the mass balance equation (6.2.16) for the groundwater problem.
 - (a) Derive an equation for the concentration C when the flux is given by $q = -DC_x + \phi VC$, and $N = \alpha C$, where V is the velocity of the bulk movement of the water through the soil and α is a positive constant. Do traveling wave solutions exist in this case?
 - (b) Along with (6.2.16) and (6.2.17), assume that the rate of adsorption is given by $N_t = \alpha C - \gamma N$, where α and γ are positive constants. Show that traveling waves for the resulting system exist, and describe them.
3. Consider the reactive-diffusive system

$$u_t - Du_{xx} = u(1 - u - v), \quad v_t - v_{xx} = av(u - b),$$
 where a , b , and D are positive constants.
 - (a) Interpret this scaled model in a predator-prey setting.
 - (b) In the special case $D = 0$, investigate the existence of traveling waves (Dunbar 1983).
4. A simplified, scaled model of a detonation process is given by the system

$$u_t + uu_x + \lambda_x = Du_{xx}, \quad \lambda_t = r(u, \lambda),$$

where u is a scaled temperature-like variable and λ is the mass fraction of the product species \mathbf{P} in a reversible chemical reaction $\mathbf{R} \leftrightarrow \mathbf{P}$ with reaction rate r given by

$$r(u, \lambda) = 1 - \lambda - \lambda e^{-1/u},$$

and D is a positive constant. Show that there exist positive traveling wave solutions to this nonlinear advection–reaction–diffusion system if the wave speed c exceeds the value of u at $+\infty$ (Logan & Dunbar 1992).

5. Show that traveling waves exist for wave speeds $c > 2$ for the reaction–advection system

$$u_t + uu_x + v_x = (2 - u)(u - 1), \quad v_t = 1 - v - ve^{-1/u}.$$

(Logan & Shores 1993).

6. Determine all nonnegative traveling waves for the system

$$u_t + (u + v)_x = u_{xx}, \quad v_t = \frac{1 + u^2}{(1 + 2u^2)} - v.$$

[Note: A large number of examples and exercises on traveling waves in biological and chemical systems can be found in Murray (2002, 2003).]

7. A model for the burning of a solid waste material is given by the system

$$u_t = u_{xx} + kar(u) \quad a_t = -ar(u),$$

where $u = u(x, t)$ is the temperature; $a = a(x, t)$ is the concentration of the immobile, combustible, unburned waste material; k is a positive constant; and $r(u)$ is the burning rate. Investigate the existence of traveling wave solutions of speed c that satisfy the boundary conditions $(u, a) \rightarrow (0, a_+)$ at plus infinity and $(u, a) \rightarrow (u_-, 0)$ at minus infinity, where a_+ and u_- are some constants. Specifically:

- Show that if such solutions exist, then necessarily $u_- = ka_+$.
- Assume that $r(u) = r_0 u$, $r_0 > 0$, and show that traveling waves exist for wave speeds c satisfying $c^2 > 4r_0 u_-$. Interpret this result physically.
- Examine the problem assuming that the burning rate is $r(u) = 0$ if $u \leq u_b$, and $r(u) = r_0(u - u_b)$ if $u > u_b$, where u_b is an ignition temperature where the combustion process switches on (Grindrod 1996, pp 64–65).

8. Consider the system of reaction-advection-diffusion equations

$$\begin{aligned} u_t &= Du_{xx} - \gamma u_x - aF(u, v), \\ v_t &= -bF(u, v), \end{aligned}$$

where a , b , D , and γ are positive constants. Find a system of ODEs for traveling waveforms $u = U(z)$, $v = V(z)$, $z = x - ct$. If boundary conditions are given by

$$U(-\infty) = u_l, \quad U(+\infty) = 0, \quad V(-\infty) = 0, \quad V(+\infty) = v_r,$$

what conditions must the reaction term F satisfy for wavefront solutions to exist? Sketch possible wave profiles. Assuming that wavefronts exist, show that the speed c of the wave is less than the advection speed γ .

9. The following system of PDEs arises in the study of bioremediation of aquifer systems where immobile, indigenous microbes attached to the soil are stimulated to consume a contaminant and produce nontoxic products (Logan 2001, p. 107)

$$RS_t = -vS_x - F, \quad A_t = -vA_x - rF, \quad M_t = yF - b(M - M_0),$$

where $M(x, t)$ is the density of the microbes, $S(x, t)$ is the density of the contaminant (e.g., a hydrocarbon), $A(x, t)$ is the density of a nutrient stimulant (e.g., oxygen), and $F = qSAM/[(K_s + S)(K_a + A)]$ is the biodegradation rate. $R > 1$ is the retardation constant, v is the average velocity of the subsurface flow, b is the decay rate of the bacteria, y is the yield, M_0 is a reference microbe density, and r is the mass of the nutrient used per mass contaminant degraded; q , K_s , and K_a are rate constants. Write a short paragraph discussing the origin of the various terms in the model equations. Find the speed c of an assumed wavefront that satisfies the boundary conditions $S = 0$, $A = A_0$ at $-\infty$, $S = S_r$, $A = 0$ at $+\infty$, and $M = M_0$ at $\pm\infty$. Without solving the traveling wave differential equations, sketch anticipated profiles of S , A , and M as a function of the variable $z = x - ct$.

10. Investigate the existence of TWS for the system

$$p_t = k_1 sp^2 + k_2 q, \quad p_t = -vq_x - k_2 q + k_1 sp^2,$$

with $q \rightarrow 0$ at $+\infty$ and $p \rightarrow 0$ at $-\infty$.

11. A chemical wave moves through a tube $x \geq 0$, where a free protein in solution of concentration u becomes irreversibly bound to the immobile matrix structure of the medium. The concentration of the bound protein is

v , and the total concentration of binding sites is B . The dynamics is given by

$$u_t = Du_{xx} - \alpha u_x - f(u, v), \quad v_t = f(u, v),$$

where

$$f(u, v) = k_1(B - v)u - k_2v$$

is the rate of chemical binding of the protein.

- (a) If $u = u_0$ and $v = 0$ at $x = 0$ for all time t , does there exist a nonuniform equilibrium solution?
- (b) Does there exist a uniform equilibrium?
- (c) In \mathbb{R} show that there is a TWS with $u, v \rightarrow 0$ as $x \rightarrow +\infty$, provided

$$\frac{(\alpha - c)k_2}{c} < k_1B, \quad c < \alpha,$$

where c is the wave speed. What are the states at $-\infty$?

6.3 Existence of Solutions

We now consider the question of existence of a solution to a nonlinear initial value problem

$$u_t - Du_{xx} = f(u), \quad x \in \mathbb{R}, \quad t > 0, \quad (6.3.1)$$

$$u(x, 0) = u_0(x), \quad x \in \mathbb{R}. \quad (6.3.2)$$

For the present we consider the scalar case, and conditions on u_0 and the nonlinear reaction term $f(u)$ will be imposed later.

It is easy to write some simple examples of problems for which a solution does not exist for all $t > 0$. The initial value problem

$$u_t - Du_{xx} = u^2, \quad x \in \mathbb{R}, \quad t > 0,$$

$$u(x, 0) = u_0, \quad x \in \mathbb{R},$$

where u_0 is a positive constant, has a spatially independent solution $u(x, t) = u_0/(1 - u_0t)$, which blows up in finite time. A less trivial example is given in Section 6.5. Further, if a solution exists, it need not be unique, as the following example shows. Consider the initial-boundary value problem

$$u_t - u_{xx} = 2u^{1/2}, \quad x \in \mathbb{R}, \quad t > 0$$

$$u(x, 0) = 0, \quad x \in \mathbb{R}.$$

It is easy to check that both $u(x, t) = 0$ and $u(x, t) = t^2$ are solutions. Therefore, to guarantee existence and uniqueness it is evidently necessary to impose conditions on the initial data and the type of nonlinearity.

6.3.1 Fixed-Point Iteration

Of course, the best strategy for showing that a solution to a given problem exists is actually to exhibit a formula for the solution. For certain linear problems, we can actually proceed in this manner. For example, using Fourier transforms one can derive the solution to the linear, nonhomogeneous diffusion problem

$$u_t - Du_{xx} = g(x, t), \quad x \in \mathbb{R}, \quad t > 0, \quad (6.3.3)$$

$$u(x, 0) = u_0(x), \quad x \in \mathbb{R}. \quad (6.3.4)$$

For reference (we need this solution in the sequel), it is given by

$$u(x, t) = \int_{\mathbb{R}} K(x - y, t) u_0(y) dy + \int_0^t \int_{\mathbb{R}} K(x - y, t - s) g(y, s) dy ds, \quad (6.3.5)$$

where $K(x, t)$ is the diffusion kernel

$$K(x, t) = \frac{1}{\sqrt{4\pi Dt}} e^{-x^2/4Dt}, \quad (6.3.6)$$

and where u_0 and g are continuous bounded functions. For nonlinear problems, however, it is usually impossible to find such formulas, and alternative methods must be found to prove existence. Such existence questions are common in applied analysis, and in this section we formulate a general method that applies to many nonlinear problems. This method, called fixed-point iteration, is an impressive illustration of the unifying power of abstraction in analysis. The basic idea is to produce a sequence, through iteration of a certain map, that converges to the solution of the problem, thus showing existence.

We indicate, for the purpose of motivation, the methodology in two different settings, nonlinear algebraic equations and ordinary differential equations, before addressing the existence question for nonlinear PDEs.

Example. (Fixed-Point Iteration) Consider a nonlinear algebraic equation in the form

$$x = \phi(x), \quad (6.3.7)$$

where ϕ is a given continuous function defined on \mathbb{R} . By selecting an initial approximate x_0 and then iterating the map ϕ repeatedly via

$$x_{n+1} = \phi(x_n), \quad n = 0, 1, 2, \dots, \quad (6.3.8)$$

we produce a sequence of numbers $x_0, x_1, x_2, x_3, \dots$, which, under certain conditions, converges to a root of (6.3.7). For, suppose that x_n converges to some number x' (i.e., $\lim x_n = x'$). Then, taking the limit of both sides of (6.3.8) as

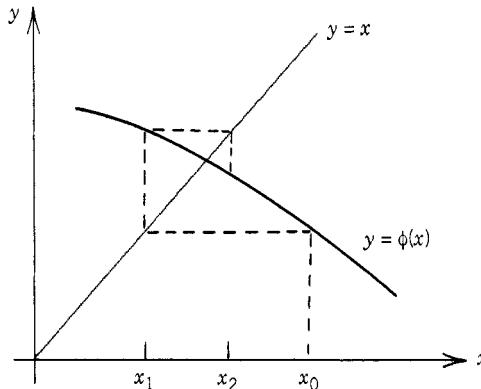


Figure 6.9 A cobweb diagram showing fixed-point iteration in the convergent case.

$n \rightarrow \infty$ and using the continuity of ϕ gives $x' = \lim \phi(x_n) = \phi(\lim x_n) = \phi(x')$, and therefore x' is a root of (6.3.7). Thus existence of a root can be proved from a limiting process. Because $\phi(x') = x'$, the function ϕ maps x' to itself, and therefore x' is called a *fixed point* of the mapping ϕ . The process (6.3.8) is often called *fixed-point iteration*. The situation is shown geometrically in Figure 6.9; a root x' is the x coordinate of the intersection of the graphs of $y = x$ and $y = \phi(x)$. The sequence x_n can be determined geometrically by drawing a cobweb diagram, depending on the shape of the graph of ϕ . If the graph is ϕ is too steep at x' , as shown in Figure 6.10, the sequence defined by the iterative process (6.3.8) will not converge to the root, or fixed point, x' . When will the fixed-point iteration converge? To fix the idea, let ϕ be a continuous function on \mathbb{R} and satisfy a *Lipschitz condition* of the form

$$|\phi(x) - \phi(z)| \leq k|x - z| \quad \text{for all } x, z \in \mathbb{R}, \quad (6.3.9)$$

where k is a constant with $0 < k < 1$. Notice that the Lipschitz condition puts a bound on all secant lines or chords connecting two points on the graph of ϕ , thereby restricting the steepness of the graph. Under condition (6.3.9) it is not difficult to observe that for any choice of x_0 , the iterative process (6.3.8) will converge to the unique root of (6.3.7). To prove this fact, we show that x_n is a Cauchy sequence of real numbers, and hence convergent. First we calculate the distance between consecutive iterates as

$$|x_{n+1} - x_n| = |\phi(x_n) - \phi(x_{n-1})| \leq k|x_n - x_{n-1}|.$$

Then, if $|x_1 - x_0| = a$, we have

$$|x_2 - x_1| \leq ka, \quad |x_3 - x_2| \leq k^2a, \dots, |x_{n+1} - x_n| \leq k^n a,$$

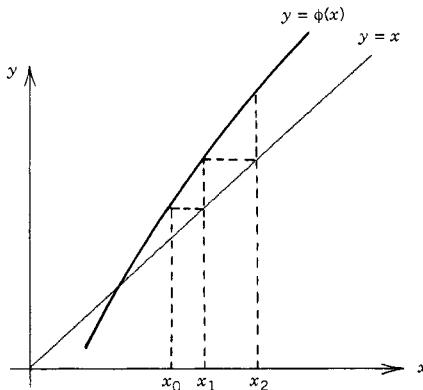


Figure 6.10 Cobweb diagram for fixed-point iteration in the divergent case.

whereupon, for any positive integer p , we obtain

$$\begin{aligned}
 |x_{n+p} - x_n| &\leq |x_{n+p} - x_{n+p-1}| + |x_{n+p-1} - x_{n+p-2}| + \cdots + |x_{n+1} - x_n| \\
 &\leq ak^{n+p-1} + ak^{n+p-2} + \cdots + ak^n \\
 &\leq ak^n[k^{p-1} + k^{p-2} + \cdots + 1] \\
 &< \frac{ak^n}{1-k}.
 \end{aligned}$$

Here, to get the first line we added and subtracted all the terms between x_{n+p} and x_n and then used the triangle inequality; to get the final inequality we applied the fact that the sum of the geometric series is $\sum k^n = 1/(1-k)$. Thus, because $k < 1$, the difference $|x_{n+p} - x_n|$ can be made arbitrarily small provided that n is chosen large enough, for all $p > 0$. In other words, for any $\varepsilon > 0$ there is an integer N such that $|x_{n+p} - x_n| < \varepsilon$ for $n > N$ and any $p > 0$. So, by definition, the sequence x_n is a Cauchy sequence, and therefore it converges to some x' . Our earlier argument showed that $x' = \lim x_n$ is a root of (6.3.7). It is straightforward to show that the root x' is unique. By way of contradiction, let x' and x'' be two distinct roots of (6.3.7); from (6.3.9) we have $|x' - x''| = |f(x') - f(x'')| \leq k|x' - x''|$, for $k < 1$, which is a contradiction; consequently, only one root can exist.

To summarize, we have shown that if ϕ is a continuous function on \mathbb{R} satisfying the Lipschitz condition (6.3.9), then (6.3.7) has a unique real root. Therefore, we proved the existence of a solution without exhibiting it explicitly; the argument is clearly based on the fact that the set of real numbers \mathbb{R} is complete (Cauchy sequences converge) and the assumption that the iteration function ϕ satisfies a Lipschitz condition. The latter condition, equation (6.3.9), can be interpreted alternatively by noting that the distance between $\phi(x)$ and $\phi(z)$

is strictly less than the distance between x and z ; that is, ϕ is a *contraction mapping*. The ideas of fixed-point iteration, contraction mappings, and completeness can be generalized to a broad setting where they apply to ordinary and partial differential equations, integral and functional equations, and other equations of interest in applied analysis.

Example. (Picard Iteration) In ODEs we are interested in the existence of a solution to the initial value problem

$$y' = f(t, y), \quad y(t_0) = y_0, \quad (6.3.10)$$

where t_0 and y_0 are given constants; f is assumed to be a continuous function. We can easily rewrite the initial value problem (16.3.10) as an integral equation by integrating from t_0 to t to obtain

$$y(t) = y_0 + \int_{t_0}^t f(s, y(s)) ds. \quad (6.3.11)$$

The right side of (6.3.11) can be regarded as a mapping Φ on the set of continuous functions; that is, with each continuous function y there is associated another continuous function $\Phi(y)$ defined at each t by

$$\Phi(y)(t) \equiv y_0 + \int_{t_0}^t f(s, y(s)) ds.$$

Therefore, the integral equation (6.3.11) may be written in the form

$$y(t) = \Phi(y)(t),$$

which formally has the same structure as the algebraic equation (6.3.7) in the preceding example. Therefore, in the same manner we may fashion a sequence of functions $y_0(t), y_1(t), y_2(t), \dots$ by the iteration formula

$$\begin{aligned} y_{n+1}(t) &= \Phi(y_n)(t), \\ &= y_0 + \int_{t_0}^t f(s, y_n(s)) ds, \quad n = 0, 1, 2, \dots, \end{aligned} \quad (6.3.12)$$

where the first iterate is $y_0(t) = y_0$. In the context of ODEs, the sequence of iterates $y_n(t)$ obtained in this way are called *Picard iterates*, and the fixed-point iteration procedure is *Picard iteration*. One is left with the task of showing that the sequence converges to a solution of the given integral equation, or equivalently, the given initial value problem.

It is not difficult to prove the following theorem [see, e.g., Birkhoff & Rota 1965]:

Theorem. Let $f(t, y)$ be a continuous function and assume that f satisfies the Lipschitz condition

$$|f(t, y) - f(t, z)| \leq k|y - z| \quad \text{for all } y, z \in \mathbb{R} \quad \text{and} \quad |t - t_0| \leq T.$$

Then, for any y_0 the Picard iterates $y_n(t)$ converge uniformly on $|t - t_0| \leq T$, and the limiting function is a unique solution to the initial value problem (6.3.10) on the interval $|t - t_0| \leq T$.

6.3.2 Semilinear Equations

With the ideas in the preceding paragraphs as motivation, we now consider the question of existence of a solution to the nonlinear initial value problem (6.3.1)–(6.3.2). The last example on ordinary differential equations suggests the strategy of writing the initial value problem as an integral equation. A hint on how to accomplish this is found in the solution (6.3.5) to the linear, nonhomogeneous problem (6.3.3)–(6.3.4). For the moment, assume that f and u_0 are bounded, continuous functions on \mathbb{R} . If $u(x, t)$ is a solution of (6.3.1)–(6.3.2), then

$$u_t(x, t) - Du_{xx}(x, t) = f(u(x, t)) = g(x, t), \quad u(x, 0) = u_0(x). \quad (6.3.13)$$

Thus, from (6.3.5) we expect that

$$u(x, t) = \int_{\mathbb{R}} K(x - y, t)u_0(y) dy + \int_0^t \int_{\mathbb{R}} K(x - y, t - s)g(y, s) dy ds,$$

or

$$u(x, t) = \int_{\mathbb{R}} K(x - y, t)u_0(y) dy + \int_0^t \int_{\mathbb{R}} K(x - y, t - s)f(u(y, s)) dy ds, \quad (6.3.14)$$

which is a nonlinear integral equation for $u = u(x, t)$. One can show that $u = u(x, t)$ is a solution of (6.3.1)–(6.3.2) if, and only if, $u = u(x, t)$ is a solution of (6.3.14). Equation (6.1.14) has the form

$$u = \Phi(u),$$

where Φ is the mapping defined on the set of bounded continuous functions by

$$\begin{aligned} \Phi(u)(x, t) &\equiv \int_{\mathbb{R}} K(x - y, t)u_0(y) dy \\ &+ \int_0^t \int_{\mathbb{R}} K(x - y, t - s)f(u(y, s)) dy ds. \end{aligned}$$

Thus one can define an iterative process by $u_{n+1} = \Phi(u_n)$, or

$$\begin{aligned} u_{n+1}(x, t) &= \int_R K(x - y, t) u_0(y) dy \\ &\quad + \int_0^t \int_R K(x - y, t - s) f(u_n(y, s)) dy ds, \\ n &= 1, 2, 3, \dots, \end{aligned} \tag{6.3.15}$$

with

$$u_0(x, t) = \int_R K(x - y, t) u_0(y) dy. \tag{6.3.16}$$

The initial approximation $u_0(x, t)$ given by (6.3.16) is just the first term on the right side of (6.3.14), and it is the solution to the linear, homogeneous diffusion equation with initial condition $u(x, 0) = u_0(x)$. There should be no confusion in using the notation $u_0(x, t)$ and $u_0(x)$ for these two different objects; we shall always indicate the arguments so that the context is clear. We are now in a position to prove an existence theorem for the initial value problem (6.3.1)–(6.3.2) under suitable assumptions. This version of the theorem was adapted from Kolmogorov et al. (1937) in their celebrated paper on nonlinear reaction–diffusion equations. In the proof of the following theorem we use the fact that the diffusion kernel K is strictly positive and that

$$\int_R K(x - y, t - s) dx = 1 \quad \text{for all } y \text{ and all } s < t. \tag{6.3.17}$$

Theorem. (*Existence–Uniqueness*) Consider the initial value problem (6.3.1)–(6.3.2) where $u_0(x)$ is a bounded continuous function on \mathbb{R} , and where f is a bounded continuous function on \mathbb{R} that satisfies the global Lipschitz condition

$$|f(u) - f(v)| \leq k|u - v| \quad \text{for all } u, v \in \mathbb{R}, \tag{6.3.18}$$

where k is a positive constant independent of u and v . Then for any $T > 0$, there exists a unique, bounded, solution $u(x, t)$ of (6.3.1)–(6.3.2) for $x \in \mathbb{R}$ and $0 \leq t \leq T$.

Proof: We show that the sequence $u_n(x, t)$ defined by (6.3.15) and (6.3.16) converges uniformly on $\mathbb{R} \times [0, T]$ to a function that is a solution of (6.3.1)–(6.3.2). First we note that the Lipschitz condition (6.3.18) implies that

$$\begin{aligned} |f(u_0(x, t))| &\leq |f(0)| + k|u_0(x, t)| \leq (1 + k)m, \\ m &= \max\{|f(0)|, \sup|u_0(x)|\}. \end{aligned}$$

Then, from (6.3.15) and (6.3.17) we obtain

$$\begin{aligned} |u_1(x, t) - u_0(x, t)| &\leq \int_0^t \int_{\mathbb{R}} K(x - y, t - s) |f(u_0(y, s))| dy ds \\ &\leq \int_0^t (1 + k)m ds = Mt \end{aligned} \quad (6.3.19)$$

For convenience we introduce the notation

$$M_n(t) = \sup\{|u_n(x, s) - u_{n-1}(x, s)| : x \in \mathbb{R}, s \leq t\}, \quad t \leq T.$$

Then, from (6.3.19), we get

$$M_1(t) \leq Mt, \quad 0 \leq t \leq T. \quad (6.3.20)$$

Next we obtain a bound for $|u_{n+1} - u_n|$. From (6.3.15) we have

$$\begin{aligned} &|u_{n+1}(x, t) - u_n(x, t)| \\ &\leq \int_0^t \int_{\mathbb{R}} K(x - y, t - s) |f(u_n(y, s)) - f(u_{n-1}(y, s))| dy ds \\ &\leq \int_0^t \int_{\mathbb{R}} K(x - y, t - s) k |u_n(y, s) - u_{n-1}(y, s)| dy ds \\ &\leq \int_0^t \int_{\mathbb{R}} K(x - y, t - s) k M_n(s) dy ds \\ &= k \int_0^t M_n(s) ds. \end{aligned}$$

Taking the supremum, we obtain

$$M_{n+1}(t) \leq k \int_0^t M_n(s) ds, \quad 0 \leq t \leq T, \quad n = 1, 2, 3, \dots \quad (6.3.21)$$

Using (6.3.20) and (6.3.21), we can obtain bounds on each M_n . To this end, observe that

$$\begin{aligned} M_2(t) &\leq k \int_0^t Ms ds = \frac{kMt^2}{2}, \\ M_3(t) &\leq k \int_0^t \frac{kMs^2}{2} ds = \frac{M(kt)^3}{3!}, \end{aligned}$$

and so on, to finally obtain

$$M_n(t) \leq \frac{Mk^n t^n}{n!}, \quad 0 \leq t \leq T. \quad (6.3.22)$$

Equation (6.3.22) implies that the sequence $u_n(x, t)$ converges uniformly on $\mathbb{R} \times [0, T]$ to some continuous, bounded function $u(x, t)$. Then, taking the limit of

both sides of (6.3.15) and using the uniform convergence to pull the limit under the integral, we observe that the limit function $u(x, t)$ satisfies the integral equation (6.3.14) and therefore is a solution to the initial value problem (6.3.1)–(6.3.2).

Uniqueness of $u(x, t)$ can be proved by the standard contradiction argument. Let $v(x, t)$ be another bounded, continuous solution of (6.3.14). Then

$$\begin{aligned} |u(x, t) - v(x, t)| &\leq \int_0^t \int_{\mathbb{R}} K(x - y, t - s) |f(u(y, s)) - f(v(y, s))| dy ds \\ &\leq \int_0^t \int_{\mathbb{R}} K(x - y, t - s) k |u(y, s) - v(y, s)| dy ds. \end{aligned}$$

Now let $M(t) = \sup |u(x, s) - v(x, s)|$, taken over all $x \in \mathbb{R}$ and $s \leq t$. It follows that

$$M(t) \leq k \int_0^t \int_{\mathbb{R}} K(x - y, t - s) M(s) dy ds = k \int_0^t M(s) ds,$$

which is impossible unless $M(t) \equiv 0$. Therefore, $u = v$ and solutions are unique, completing the proof. \square

The preceding existence–uniqueness theorem has a very strict hypothesis, namely, the global Lipschitz condition on the reaction term $f(u)$. The reaction term $f(u) = u(1 - u)$ for the Fisher equation, for example, does not satisfy this because, in this case

$$|f(u) - f(v)| = |u - u^2 - v + v^2| = |1 - u - vu - v|,$$

and the right side cannot be bounded by $k|u - v|$ for all u and v in \mathbb{R} . However, if u and v are restricted, the factor $|1 - u - v|$ can be bounded and we obtain a *local form* of a Lipschitz condition. Thus, we want to formulate an existence–uniqueness theorem by weakening the hypothesis to a local Lipschitz condition, thereby obtaining a theorem with broader applicability. We also take the opportunity to introduce some basic concepts from elementary functional analysis that permits us to resolve existence–uniqueness issues in a variety of contexts that includes many of the equations occurring in applied mathematics. As mentioned earlier, this unifying approach shows the power of abstraction in mathematics and should convince even the most skeptical of the value of abstract methods.

6.3.3 Normed Linear Spaces

We recall from elementary linear algebra that a *real linear space* (or vector space) is a set of objects (numbers, vectors, matrices, functions, or whatever)

on which an addition is defined and in which multiplication of the objects by scalars (real numbers) is defined. In other words, if X denotes the set of objects and u and v are in X , then $u + v$ is defined and belongs to X , and au is defined and belongs to X for any u in X and any real number a . The addition must satisfy the usual rules of addition: commutativity, associativity, existence of a zero element, and existence of an inverse element $(-u)$ corresponding to each u in X . More specifically:

1. $u + v = v + u$, $u + (v + w) = (u + v) + w$.
2. There is an element 0 in X such that $u + 0 = u$ for all u in X .
3. For each u in X there is an inverse (denoted by $-u$) in X for which $u + (-u) = 0$.

Similarly, the scalar multiplication must satisfy a minimal set of rules:

4. $1u = u$ for all u in X .
5. $(ab)u = a(bu)$, $(a + b)u = au + bu$, $a(u + v) = au + av$, for all u and v in X and all a and b in \mathbb{R} .

Any set of objects on which addition and scalar multiplication is defined and the rules (1)–(5) hold is called a real *linear space*. Thus a linear space is a set on which there is well-defined algebraic structure.

Given a linear space X with its algebraic structure, one may also impose geometry, that is, a measure of the size of a given element of the space. If to each u in a linear space X there is associated a nonnegative number, denoted by $\|u\|$, satisfying the three properties

$$\|u\| = 0 \Leftrightarrow u = 0; \quad \|au\| = |a|\|u\| \quad \text{for } a \in \mathbb{R}, u \in X; \quad (6.3.23)$$

$$\|u + v\| \leq \|u\| + \|v\| \quad \text{for } u, v \in \mathbb{R}, \quad (6.3.24)$$

then X is a *normed linear space*, and the measure of size $\|\cdot\|$ is called a *norm* on X . In a natural way, a *distance function* may be defined in a normed linear space by the formula

$$\text{dist}(u, v) = \|u - v\|.$$

Common examples of normed linear spaces are

1. $X = \text{the real numbers } \mathbb{R}$, with $\|x\| = |x|$; the distance between x and y is $|x - y|$.
2. $X = \mathbb{R}^n = \{(x_1, x_2, \dots, x_n) | x_i \in \mathbb{R}\}$ = the set of all ordered n -tuples of real numbers; one norm is the *Euclidean norm* $\|(x_1, \dots, x_n)\| = (x_1^2 + \dots + x_n^2)^{1/2}$; two other norms are $\|(x_1, \dots, x_n)\| = |x_1| + \dots + |x_n|$ and $\|(x_1, \dots, x_n)\| = \max\{|x_1|, \dots, |x_n|\}$. Thus it is possible to define different norms on a given linear space.

3. $X =$ the set of real n by n matrices, with $\|M\| = \Sigma |m_{ij}|$, where $M = (m_{ij})$.

Many of the normed linear spaces of interest in applied mathematics are spaces of functions. For example:

4. $X =$ the set of continuous functions defined on a closed interval $[a, b]$. This linear space is denoted by $C[a, b]$. It can be made into a normed linear space by defining $\|u\| = \max |u(x)|$, where the maximum is taken over $a \leq x \leq b$. Another norm on $C[a, b]$ is $\|u\| = \int_a^b |u(x)| dx$.
5. $X =$ the set of bounded, uniformly continuous functions on \mathbb{R} . This linear space is often denoted by BC . A norm on BC is $\|u\| = \sup |u(x)|, x \in \mathbb{R}$.

The reader should verify that some of the norms defined in the preceding examples do indeed satisfy the conditions (6.3.23).

It is a fundamental property of the real number system that Cauchy sequences converge. Specifically, if x_n is a sequence of real numbers having the property that for any $\varepsilon > 0$ there exists a positive integer N such that $|x_{n+p} - x_n| < \varepsilon$ for all $n > N$ and all $p > 0$, the sequence x_n must converge to a real number. This fact follows from the completeness axiom of the real-number system, which states that every bounded set of real numbers must have a least upper bound (supremum) and a greatest lower bound (infimum). We can generalize this notion to an arbitrary normed linear space X in the following manner. Let u_n be a sequence of elements in a normed linear space X . We say that u_n converges to an element u if for any $\varepsilon > 0$ there is a positive integer N such that $\|u_n - u\| < \varepsilon$ for all $n > N$; this type of convergence is also called *norm convergence*, or convergence in the norm topology. We say that u_n is a *Cauchy sequence* in X if, and only if, for any $\varepsilon > 0$ there is a positive integer N such that $\|u_{n+p} - u_n\| < \varepsilon$ for all $n > N$ and $p > 0$. A normed linear space X is said to be *complete* if it has the property that every Cauchy sequence in X converges to an element of X . A complete normed linear space is called a *Banach space*.

Example. Consider the normed linear space $C[0, 1]$ of continuous functions on $[0, 1]$ with the norm $\|u\| = \max |u(x)|$. It is not difficult to prove that $C[0, 1]$ with this norm is a Banach space. On the other hand, if the norm is defined by $\|u\| = \int_0^1 |u(x)| dx$, then $C[0, 1]$ with this norm is not complete (the reader is asked to demonstrate this fact in an exercise at the end of the section); however, $C[0, 1]$ with this norm may be completed in much the same way that the rational numbers are completed to obtain the real numbers, namely, by appending to $C[0, 1]$ the limits of all Cauchy sequences of functions in $C[0, 1]$. The resulting space is denoted by $L^1[0, 1]$ and is known as the space of Lebesgue integrable functions on $[0, 1]$. We remark that limits of Cauchy sequences in $C[0, 1]$ may not

be Riemann integrable, and therefore the integral in the definition of the norm must be generalized to the Lebesgue integral. In a similar manner, the $L^p[a, b]$ spaces are Banach spaces that arise by completing the continuous functions in the norm

$$\|u\|_p = \left[\int_a^b |u(x)|^p dx \right]^{1/p}.$$

Now we state the fundamental result, which is the vehicle to an existence theorem for nonlinear partial differential equations. Let X be a Banach space and suppose that $\phi : X \rightarrow X$ is a mapping on X . If ϕ has the property that $\|\phi(u) - \phi(v)\| \leq \alpha \|u - v\|$ for all u and v in X , for some positive constant $\alpha < 1$, then ϕ is called a *contraction mapping*. Then we have the following theorem.

Theorem. (Banach Fixed-Point Theorem) If $\phi : X \rightarrow X$ is a contraction mapping on a Banach space X , then ϕ has precisely one fixed point [i.e., there exists a unique $u \in X$ such that $\phi(u) = u$].

Proof: Earlier in this section we proved this theorem for a contraction mapping on the reals. The general proof for a Banach space is exactly the same argument with the absolute values replaced by norms. \square

Remark. In applications, ϕ is usually not a contraction on the entire Banach space, but rather only on a closed subset in the space. (A set G in a Banach space is *closed* if any convergent sequence in G has its limit in G .) It is straightforward to prove that the fixed-point theorem remains valid on closed subsets of a Banach space.

6.3.4 General Existence Theorem

Using the machinery of a Banach space and the fixed-point theorem, we now formulate and prove an existence theorem for the semilinear problem (6.3.1)–(6.3.2), where only a local Lipschitz condition is required. First we establish some notation. Consider a function $u = u(x, t)$. For each *fixed* time t we can regard u as a function of x (on \mathbb{R}) representing a slice of the surface $u(x, t)$ at time t , or equivalently, a time snapshot of the wave. The underlying Banach space B in the following formulation will specify the type of functions that these time snapshots or wave profiles can be. Therefore, let B be the Banach space of all bounded, continuous functions $u(x, t)$ on \mathbb{R} (t fixed), and let $\|u(\cdot, t)\|_B$ denote the norm of a function $u(x, t)$ in B . Specifically, we take

$$\|u(\cdot, t)\|_B = \sup_{x \in \mathbb{R}} |u(x, t)| \quad \text{for } t \text{ fixed.} \quad (6.3.25)$$

This norm is denoted as the *sup norm* (this norm is usually denoted with the subscript ∞ rather than B ; however, we shall maintain this generic notation because the following results can be extended to other Banach spaces with other norms). Now, let T be positive and let $C([0,T]; B)$ be the set of all continuous functions defined on the interval $0 \leq t \leq T$ that have values in the Banach space B . Thus, to each $t \in [0, T]$ we associate a bounded, continuous function $u(x, t)$ on $x \in \mathbb{R}$ (t fixed) that is an element of the Banach space B . The set $C([0,T]; B)$, whose elements will be denoted by either u or $u(x, t)$, is a Banach space in its own right with norm

$$\|u\| = \sup_{t \in [0, T]} \|u(\cdot, t)\|_B. \quad (6.3.26)$$

We must be careful not to confuse the various objects that we have defined; there are two Banach spaces here, B and $C([0, T], B)$. The objects in B are denoted by either $u(x, t)$ or $u(\cdot, t)$ and are considered to be a functions of x , with t fixed; the norm in B is denoted by the subscript B , and we prefer to use the norm notation $\|u(\cdot, t)\|_B$ with the generic *dot* argument because the spatial variable has been “sup-ed out” in taking the norm. Regarded as an object in $C([0, T], B)$, $u = u(x, t)$ is a function of both x and t , such that for each t on $[0, T]$ the association of $u(x, t)$ in B is continuous; in this space there is no subscript on the norm symbol. Another way of thinking about the objects in $C([0, T], B)$ is as continuous curves in the Banach space B ; that is, to each t in the parameter interval $[0, T]$ we associate an element $u(\cdot, t)$ of B . Interpreted still differently, to each t we associate a wave profile in B , and the totality of all the wave profiles forms the surface $u = u(x, t)$. We also introduce the convolution operation

$$(K * u)(x, t) = \int_{\mathbb{R}} K(x - y, t)u(y, t) dy,$$

where $K(x, t)$ is the diffusion kernel. $K * u$ is the convolution of K with u , and $(K * u)(\cdot, t)$ also belongs to B . Perhaps the greatest difficulty in proving the existence theorem is sorting out the notation.

Theorem. (Local Existence) Consider the initial value problem (6.3.1)–(6.3.2) where $u_0 \in B$ and f satisfies the following conditions

1. $f \in C^1(\mathbb{R})$.
2. $f(0) = 0$, and for each fixed t in $[0, T]$, $f(u(x, t)) \in B$ for each $u(x, t) \in B$.
3. For any $M > 0$ there exists a constant k , depending only on M , such that

$$\|f(u(\cdot, t)) - f(v(\cdot, t))\|_B \leq k\|u(\cdot, t) - v(\cdot, t)\|_B$$

for all $t \in [0, T]$ and all $u(x, t)$ and $v(x, t)$ in B with $\|u(\cdot, t)\|_B \leq M$ and $\|v(\cdot, t)\|_B \leq M$.

Then there exists $t_0 > 0$, where t_0 depends only on f and $\|u_0(\cdot)\|_B$, such that the initial value problem (6.3.1)–(6.3.2) has a unique solution $u = u(x, t)$ in $C([0, t_0]; B)$, and $\|u\| \leq 2\|u_0(\cdot)\|_B$.

Proof: We define a closed subset G of the Banach space $C([0, t_0]; B)$ and show that the mapping

$$\begin{aligned}\phi(u)(x, t) = & \int_{\mathbb{R}} K(x - y, t)u_0(y) dy \\ & + \int_0^t \int_{\mathbb{R}} K(x - y, t - s)f(u(y, s)) dy ds\end{aligned}\quad (6.3.27)$$

is a contraction mapping on G . Then we apply the Remark after the statement of the Banach fixed-point theorem to produce a solution to $u = \phi(u)$, which by our previous remarks is a solution to the initial value problem. To this end, define

$$\begin{aligned}G = \{u \in C([0, T]; B) : \\ \|u(\cdot, t) - (K * u_0)(\cdot, t)\|_B \leq \|u_0(\cdot)\|_B, \text{ for } 0 \leq t \leq t_0\},\end{aligned}$$

where $t_0 = 1/2k$. The set G is closed and nonempty (e.g., zero is in G). Also, the defining property of G , the triangle inequality, and the fact that

$$\|(K * u)(\cdot, t)\|_B \leq \|u(\cdot, t)\|_B, \quad (6.3.28)$$

imply that $\|u(\cdot, t)\|_B \leq 2\|u_0(\cdot)\|_B$, whereupon taking the supremum on t gives

$$\|u\| \leq 2\|u_0(\cdot)\|_B. \quad (6.3.29)$$

This proves the last statement in the theorem. Now, we have from rule(2) listed at the beginning of Section 6.3.3, that for any $0 \leq t \leq t_0$,

$$\begin{aligned}\|f(u(\cdot, t)) - f(v(\cdot, t))\|_B & \leq k\|u(\cdot, t) - v(\cdot, t)\|_B \\ & \leq k\|u - v\|\end{aligned}\quad (6.3.30)$$

Because k depends only on the sup norm of u and v , and of course on f , it is clear that t_0 depends only on f and the sup norm of u_0 , by virtue of the inequality (6.3.29).

The fact that ϕ maps G into G [i.e., $\phi(u) \in G$ if $u \in G$] follows from the sequence of inequalities

$$\begin{aligned}\|\phi(u)(\cdot, t) - (K * u_0)(\cdot, t)\|_B &= \left\| \int_0^t (K * f(u))(\cdot, t-s) ds \right\|_B \\ &= \int_0^t \|(K * f(u))(\cdot, t-s)\|_B ds \\ &\leq \int_0^t \|f(u(\cdot, s))\|_B ds \\ &\leq \int_0^t 2k \|u_0(\cdot)\|_B ds \\ &= 2kt \|u_0(\cdot)\|_B \\ &< \|u_0(\cdot)\|_B.\end{aligned}$$

Now we prove that ϕ is a contraction. We have

$$\begin{aligned}\|\phi(u)(\cdot, t) - \phi(v)(\cdot, t)\|_B &\leq \sup_{x \in \mathbb{R}} \int_0^t \int_{\mathbb{R}} K(x-y, t-s) |f(u(y, s)) - f(v(y, s))| dy ds \\ &\leq \int_0^t \|f(u(\cdot, s)) - f(v(\cdot, s))\|_B ds \\ &\leq \int_0^t k \|u - v\| ds \\ &\leq kt_0 \|u - v\| = \frac{\|u - v\|}{2}.\end{aligned}$$

Taking the supremum over $t \in [0, T]$ then gives

$$\|\phi(u) - \phi(v)\| \leq \frac{1}{2} \|u - v\|.$$

Therefore, ϕ is contraction on the closed subset G of the Banach space $C([0, T]; B)$. By the Remark following the fixed-point theorem it follows that there is a unique fixed point in G [i.e., a unique solution to the initial value problem (6.3.1)–(6.3.2) in G].

It remains to show that there are no solutions outside the set G . This fact results from the following general argument. If u and v are two solutions lying in $C([0, T]; B)$ that satisfy the same initial condition, then

$$u(x, t) - v(x, t) = \int_0^t \int_{\mathbb{R}} K(x-y, t-s) [f(u(y, s)) - f(v(y, s))] dy ds,$$

from which it easily follows that

$$\|u(\cdot, t) - v(\cdot, t)\|_B \leq k \int_0^t \|u(\cdot, s) - v(\cdot, s)\|_B ds.$$

Multiplying this inequality by e^{-kt} gives the result that

$$\frac{d}{dt}\{e^{-kt}\|u(\cdot, t) - v(\cdot, t)\|_B\} \leq 0,$$

which in turn implies that $\|u(\cdot, t) - v(\cdot, t)\|_B = 0$, or $u = v$. Consequently, the theorem is proved. \square

The theorem is only a local existence result, guaranteeing a solution for $0 \leq t \leq t_0$, for some t_0 . Under certain conditions we may extend the solution to any finite time T . We have the following result.

Theorem. (Global Existence) Under the same hypotheses of the last theorem, assume in addition that the solution is a priori bounded in the sup norm on $0 \leq t \leq T$; that is, assume that there is a constant $C > 0$ depending only on $\sup_{x \in R}|u_0(x)|$ such that if u is any solution of (6.3.1)–(6.3.2) in $0 \leq t \leq T$, then $\sup_{x \in R}|u(x, t)| \leq C$. Then the solution to (6.3.1)–(6.3.2) exists for all t in $[0, T]$, and $u(x, t) \in B$. Here T may be infinity, giving global existence.

Proof: The local theorem guarantees a solution u on $[0, t_0]$. Then we can apply the local theorem again with initial condition $u(x, t_0)$ to get a solution on $[t_0, 2t_0]$. Continuing in this manner we can obtain, after a finite number of steps, a solution on $[0, T]$. \square

Example. Consider the initial value problem for Fisher's equation:

$$\begin{aligned} u_t - Du_{xx} &= u(1 - u), \quad x \in \mathbb{R}, \quad t > 0, \\ u(x, 0) &= u_0(x), \quad x \in \mathbb{R}. \end{aligned}$$

Here $f(u) = u(1 - u)$. Clearly, $f(0) = 0$, f is smooth, and

$$\sup|f(u) - f(v)| = \sup(|1 - u - vu - v|) \leq k \sup|u - v|,$$

where $k = 1 + \sup|u| + \sup|v|$, and the sup is taken over $x \in \mathbb{R}$. Therefore, the local Lipschitz condition holds. Further, if $u \in B$, so is $f(u)$. The local existence theorem guarantees a solution on $[0, t_0]$, for some positive t_0 . To prove existence for all $t > 0$ we would need to have an *a priori* bound on solutions. Such a bound can be obtained from the comparison theorems in the next section. \square

Proofs proceed in the same way for *systems* of reaction–diffusion equations of the form

$$\begin{aligned} \mathbf{u}_t - D\mathbf{u}_{xx} &= \mathbf{f}(\mathbf{u}), \quad x \in \mathbb{R}, \quad t > 0, \\ \mathbf{u}(x, 0) &= \mathbf{u}_0(x), \quad x \in \mathbb{R}, \end{aligned}$$

where \mathbf{u} and \mathbf{f} are n -vectors, and $D = \text{diag}(d_1, \dots, d_n)$ is a diagonal matrix with positive diagonal entries (see Smoller 1994, Chapter 14, Section A). Now, the Banach space B is a space of vector-valued functions on \mathbb{R} with values in \mathbb{R}^n , and each component is bounded, continuous function. The equivalent integral equation is

$$\mathbf{u}(x, t) = \int_R \mathbf{K}(x - y, t) \mathbf{u}_0(y) dy + \int_0^t \int_R \mathbf{K}(x - y, t - s) \mathbf{f}(\mathbf{u}(y, s)) dy ds,$$

where $\mathbf{K}(x, t)$ is a diagonal matrix with diagonal elements

$$k_i(x, t) = \frac{1}{\sqrt{4\pi d_i t}} e^{-x^2/4d_i t}, \quad i = 1, \dots, n.$$

EXERCISES

1. Consider the linear initial value problem

$$u_t - Du_{xx} = au, \quad x \in \mathbb{R}, \quad t > 0; \quad u(x, 0) = u_0(x), \quad x \in \mathbb{R},$$

where $u_0 \in B$, where a and D are positive constants, and where B is the space of all continuous, bounded functions on \mathbb{R} with the sup norm. Use the machinery in this section to prove that a solution exists for all $t \geq 0$.

2. Consider the initial value problem

$$u_t - Du_{xx} = u^p, \quad x \in \mathbb{R}, t > 0; \quad u(x, 0) = u_0(x), \quad x \in \mathbb{R},$$

where $p > 1$, and u_0 is a bounded continuous function on \mathbb{R} .

- (a) Use the local existence theorem to prove that a unique solution exists for $0 \leq t \leq t_0$.
- (b) Show that a global solution does not, in general, exist.

3. Prove that the linear space $C[0, 1]$ with norm $\|u\| = \int_0^1 |u(x)| dx$ is not a Banach space by considering the Cauchy sequence of continuous functions

$$u_n(x) = \begin{cases} 0, & 0 \leq x \leq \frac{1}{2} \\ 1, & \frac{1}{2} + \frac{1}{n} \leq x \leq 1 \\ \text{linear,} & \frac{1}{2} \leq x \leq \frac{1}{2} + \frac{1}{n} \end{cases}.$$

4. Consider the advection–diffusion problem

$$\begin{aligned} u_t - Du_{xx} &= F(u)_x, \quad x \in \mathbb{R}, \quad t > 0, \\ u(x, 0) &= u_0(x), \quad x \in \mathbb{R}. \end{aligned}$$

- (a) Show that the integral representation of a solution to this initial value problem is given by

$$\begin{aligned} u(x, t) = & \int_{\mathbb{R}} K(x - y, t) u_0(y) dy \\ & + \int_0^t \int_{\mathbb{R}} K_x(x - y, t - s) F(u(y, s)) dy ds, \end{aligned}$$

where K is the diffusion kernel.

- (b) Let $F(u) = u^2/2$. Formulate and prove a local existence theorem for the initial value problem.

6.4 Maximum Principles and Comparison Theorems

Section 6.3 dealt with the question of existence of solutions to the initial value problem for certain nonlinear reaction–diffusion equations. Now we want to formulate two other theoretical concepts associated with diffusion problems and parabolic equations: maximum principles and comparison theorems.

A maximum principle is a statement about where a solution to a problem acquires its maximum value. There are correspondingly minimum principles. The problems can be initial value problems or boundary value problems, or mixed problems. For example, diffusion processes, by their very physical nature, tend to smear out the density function u , and this precludes clumping of the solution in the interior of the given spacetime domain; therefore, diffusion problems have solutions whose maximum occurs on the boundary of the domain. Comparison theorems, on the other hand, are statements that compare the solutions to two similar problems, say, differing only in their boundary or initial values. If one of the problems can be solved, a comparison theorem can give bounds on the solution of the problem that perhaps cannot be solved. This information can be used to obtain the asymptotic behavior of the solution as t increases or produce a priori bounds that guarantee the existence of global solutions.

6.4.1 Maximum Principles

A *maximum principle* is a theorem indicating where a solution to a given PDE assumes its maximum value. A simple example involving the linear diffusion equation illustrates the concept.

Example. Consider a solution $u = u(x, t)$ to the equation

$$Hu \equiv u_t - ku_{xx} = f(x, t) \quad (6.4.1)$$

on the bounded spacetime domain D shown in Figure 6.11. Assume that D is open and its boundary consists of a segment S (excluding the endpoints) at time $t = T > 0$ and a lower boundary B that does include the endpoints of S . We also assume D is convex (contains the line segment connecting any two points in the domain) and lies strictly in the upper half-plane $t > 0$. Further, assume that $u \in C(\overline{D})$ and $u \in C^2(D)$, where $\overline{D} = D \cup S \cup B$ is the closure of D , that is, D along with its boundary. Portions of \overline{D} can lie on the initial time line $t = 0$; for example, D may be a rectangular region with a segment of B lying along $t = 0$. Now, if the source term f in (6.4.1) is strictly negative, that is

$$Hu = f < 0, \quad (x, t) \in D, \quad (6.4.2)$$

it is easy to show that a local maximum of u cannot occur in D or on S . By way of contradiction, assume that there is such a point (x_0, t_0) in $D \cup S$. Then, from calculus, we must have

$$u_x(x_0, t_0) = 0, \quad U_{xx}(x_0, t_0) \leq 0, \quad u_t(x_0, t_0) \geq 0.$$

But then $Hu \geq 0$ at (x_0, t_0) , contradicting (6.4.2). Consequently, if the source term is strictly negative, the maximum value of any solution to (6.4.1) must occur on B , the lower portion of the boundary. \square

We extend the ideas in this example to other cases with more general differential operators, or perhaps allowing equality in (6.4.2). Therefore, define the operator P defined by

$$Pu = a(x, t)u_{xx} + b(x, t)u_x - u_t, \quad (6.4.3)$$

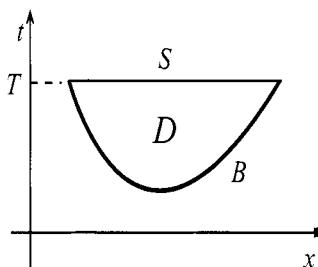


Figure 6.11 Open domain D .

where a and b are continuous functions on \bar{D} . We say that P is *uniformly parabolic* in D if $a(x, t) \geq m > 0$ in D , for some constant m . First, we remark that the result of the last example can be extended at once to the operator P . If $Pu > 0$, a local maximum of u cannot occur in D or on S because, at such a point, $u_{xx} \leq 0$, $u_x = 0$, and $u_t \geq 0$, which violate $Pu > 0$.

We now state and prove the *weak maximum principle*, which essentially states that functions u satisfying $Pu + c(x, t)u \geq 0$ on D , where $c \leq 0$ on D , must assume their maximum on B . The result is called *weak* because it does not preclude the maximum also occurring at an interior point in D . There is also a *strong maximum principle*, which we shall formulate later, but not prove, that does preclude a maximum occurring in D unless u is constant up to that point of time.

Theorem. (*Weak Maximum Principle*) Let $u \in C(\bar{D}) \cap C^2(D)$ be a solution of the equation

$$Pu + cu = f(x, t), \quad (x, t) \text{ in } D, \quad (6.4.4)$$

where c and f are $C(D)$, $c(x, t) \leq 0$, $f(x, t) \geq 0$ in D , and P is uniformly parabolic in D . Then u assumes its positive maximum on B , provided that it exists.

Proof: The idea of the proof is to assume the contrary and construct a function that violates the usual calculus conditions for a maximum. Most proofs of maximum principles use this strategy in one way or another, but some proofs require clever constructions.

Suppose that u has a positive maximum in \bar{D} . To show that this maximum must occur somewhere on B , we first define the auxiliary function

$$v(x, t) = u(x, t) - \varepsilon t.$$

If u has a positive maximum in \bar{D} , so does v for ε sufficiently small. By way of contradiction, assume that the positive maximum of v occurs at a point (x_0, t_0) in $D \cup S$. Then

$$v_x(x_0, t_0) = 0, \quad v_{xx}(x_0, t_0) \leq 0.$$

Now calculate v_t . We have

$$\begin{aligned} v_t &= u_t - \varepsilon = au_{xx} + bu_x + cu - f - \varepsilon \\ &= av_{xx} + bv_x + c(v + \varepsilon t) - f - \varepsilon \\ &\leq cv + \varepsilon tc - \varepsilon \quad \text{at } (x_0, t_0) \\ &\leq -\varepsilon. \end{aligned}$$

Next, choose a number $d > 0$ such that $v_t(x_0, t) \leq -\varepsilon/2$ for $t_0 - d \leq t \leq t_0$. Then integration gives

$$v(x_0, t_0) - v(x_0, t_0 - d) \leq \int_{t_0-d}^{t_0} -\frac{\varepsilon}{2} dt = -\frac{\varepsilon d}{2} < 0.$$

This contradicts the assumption that v has a positive maximum at (x_0, t_0) . So the positive maximum of v must occur on B . Now we can make the following estimates:

$$\begin{aligned} \max_{\bar{D}} u &= \max_{\bar{D}}(v + \varepsilon t) \leq \max_{\bar{D}} v + \varepsilon T \\ &= \max_B v + \varepsilon T = \max_B(u - \varepsilon t) + \varepsilon T \leq \max_B u + \varepsilon T. \end{aligned}$$

Because ε is arbitrary, $\max_{\bar{D}} u \leq \max_B u$, and therefore the maximum must occur someplace on B , completing the proof. \square

If $c = 0$ and $f = 0$, we have the linear homogeneous equation $Pu = 0$. If u is a solution to this equation, so is $u + \text{constant}$, for any constant. Thus, for this equation we may always assume that u has a positive maximum, and this hypothesis in the weak maximum principle may be deleted. We therefore have the following corollary.

Corollary. Let $u \in C(\bar{D}) \cap C^2(D)$ be a solution of

$$Pu = 0 \quad \text{on } D,$$

where P is uniformly parabolic on D . Then u assumes its maximum and minimum on B .

We can take the analysis one step further and prove the following theorems with no assumption on the sign of the function $c = c(x, t)$. This theorem will yield simple proofs of uniqueness theorems, which are requested in the Exercises.

Theorem. Let $u \in C(\bar{D}) \cap C^2(D)$ be a solution of

$$\begin{aligned} Pu + cu &= f(x, t) \quad \text{on } D, \\ u &= 0 \quad \text{on } B, \end{aligned}$$

where $f \geq 0$ on D . Then $u = 0$ in D .

Proof: First consider the case $c \leq 0$ in D . By way of contradiction, assume that u is strictly positive at some point of D . Then u has a positive maximum

in D , and by the weak maximum principle u must have a positive maximum on B , contradicting the hypothesis that $u = 0$ on B . If $u < 0$ at some point in D , then in the same way it must have a negative minimum on B , again a contradiction. So in the case $c \leq 0$ we must have $u = 0$ in D .

Now assume that $c \leq 0$ is not satisfied. In this case choose a constant $\lambda \geq c$ in D and set $u = we^{\lambda t}$. Then w satisfies the equation

$$Pw + (c - \lambda)w = fe^{-\lambda t} \quad \text{in } D$$

$$w = 0 \quad \text{on } B$$

But now $c - \lambda \leq 0$, and therefore the weak maximum principle may be applied to get $w = 0$ in D . Hence we must have $u = 0$ in D , which completes the argument. \square

As mentioned earlier, there is a strong maximum principle that states the consequences of having a maximum inside D . We state, without proof, a simplified version of this theorem that applies to the simply connected, convex region D shown in Figure 6.11. An easily accessible, general version is stated and proved in Smoller (1994), Protter & Weinberger (1967), or Friedman (1964).

Theorem. (*Strong Maximum Principle*) Let $u \in C(\bar{D}) \cap C^2(D)$ be a solution of the equation

$$Pu + c(x, t)u = f(x, t) \quad \text{in } D,$$

where P is uniformly parabolic in D , where $c \leq 0$ and $f \geq 0$ (respectively, $f \leq 0$) in D , and c and f are continuous in D . Let m be a nonnegative maximum (respectively, nonpositive minimum) of u in D , and suppose that $u(x_0, t_0) = m$ at some point of D . Then $u = m$ at all points (x, t) in D with $t < t_0$.

The strong maximum principle is also valid on unbounded domains (that are open and connected), provided that the functions a , b , c , and f are continuous and bounded on D , and the terms max and min in the statement of the theorem above are replaced by sup and inf.

Example. Maximum principles do not necessarily hold for *systems* of diffusion equations. Consider the system

$$u_{xx} - u_t \geq 0, \quad v_{xx} - 9u_x - vt \geq 0,$$

on the unit square $0 \leq x \leq 1$, $0 \leq t \leq 1$. A solution is

$$u(x, t) = -e^{x+t}, \quad v(x, t) = t - 4(x - 0.5)^2.$$

Here, u and v are negative on the bottom and side boundaries, yet $v > 0$ along $x = \frac{1}{2}$. \square

6.4.2 Comparison Theorems

We alluded to the fact that a comparison theorem is a result that compares solutions to similar problems. For example, consider the semilinear equation

$$u_t - a(x, t)u_{xx} - b(x, t)u_x = f(x, t, u) \quad (6.4.5)$$

on the convex domain D shown in Figure 6.11. Here, a and b are continuous functions on \bar{D} and $a(x, t) \geq \mu > 0$. The reaction term f is continuously differentiable on $D \times \mathbb{R}$. Let $v = v(x, t)$ and $u = u(x, t)$ be two solutions of (6.4.5) in D , and suppose that $u \leq v$ on the lower boundary B of D . Then it follows that $u \leq v$ on the entire domain D . To prove this fact, let $w = u - v$. Then $w \leq 0$ on B and

$$aw_{xx} + bw_x - w_t = f(x, t, v) - f(x, t, u) \quad \text{on } D.$$

By the mean value theorem, we obtain

$$\begin{aligned} f(x, t, u) - f(x, t, v) &= f_u(x, t, u^*)w, \\ u^* &= \theta u + (1 - \theta)v, \quad 0 < \theta < 1. \end{aligned}$$

Consequently,

$$\begin{aligned} aw_{xx} + bw_x + f_u w - w_t &= 0 \quad \text{on } D, \\ w &\leq 0 \quad \text{on } B. \end{aligned}$$

We now have a form of the problem for w in which we can apply the weak maximum principle. Suppose first that $f_u \leq 0$ on D . Directly from the weak maximum principle we conclude that $w \leq 0$ in D (otherwise, if w were positive at some point of D , a positive maximum would have to occur on B , a contradiction). If the coefficient f_u does not satisfy the condition $f_u \leq 0$ on D , choose a number λ such that $\lambda \geq f_u$ in D and let $w = W \exp(\lambda t)$. Then W satisfies the problem

$$\begin{aligned} aW_{xx} + bW_x + (f_u - \lambda)W - W_t &= 0 \quad \text{on } D, \\ W &\leq 0 \quad \text{on } B. \end{aligned}$$

Because $f_u - \lambda \leq 0$, we can apply the weak maximum principle to conclude that $W \leq 0$ on D and thus ascertain $w \leq 0$ on D as well. Here is a formal statement of what we proved.

Theorem. (Comparison Theorem) Let $u, v \in C(\bar{D}) \cap C^2(D)$ be two solutions of the semilinear equation (6.4.5) under the stated assumptions. If $u \leq v$ on B , then $u \leq v$ in D .

It is not difficult to generalize and extend this result to general nonlinear equations. To this end we require one bit of terminology. Let us assume that $F(x, t, u, p, r)$ is a given continuously differentiable function of its five arguments. We say that F is *elliptic* with respect to a function $u = u(x, t)$ at a point (x, t) if $F_r(x, t, u(x, t), u_x(x, t), u_{xx}(x, t)) > 0$.

Theorem. (*General Comparison Theorem*) Let D be the domain in Figure 6.11, and consider the equation

$$L[u] = F(x, t, u, u_x, u_{xx}) - u_t = f(x, t) \quad \text{in } D, \quad (6.4.6)$$

where $F = F(x, t, u, p, r)$ is a given continuously differentiable function of its five arguments, and f is continuous on \bar{D} . Suppose that u , w , and W are continuous in \bar{D} and C^2 on D and that u is a solution of (6.4.6). Further, assume that F is elliptic on D with respect to the functions $\theta w + (1 - \theta)u$ and $\theta W + (1 - \theta)u$, where $\theta \in [0, 1]$. If

$$L[W] \leq f(x, t) \leq L[w] \quad \text{in } D,$$

and

$$w \leq u \leq W \quad \text{on } B,$$

then

$$w \leq u \leq W \quad \text{in } D. \quad (6.4.7)$$

Proof: Let $v = w - u$. Then $v \leq 0$ on B , and we must show that $v \leq 0$ in D . We have

$$f(x, t) = F(x, t, u, u_x, u_{xx}) - u_t \leq F(x, t, u, w, w_x, w_{xx}) - w_t,$$

or

$$F(x, t, w, w_x, w_{xx}) - F(x, t, u, u_x, u_{xx}) - v_t \geq 0 \quad \text{on } D.$$

By Taylor's theorem we obtain

$$F(x, t, w, w_x, w_{xx}) - F(x, t, u, u_x, u_{xx}) = F_u v + F_p v_x + F_r v_{xx},$$

where the partial derivatives F_u , F_p , and F_r are evaluated at $(x, t, \theta w + (1 - \theta)u, \theta w_x + (1 - \theta)u_x, \theta w_{xx} + (1 - \theta)u_{xx})$ and $0 < \theta < 1$. Therefore

$$F_r v_{xx} + F_p v_x + F_u v - v_t \geq 0 \quad \text{on } D.$$

This equation is now in the form, with its coefficients as continuous functions of x and t on D , where we can apply the weak maximum principle. As in the proof of the comparison theorem, the principle can be applied regardless of the sign of F_u . All we require is that F_r , evaluated at the point indicated above, be

positive and bounded away from zero on D . Thus the weak maximum principle implies that $v \leq 0$ in D , giving the leftmost inequality in (6.4.7). We leave the right inequality as an exercise. \square

EXERCISES

1. Consider the nonlinear diffusion problem

$$\begin{aligned}(K(u)u_x)_x - u_t &= 0, \quad 0 < x < 1, \quad 0 < t < T, \\ u(x, 0) &= 1 + x(1 - x), \quad 0 < x < 1, \\ u(0, t) &= u(1, t) = 1 \quad 0 < t < T,\end{aligned}$$

where the diffusion coefficient $K = K(u)$ is a continuously differentiable function and $K(u) > 0$. Show that if u is a solution to this problem, then $1 \leq u(x, t) \leq \frac{5}{4}$.

2. Prove the rightmost inequality in (6.4.7).
3. Let D be the infinite domain $\mathbb{R} \times (0, T)$, and let u be continuous on \bar{D} and C^2 on D , and assume that

$$u_t - u_{xx} = 0 \quad \text{in } D.$$

Let $M = \sup_D u$. Prove that if M is finite, $\sup_{x \in \mathbb{R}} u(x, 0) = M$. Hint: Proceed by contradiction and apply the weak maximum principle on a sufficiently large bounded domain to the function $v = u - \varepsilon(2t + x^2)$.

4. Consider the initial-boundary value problem

$$\begin{aligned}u_{xx} + 2u - tut &= 0, \quad 0 < x < \pi, \quad t > 0, \\ u(x, 0) &= 0, \quad 0 < x < \pi, \quad u(0, t) = u(\pi, t) = 0, \quad t > 0.\end{aligned}$$

Use the method of separation of variables to show that this problem has nonzero solutions. Does the maximum principle apply?

5. Consider the equation

$$x^2 u_{xx} - u_t = 0, \quad -1 < x < 1, \quad 0 < t < \frac{1}{2}.$$

Does the maximum principle apply? Hint: Consider $u = -x(x + 2t)$.

6. (a) Let D be the infinite strip $-\infty < x < \infty$, $0 < t < T$, and consider the differential operator P defined by

$$Pu = a(x, t)u_{xx} + b(x, t)u_x + c(x, t)u - u_t,$$

where $a(x, t) \geq \mu > 0$ in D , $c(x, t) \leq 0$ in D , and a , b , and c are continuous functions on D . Prove that if $Pu > 0$ in D , or if $Pu \geq 0$ and $c < 0$ in D , then u cannot have a positive maximum in D .

- (b) Formulate and prove a corresponding theorem regarding a negative minimum.
7. Prove the following one-dimensional version of a comparison theorem. Let f and g be continuous functions on $[a, b]$ with continuous second derivatives on (a, b) . Assume that $f(a) \geq g(a)$, $f(b) \geq g(b)$, and $g''(x) \geq f''(x)$ for x in (a, b) . Prove that $f \geq g$ on (a, b) .

6.5 Energy Estimates and Asymptotic Behavior

Now we introduce an important technique that is applicable to all types of PDEs. This technique and its variations, collectively called *energy methods*, allow us to obtain bounds on certain quantities associated with the solution, prove uniqueness, show that solutions blow up, and obtain other important information about the behavior of solutions. One vehicle for establishing energy estimates is integration by parts, one of the basic techniques in partial differential equations; another is a set of key inequalities.

Let $u = u(x, t)$ be the solution of a given evolution problem on the domain $0 \leq x \leq 1, t \geq 0$. The quantity $E(t)$, defined by

$$E(t) \equiv \int_0^1 u^2(x, t) dx,$$

is called the *energy* at time t ; it is the total area under the wave profile squared. One of the important problems for solutions of evolution equations is to obtain bounds on $E(t)$. For example, if we could obtain an estimate of the form $E(t) \leq C/t$, where C is a positive constant, we could conclude that the energy decays like $1/t$, and in fact, goes to zero as t gets large. In a generalized sense this would mean that the solution u itself would have to go to zero asymptotically. Another quantity of interest is the integral of the square of the gradient, or

$$Q(t) \equiv \int_0^1 u_x^2(x, t) dx.$$

If one can show, for example, that $Q(t)$ tends to zero as t goes to infinity, then u_x must tend to zero in a generalized sense and thus u must tend to a constant. Results like this are important, for example, in showing that reaction-diffusion equations occurring in the biological sciences cannot give rise to spatial patterns that involve density variations.

6.5.1 Calculus Inequalities

To obtain energy and gradient estimates, we require some basic inequalities. One cannot underestimate the role that inequalities play in the theory of PDEs. Here we introduce only a few of the basic inequalities; a detailed development of other important inequalities, for example, the Sobolev inequalities, can be found in the references.

The first result is Young's inequality, which permits a product to be bounded by a sum; it is a generalization of the inequality between the arithmetic and geometric mean.

Young's Inequality. Let f and g be positive quantities, and let p and q be positive real numbers satisfying $1/p + 1/q = 1$. Then

$$fg \leq \frac{f^p}{p} + \frac{g^q}{q}.$$

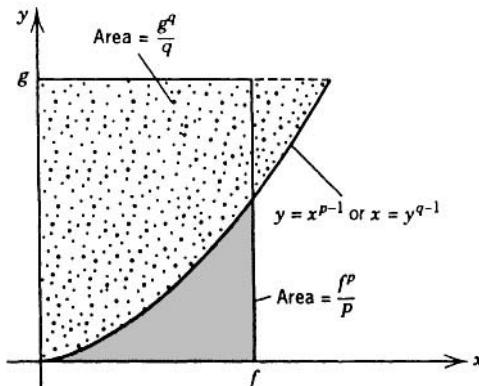


Figure 6.12 Geometric proof of Young's inequality.

Example. If we take $f = \int_a^b u(x)^2 dx$ and $g = \int_a^b v(x)^2 dx$, and $p = q = 2$, then Young's inequality gives

$$\int_a^b u(x)^2 dx \cdot \int_a^b v(x)^2 dx \leq \frac{1}{2} \left[\int_a^b u(x)^2 dx \right]^2 + \frac{1}{2} \left[\int_a^b v(x)^2 dx \right]^2. \quad \square$$

A geometric proof of Young's inequality follows easily from the graph in Figure 6.12. Here fg is the area of the rectangle, while f^p/p is the area under

the curve $y = x^{p-1}$ from 0 to f (shaded). Because $q - 1 = 1/(p - 1)$, the area under the curve $x = y^{q-1}$ from 0 to g , which is g^q/q , is the dotted area. Geometrically, the sum of the shaded area and the dotted area exceeds the area of the rectangle. \square

Another useful inequality is the Hölder inequality, which is a generalization of the Cauchy–Schwarz; we leave its proof, which follows from Young’s inequality, as an exercise.

Hölder’s Inequality. If p and q are positive real numbers satisfying $1/p + 1/q = 1$, then

$$\int_a^b |u(x)v(x)| dx \leq \left[\int_a^b |u(x)|^p dx \right]^{1/p} \left[\int_a^b |v(x)|^q dx \right]^{1/q},$$

provided that the two integrals on the right exist.

Finally, we need the following results, known as *Poincaré inequalities*. These inequalities relate integrals of functions to integrals of their derivatives when certain boundary conditions hold true. We state the inequalities on the interval $[0, 1]$, but they hold true on any interval with appropriate adjustment of the constants.

Poincaré Inequalities. Let $u = u(x)$ be twice continuously differentiable on $[0, 1]$.

1. If $u(0) = u(1) = 0$, then $\int_0^1 u'(x)^2 dx \geq \pi^2 \int_0^1 u(x)^2 dx$.
2. If $u'(0) = u'(1) = 0$, then $\int_0^1 u''(x)^2 dx \geq \pi^2 \int_0^1 u'(x)^2 dx$.

We prove (1) and leave (2) as an exercise. First, consider the eigenvalue problem

$$u'' + \lambda u = 0 \quad (0 < x < 1), \quad u(0) = u(1) = 0,$$

which has eigenvalues $\lambda_n = n^2\pi^2$ and corresponding orthonormal eigenfunctions $u_n(x) = \sqrt{2} \sin n\pi x$, $n = 1, 2, \dots$. Expanding u in terms of these eigenfunctions, we obtain

$$u(x) = \sum a_n u_n(x),$$

where the a_n are the Fourier coefficients

$$a_n = (1/\sqrt{2}) \int_0^1 u(x) \sin n\pi x dx.$$

Using integration by parts and the orthogonality property of the $u_n(x)$, we have

$$\begin{aligned} \int_0^1 u'(x)^2 dx &= - \int_0^1 u''(x)u(x) dx \\ &= - \int_0^1 \left(\sum a_n u_n'' \right) \left(\sum a_n u_n \right) dx \\ &= \int_0^1 \left(\sum a_n \lambda_n u_n \right) \left(\sum a_n u_n \right) dx \\ &= \sum a_n^2 \lambda_n \geq \pi^2 \sum a_n^2. \end{aligned}$$

But

$$\int_0^1 u(x)^2 dx = \int_0^1 \left(\sum a_n u_n \right) \left(\sum a_n u_n \right) dx = \sum a_n^2,$$

again using orthogonality. This completes the argument. \square

The Poincaré inequalities can be generalized to higher dimensions where they relate the norms (the L_2 -norm) of functions to the norms of their gradients (Smoller 1994, pp. 112 ff).

6.5.2 Energy Estimates

We are now in a position to obtain estimates of $E(t)$ and $Q(t)$ for various initial-boundary value problems. The four examples in this section illustrate the various techniques.

Example. Consider the semilinear reaction–diffusion equation

$$u_t - u_{xx} = f(u), \quad 0 < x < 1, \quad t > 0, \tag{6.5.1}$$

$$u(x, 0) = u_0(x), \quad 0 < x < 1, \tag{6.5.2}$$

$$u_x(0, t) = u_x(1, t) = 0, \quad t > 0, \tag{6.5.3}$$

where the reaction term f is assumed to be continuously differentiable and $\sup_{u \in \mathbb{R}} |f'(u)| = M < \infty$. Let

$$Q(t) = \int_0^1 u_x^2 dx.$$

Then, differentiating Q gives

$$\begin{aligned} Q'(t) &= 2 \int_0^1 u_x u_{xt} dx = 2u_x u_t|_0^1 - 2 \int_0^1 u_{xx} u_t dx \\ &= -2 \int_0^1 u_{xx}(u_{xx} + f(u)) dx \\ &= -2 \int_0^1 u_{xx}^2 dx + 2 \int_0^1 u_x^2 f'(u) dx, \end{aligned}$$

where we used integration by parts and applied (6.5.1) and (6.5.2). We use Poincaré's inequality on the first term on the right side of the last equation, and we bound the second integral by

$$\int_0^1 u_x^2 f'(u) dx \leq \int_0^1 u_x^2 |f'(u)| dx \leq M \int_0^1 u_x^2 dx.$$

Therefore

$$Q'(t) \leq -2\pi^2 \int_0^1 u_x^2 dx + 2M \int_0^1 u_x^2 dx = 2(M - \pi^2)Q(t) = \mu Q(t),$$

where $\mu \equiv 2(M - \pi^2)$. We obtained a differential inequality for the quantity $Q(t)$ of the form

$$Q'(t) - \mu Q(t) \leq 0.$$

Multiplying by the integrating factor $e^{-\mu t}$ gives

$$(Qe^{-\mu t})' \leq 0,$$

and then integrating from 0 to t yields

$$Q(t) \leq Q(0)e^{\mu t}.$$

So, we obtained a bound on $Q(t)$. If $\mu < 0$ (i.e., $M < \pi^2$), then $Q(t)$ decays exponentially as t tends to infinity. So if the nonlinear problem (6.5.1)–(6.5.3) has a solution $u = u(x, t)$, and if f is a smooth function with the property that $\sup|f'(u)| = M < \pi^2$, then

$$\lim_{t \rightarrow \infty} \int_0^1 u_x^2(x, t) dx = 0.$$

Biologically, for example, this result precludes spatial pattern formation (e.g., spatial striations) in reactive-diffusive systems governed by (6.5.1)–(6.5.3). \square

We now consider a reaction-advection-diffusion equation where inequalities are used in an ingenious way to show that if the initial wave profile is square-integrable, then the wave profile for any $t > 0$ is square integrable. Another

way of stating this is to say that the L_2 norm of the solution remains bounded, or stays under control, as t increases.

Example. Consider the initial-boundary value problem

$$u_t - Du_{xx} + uu_x = f(u), \quad 0 < x < 1, \quad t > 0, \quad (6.5.4)$$

$$u(x, 0) = u_0(x), \quad 0 < x < 1, \quad (6.5.5)$$

$$u(0, t) = u(1, t) = 0, \quad t > 0. \quad (6.5.6)$$

Assume that $u_0(0) = u_0(1) = 0$ and u_0 is continuous and nonnegative on $[0, 1]$; further assume that the reaction term f is a bounded continuous function on \mathbb{R} , $M = \sup |f|$, and $D > 0$. We prove that $E(t) = \int_0^1 u(x, t)^2 dx$ stays bounded for all t and, in fact, we calculate an upper estimate. Using (6.5.4) and (6.5.6), we have

$$\begin{aligned} E'(t) &= 2 \int_0^1 uu_t dx \\ &= 2 \int_0^1 u(Du_{xx} - uu_x + f(u)) dx \\ &= 2D \int_0^1 uu_{xx} dx - 2 \int_0^1 \left(\frac{u^3}{3}\right)_x dx + 2 \int_0^1 uf(u) dx. \end{aligned}$$

The second integral is zero by the boundary conditions on u at $x = 0$ and $x = 1$; the first integral can be integrated by parts to obtain

$$\int_0^1 uu_{xx} dx = uu_x|_0^1 - \int_0^1 u_x^2 dx.$$

The third integral can be bounded by the following sequence of inequalities:

$$\begin{aligned} \int_0^1 uf(u) dx &= \left(\int_0^1 u^2 dx \right)^{1/2} \left[\int_0^1 f(u)^2 dx \right]^{1/2} && \text{(Hölder's inequality)} \\ &\leq M \left(\int_0^1 u^2 dx \right)^{1/2} \\ &\leq \frac{M}{\pi} \left(\int_0^1 u_x^2 dx \right)^{1/2} && \text{(Poincaré's inequality)} \\ &= \frac{MD^{-1/2}}{\pi} \left(D \int_0^1 u_x^2 dx \right)^{1/2} \\ &\leq \frac{1}{2} \left(\frac{MD^{-1/2}}{\pi} \right)^2 + \frac{D}{2} \int_0^1 u_x^2 dx. && \text{(Young's inequality)} \end{aligned}$$

Putting all the inequalities together yields

$$\begin{aligned} E'(t) &\leq -2D \int_0^1 u_x^2 dx + \frac{M^2 D^{-1}}{\pi^2} + D \int_0^1 u_x^2 dx \\ &= -D \int_0^1 u_x^2 dx + \frac{M^2 D^{-1}}{\pi^2} \\ &\leq -D\pi^2 \int_0^1 u^2 dx + \frac{M^2 D^{-1}}{\pi^2}, \end{aligned}$$

where, in the last step, Poincaré's inequality is reapplied. Therefore, we derived the differential inequality

$$E'(t) + D\pi^2 E(t) \leq \frac{M^2}{D\pi^2}.$$

Multiplying by $e^{D\pi^2 t}$ and then integrating from 0 to t yields

$$E(t) \leq E(0)e^{-\pi^2 Dt} + \frac{M^2}{D^2\pi^4} \left[1 - e^{-\pi^2 Dt} \right].$$

Therefore, the energy remains bounded for all $t > 0$, and an upper bound on the energy has been obtained. \square

Another important problem is to determine whether solutions blow up in finite time. Again, an energy-type argument is relevant.

Example. (Blowup) Consider the reactive-diffusive system

$$u_t = u_{xx} + u^3, \quad 0 < x < \pi, \quad t > 0, \quad (6.5.7)$$

$$u(0, t) = u(\pi, t) = 0, \quad t > 0, \quad (6.5.8)$$

$$u(x, 0) = u_0(x), \quad 0 < x < \pi, \quad (6.5.9)$$

where u_0 is continuous and nonnegative on $[0, \pi]$. We demonstrate that if

$$\int_0^\pi u_0(x) \sin x dx > 2, \quad (6.5.10)$$

then the solution blows up in finite time. To this end we observe that the maximum principle implies that $u \geq 0$, so long as the solution exists, and we define

$$s(t) \equiv \int_0^\pi u(x, t) \sin x dx.$$

Integration by parts gives

$$s'(t) = \int_0^\pi u_t \sin x dx = \int_0^\pi (u_{xx} \sin x + u^3 \sin x) dx = -s(t) + \int_0^\pi u^3 \sin x dx.$$

Now apply Hölder's inequality with $p = 3$ and $q = \frac{3}{2}$ to get

$$\begin{aligned} s(t) &= \int_0^\pi u \sin x \, dx = \int_0^\pi \sin^{2/3} x u \sin^{1/3} x \, dx \\ &\leq \left(\int_0^\pi (\sin^{2/3} x)^{3/2} \, dx \right)^{2/3} \left(\int_0^\pi (u \sin^{1/3} x)^3 \, dx \right)^{1/3} \\ &\leq 2^{2/3} \left(\int_0^\pi u^3 \sin x \, dx \right)^{1/3}. \end{aligned}$$

Therefore

$$s(t)^3 \leq 4 \int_0^\pi u^3 \sin x \, dx,$$

and hence

$$s'(t) \geq -s(t) + \frac{s(t)^3}{4}, \quad t > 0; \quad s(0) > 2. \quad (6.5.11)$$

We now show that the inequality (6.5.11) implies that $s(t) \rightarrow +\infty$ at a finite t , and thus the solution blows up in finite time. Expression (6.5.11) reminds us of a Bernoulli equation in ordinary differential equations, except that it is an inequality. But the same technique applies. Specifically, let $v = 1/s^2$. Then (6.5.11) becomes a linear inequality in $v(t)$ given by

$$v'(t) \leq 2v(t) - \frac{1}{2}.$$

Multiplying through by the integrating factor e^{-2t} and integrating from 0 to t yields

$$v(t) \leq \frac{1 - e^{-2t}}{4} + v(0)e^{2t}.$$

Consequently

$$s(t) \geq \left[\left(e^{2t} \left(\frac{1}{s(0)^2} - \frac{1}{4} \right) + \frac{1}{2} \right)^{-2} \right]. \quad (6.5.12)$$

But because $s(0) > 2$, the right side of (6.5.12) goes to infinity at a finite value of t , showing that $s(t)$ blows up at finite time. \square

Example. (Uniqueness) In this example we use energy estimates to show that an initial-boundary value problem associated with Burgers' equation can have at most one solution. Consider the problem

$$u_t + uu_x = Du_{xx}, \quad 0 < x < 1, \quad t > 0, \quad (6.5.13)$$

$$u(x, 0) = u_0(x), \quad 0 < x < 1, \quad (6.5.14)$$

$$u(0, t) = u(1, t) = 0, \quad 0 < x < 1. \quad (6.5.15)$$

In the usual manner assume that both u and v are solutions to (6.5.13)–(6.5.15); then $w = u - v$ satisfies the equation

$$w_t + \left(\frac{aw}{2}\right)_x = Dw_{xx}, \quad a = a(u, v) \equiv u + v, \quad (6.5.16)$$

with $w = 0$ along $x = 0, x = 1$, and $t = 0$. Take $E(t) = \frac{1}{2} \int_0^1 w(x, t)^2 dx$. Then

$$E'(t) = \int_0^1 ww_t dx = \int_0^1 Dw w_{xx} dx - \frac{1}{2} \int_0^1 (aw)_x w dx. \quad (6.5.17)$$

The first term on the right can be integrated by parts to obtain

$$\int_0^1 Dw w_{xx} dx = -D \int_0^1 w_x^2 dx.$$

The second integral on the right of (6.5.17) can be written

$$\int_0^1 (aw)_x w dx = \int_0^1 a_x w^2 dx + \int_0^1 aw w_x dx = \int_0^1 a_x w^2 dx - \int_0^1 (aw)_x w dx,$$

where integration by parts has been applied yet again. Thus

$$\int_0^1 (aw)_x w dx = \frac{1}{2} \int_0^1 a_x w^2 dx,$$

and (6.5.17) becomes

$$\begin{aligned} E'(t) &= -D \int_0^1 w_x^2 dx - \frac{1}{4} \int_0^1 a_x w^2 dx, \\ &\leq -\frac{1}{4} \int_0^1 a_x w^2 dx \leq \frac{\max |a_x|}{4} \int_0^1 w^2 dx. \end{aligned}$$

We have shown that $E'(t) \leq cE(t)$ for all $t > 0$, where c is a positive constant. As in previous examples we multiply by e^{-ct} and integrate from 0 to t to obtain the inequality $E(t) \leq E(0)e^{ct}$. But $E(0) = 0$, and therefore $E(t) = 0$ for all $t > 0$, showing that w is identically zero. Thus $u = v$ and uniqueness is established. (Uniqueness can also be proved by using the Cole–Hopf transformation to transform Burgers' equation to the diffusion equation.) \square

6.5.3 Invariant Sets

Another method used to predict the long time behavior of solutions of certain classes of reaction–diffusion equations is the method of invariant sets. To introduce the concept we first consider the plane autonomous system of ordinary differential equations

$$\frac{du}{dt} = f(u, v) \quad \frac{dv}{dt} = g(u, v). \quad (6.5.18)$$

A domain $\Sigma \subset \mathbb{R}^2$ enclosed by a smooth, simple closed curve $\partial\Sigma$ is an *invariant set* for (6.5.18) if any solution $(u(t), v(t))$ with $(u(0), v(0))$ in Σ remains in Σ for all $t > 0$. It is clear that if the vector field $\mathbf{f} = (f, g)$ along $\partial\Sigma$ points into the region Σ , which is implied by the condition $\mathbf{f} \cdot \mathbf{n} < 0$ on $\partial\Sigma$, where \mathbf{n} is the outward unit normal, then Σ is an invariant set. Having information about the existence of bounded, invariant sets allows us to extract information about the longtime behavior of solutions of (6.5.18); for example, the Poincaré–Bendixson theorem states that if $(u(t), v(t))$ remains bounded as t becomes infinite, then the orbit tends to a critical point, is a periodic solution, or approaches a periodic solution.

This concept extends in a natural way to systems of reaction–diffusion equations. The idea is to find a closed bounded region in the space of dependent variables that traps the solution for all $t > 0$, or at least up until the time the solution ceases to exist. If such a region can be found, then one automatically obtains a priori bounds on the solution (recall that a priori bounds are often required to obtain global existence of solutions). Even if the region is unbounded, useful information can often be extracted.

To fix the concept, we consider a system of two reaction–diffusion equations of the form

$$u_t = d_1 u_{xx} + f(u, v) \quad v_t = d_2 v_{xx} + g(u, v), \quad x \in I, \quad t > 0, \quad (6.5.19)$$

where $\mathbf{u} = (u, v)$ is the unknown solution vector, $\mathbf{f} = (f, g)$ is the vector field of nonlinear reaction terms, d_1 and d_2 are nonnegative diffusion constants, and I is an interval in \mathbb{R} , possibly all of \mathbb{R} . The functions f and g are continuous, and initial conditions are given by

$$u(x, 0) = u_0(x), \quad v(x, 0) = v_0(x), \quad x \in I. \quad (6.5.20)$$

If I is not all of \mathbb{R} , then we assume Dirichlet or Neumann boundary conditions at the ends of the interval.

Definition. Let Σ be a closed set in \mathbb{R}^2 . If $\mathbf{u}(x, t)$ is a solution to (6.5.19)–(6.5.20) for $0 \leq t < \delta \leq \infty$, with given boundary conditions, and the initial

values and boundary values are in Σ , and $\mathbf{u}(x, t)$ is in Σ for all $x \in I$ and $0 < t < \delta$, then Σ is called an *invariant set* for the solution $\mathbf{u}(x, t)$.

The next theorem asserts that if the vector field \mathbf{f} points inward along the boundary of a rectangle, then the rectangle must be an invariant set.

Theorem. Let $\Sigma = [a, b] \times [c, d]$ be a rectangle in uv space, and let Σ^0 denote its interior and $\partial\Sigma$ denote the boundary, with \mathbf{n} the outward unit normal. If

$$\mathbf{f}(\mathbf{u}) \cdot \mathbf{n} < 0 \quad \text{on } \partial\Sigma, \quad (6.5.21)$$

then Σ is an invariant set for (6.5.19)–(6.5.20).

Proof: We proceed by contradiction and assume that Σ is not an invariant set. Then, without loss of generality, we may suppose that $u(x_0, t_0) = b$ for some (x_0, t_0) with $u(x, t) < b$ for all $x \in I$, $0 < t < t_0$, yet

$$u_t(x_0, t_0) \geq 0. \quad (6.5.22)$$

Then the function $u(x, t_0)$, regarded as a function of x , must have a maximum at $x = x_0$. [For example, let $h(x) = u(x, t_0)$, so that $h(x_0) = b$. If $h'(x_0) > 0$, then $h(x_1) > b$ for some $x_1 > x_0$ sufficiently close to x_0 . Then $u(x_1, t_0) > b$, and because u is continuous, $u(x, t) > b$ in some neighborhood of (x_1, t_0) . In particular, $u(x, t) > b$ for some $x \in I$ and some $t < t_0$, which is contrary to our assumption. Similarly, we cannot have $h'(x_0) < 0$.] Therefore $u_{xx}(x_0, t_0) \leq 0$. Moreover, at (x_0, t_0) , we have

$$u_t = d_1 u_{xx} + f(u, v) \leq f(u, v) = \mathbf{f}(\mathbf{u}) \cdot \mathbf{n} < 0. \quad (6.5.23)$$

The latter statement following from the fact the $\mathbf{n} = (1, 0)$ on the boundary $u = b$. But (6.5.23) contradicts (6.5.22), and therefore Σ must be an invariant set, completing the proof. \square

Example. The FitzHugh–Nagumo equations are

$$u_t = u_{xx} + u(1 - u)(u - a) - v \quad v_t = \alpha v_{xx} + \sigma u - \gamma v,$$

where $0 < a < 1$, $\alpha \geq 0$, and $\sigma, \gamma > 0$. These equations model conduction along nerve fibers. Here

$$f(u, v) = u(1 - u)(u - a) - v, \quad g(u, v) = \sigma u - \gamma v.$$

are the reaction terms. Let Σ be the rectangle shown in Figure 6.13. It is easy to see that the vector field $\mathbf{f} = (f, g)$ points inward along the boundaries of the rectangle; thus (6.5.21) holds and Σ is an invariant set. \square

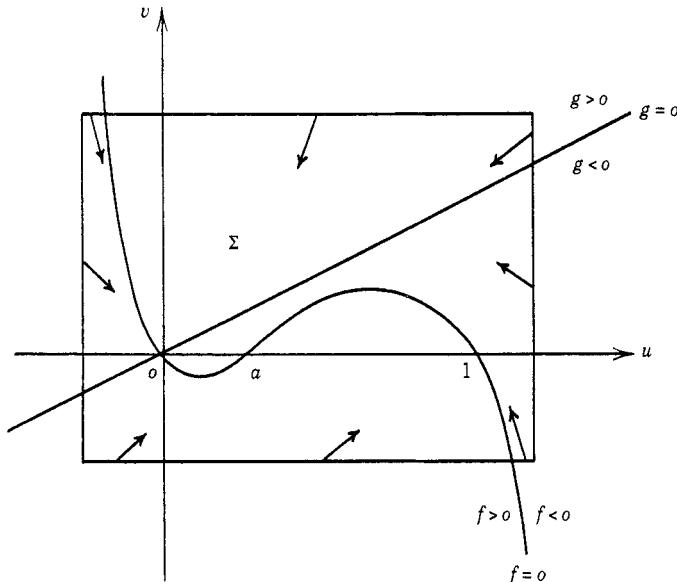


Figure 6.13 Invariant rectangle for the FitzHugh–Nagumo equations.

Example. The last theorem can be generalized immediately to an n -dimensional system

$$\mathbf{u}_t = D(\mathbf{u})\mathbf{u}_{xx} + C(\mathbf{u})\mathbf{u}_x + \mathbf{f}(\mathbf{u}, t), \quad x \in I, \quad t > 0, \quad (6.5.24)$$

where D and C are $n \times n$ diagonal matrices with $D \geq 0$ (D is positive semi-definite), \mathbf{f} is continuously differentiable, and $\mathbf{u} = (u_1, \dots, u_n)$ is the unknown solution vector. Then any region (a box in \mathbb{R}^n) of the form

$$\Sigma = [a_1, b_1] \times \cdots \times [a_n, b_n]$$

is an invariant set for (6.5.24) with appropriate initial and boundary conditions, provided that

$$\mathbf{f}(\mathbf{u}, t) \cdot \mathbf{n} < 0 \quad \text{for all } \mathbf{u} \in \partial\Sigma, \quad t > 0,$$

where \mathbf{n} is the outward unit normal to Σ . The proof is left as an exercise. \square

This last result can be generalized under certain conditions to permit invariant sets to have the form of a convex subset of \mathbb{R}^n . We now formulate this general result and refer the reader to Smoller (1994) for a simple proof.

Consider the reaction–diffusion system (6.5.24) where $D \geq 0$ and C are continuous matrices, not necessarily diagonal, and \mathbf{f} is continuously differentiable.

Let $G_i, i = 1, \dots, p$, be continuously differentiable, real-valued functions on \mathbb{R}^n and define the regions $E_i = \{\mathbf{u} \in \mathbb{R}^n : G_i(\mathbf{u}) \leq 0\}$ for $i = 1, \dots, p$. Assume that $\mathbf{grad} G_i$ is never zero. Recall that the vector $\mathbf{grad} G_i$ is normal to the locus $G_i(\mathbf{u}) = 0$, which is a level curve of the surface $z = G_i(\mathbf{u})$ in \mathbb{R}^{n+1} . Now define

$$\Sigma = \cap_{i=1}^p E_i,$$

so that Σ is the intersection of half-spaces. Clearly Σ is closed. To show that Σ is an invariant set for (6.5.24), three conditions hold at every boundary point of Σ :

1. The gradient of the G_i that defines the particular section of the boundary under consideration must be a left eigenvector of the diffusion matrix D and the advection matrix C .
2. The vector field \mathbf{f} along the boundary must point into the region.
3. On every section of the boundary the function G_i defining that portion of the boundary must be a convex function; this implies, of course, that the region Σ is convex.

To formulate these conditions mathematically, we say that a function G is *quasiconvex* at a point \mathbf{v} in \mathbb{R}^n if $\mathbf{grad} G(\mathbf{v}) \cdot \mathbf{h} = 0$ implies that $\mathbf{h}^T Q(\mathbf{v}) \mathbf{h} \geq 0$, where Q is the matrix of second partial derivatives of G ; that is, $Q = (\partial^2 G / \partial u_i \partial u_j)$. Then one can prove the following: Σ is an invariant set for (6.5.24) provided the following three conditions hold: for all $t > 0$ and for every $\mathbf{u}^* \in \partial\Sigma$ (so that $G_i(\mathbf{u}^*) = 0$ for some i):

1. $\mathbf{grad} G_i(\mathbf{u}^*)$ is a left eigenvector of $D(\mathbf{u}^*)$ and of $C(\mathbf{u}^*)$.
2. $\mathbf{grad} G_i(\mathbf{u}^*) \cdot \mathbf{f}(\mathbf{u}^*, t) < 0$.
3. If $\mathbf{grad} G_i(\mathbf{u}^*) D(\mathbf{u}^*) = \mu \mathbf{grad} G_i(\mathbf{u}^*)$ and $\mu \neq 0$, then G_i is quasiconvex at \mathbf{u}^* .

EXERCISES

1. Prove the Poincaré inequality (2).
2. Consider the initial value problem for the diffusion equation

$$\begin{aligned} u_t - Du_{xx} &= 0, \quad x \in \mathbb{R}, \quad t > 0, \\ u(x, 0) &= u_0(x), \quad x \in \mathbb{R}, \end{aligned}$$

where u_0 is bounded and continuous, and $\int_{\mathbb{R}} u_0^2(x) dx$ is finite. Prove that there exists a constant C such that $|u(x, t)| \leq C/t^{1/4}$ for all $t > 0$.

3. Consider the problem

$$\begin{aligned} u_t - u_{xx} &= f(u), \quad 0 < x < 1, \quad t > 0, \\ u(x, 0) &= u_0(x), \quad 0 < x < 1, \\ u(0, t) &= u(1, t) = 0, \quad t > 0, \end{aligned}$$

where f is continuously differentiable and $\sup |f'(u)| < \pi^2$. Prove that if a unique solution $u = u(x, t)$ exists, then

$$\lim_{t \rightarrow \infty} \int_0^1 u^2(x, t) dx = 0.$$

4. Consider the reactive-diffusive system

$$\begin{aligned} u_t - D_1 u_{xx} &= f(u, v), \quad v_t - D_2 v_{xx} = g(u, v), \quad 0 < x < 1, \quad t > 0, \\ u(x, 0) &= u_0(x), \quad v(x, 0) = v_0(x), \quad 0 < x < 1, \\ u_x &= v_x \quad \text{at } x = 0 \quad \text{and} \quad x = 1 \quad \text{for all } t > 0, \end{aligned}$$

where f and g are continuously differentiable and f_u, f_v, g_u, g_v are bounded. Let

$$D = \min(D_1, D_2), \quad M = \sup\{(f_u^2 + f_v^2 + g_u^2 + g_v^2)^{1/2} : u, v \in \mathbb{R}\}.$$

Show that if $4M - 2\pi^2 D < 0$, then $\lim_{t \rightarrow \infty} q(t) = 0$, where

$$q(t) = \frac{1}{2} \int_0^1 (u_x^2 + v_x^2) dx.$$

5. (a) Show that the boundary value problem

$$-u'' = \delta e^u, \quad 0 < x < 1, \quad u(0) = u(1) = 0,$$

has no solution for $\delta > \pi^2/e$. Hint: Show that $\int_0^1 \phi_1(\delta e^u - \lambda_1 u) dx = 0$, where λ_1 is the smallest eigenvalue of the linear problem $-\phi'' = \lambda \phi, \phi(0) = \phi(1) = 0$, and ϕ_1 is a corresponding eigenfunction.

(b) (*Gelfand Problem*) Show that the initial-boundary value problem

$$\begin{aligned} u_t - u_{xx} &= \delta e^u, \quad 0 < x < 1, \quad t > 0, \\ u(x, 0) &= 0, \quad 0 < x < 1; \quad u(0, t) = u(1, t) = 0, \quad t > 0, \end{aligned}$$

does not have a global solution if $\delta > \pi^2/e$. Hint: Show that

$$E'(t) + \lambda_1 E(t) = \delta \int_0^1 \phi_1 e^u dx \geq \delta e^{E(t)},$$

where $E(t) = \int_0^1 \phi_1(x) u(x, t) dx$.

6. Consider the initial value problem for the Schrödinger equation

$$u_t = iu_{xx}, \quad x \in \mathbb{R}, \quad t > 0; \quad u(x, 0) = u_0(x), \quad x \in \mathbb{R},$$

and assume that a solution $u = u(x, t)$ exists for all $t > 0$ and $x \in \mathbb{R}$ with u and all of its derivatives vanishing at $|x| = \infty$. Prove that $E(t) = \int_{\mathbb{R}} u \bar{u} dx$ is constant for $t > 0$.

7. Consider the system

$$\begin{aligned} \mathbf{u}_t &= A\mathbf{u}_x, \quad 0 < x < 1, \quad t > 0, \\ \mathbf{u}(0, t) &= \mathbf{u}(1, t) = \mathbf{0}, \quad t > 0; \quad \mathbf{u}(x, 0) = \mathbf{u}_0(x), \quad 0 < x < 1, \end{aligned}$$

where $\mathbf{u} = \mathbf{u}(x, t)$ is an n -vector and A is a constant matrix with $A^T = A$.

Prove that $E'(t) = 0$ where $E(t) = \int_0^1 \mathbf{u}^T \mathbf{u} dx$.

8. Consider the problem

$$\begin{aligned} u_t - u_{xx} &= f(u), \quad 0 < x < \pi, \quad t > 0, \\ u(0, t) &= u(\pi, t) = 0, \quad t > 0; \quad u(x, 0) = u_0(x), \quad 0 < x < \pi, \end{aligned}$$

where u_0 is nonnegative and continuous, and

$$f(u) > 0, \quad f'(u) > 0, \quad f''(u) > 0 \quad \text{for } u > 0,$$

and $\int_0^\infty du/f(u)$ is finite. If $\int_0^\pi u_0(x) dx$ is sufficiently large, prove that the solution blows up in finite time. Hint: Consider $s(t) = \int_0^\pi u(x, t) \sin x dx$ and utilize the fact that $\int_0^\pi f(u) \sin x dx > f(\int_0^\pi u \sin x dx)$ to get $s'(t) \geq -\lambda s(t) + f(s(t))$ for some $\lambda > 0$.

9. Assume that $u \in C^\infty(\mathbb{R})$ and that $u(x+1) = u(x)$ for all x , and denote $\|u\|^2 = \int_0^1 |u(x)|^2 dx$. Prove for all $c > 0$ and all integers $j, k \geq 1$ that

$$\|u^{(j)}\|^2 \leq c \|u^{(j+k)}\|^2 + c^{-j/k} \|u\|^2.$$

This is an example of a *Sobolev-type* inequality.

10. (*Navier–Stokes Equations*) The Navier–Stokes equations, which govern a Newtonian viscous fluid, are

$$\mathbf{u}_t - \nu \Delta \mathbf{u} + (\mathbf{u} \cdot \mathbf{grad}) \mathbf{u} + \mathbf{grad} p = \mathbf{f}(\mathbf{x}), \quad \operatorname{div} \mathbf{u} = 0,$$

where $\mathbf{x} = (x, y, z)$, and where $\mathbf{u} = \mathbf{u}(\mathbf{x}, t)$ is the velocity vector, $p = p(\mathbf{x}, t)$ the pressure, \mathbf{f} is the body force, $\nu > 0$ the viscosity, and Δ is the three-dimensional Laplacian. Let Ω be an open, bounded region in space with smooth boundary $\partial\Omega$, and assume that the Navier–Stokes equations hold in Ω with $\mathbf{u} = \mathbf{0}$ on $\partial\Omega$ for all $t > 0$, and $\mathbf{u}(\mathbf{x}, 0) = \mathbf{u}_0(\mathbf{x})$ on Ω . Prove that

$$\|\mathbf{u}(t)\|^2 \leq \|\mathbf{u}(0)\|^2 e^{-a\nu t} + \frac{C}{\nu} (1 - e^{-a\nu t}) \quad (C, a > 0),$$

for all $t > 0$, where $\|\mathbf{u}(t)\|$ denotes the L_2 norm

$$\|\mathbf{u}(t)\| = \left[\int_{\Omega} \mathbf{u}(\mathbf{x}, t)^T \mathbf{u}(\mathbf{x}, t) d\mathbf{x} \right]^{1/2} d\mathbf{x} = dx.$$

Thus the energy remains bounded for all time. *Hint:* Integrate over Ω , then use integration by parts followed by three-dimensional versions of Young's and Poincaré's inequalities.

11. The flow of an ideal, incompressible fluid of constant density in an open bounded domain Ω in \mathbb{R}^3 with a smooth boundary is governed by the momentum law and the incompressibility condition

$$\mathbf{u}_t + (\mathbf{u} \cdot \mathbf{grad})\mathbf{u} + \mathbf{grad} p = \mathbf{0}, \quad \operatorname{div} \mathbf{u} = 0, \quad x \in \Omega, \quad t > 0,$$

where \mathbf{u} is the velocity and p is the pressure. Assume that the flow is parallel to the boundary [i.e., $\mathbf{u} \cdot \mathbf{n} = 0$ on $\partial\Omega$], and initially $\mathbf{u}(\mathbf{x}, 0) = \mathbf{u}_0(\mathbf{x})$, $\mathbf{x} \in \Omega$.

Prove that

$$\|\mathbf{u}(t)\| = \|\mathbf{u}_0\|, \quad t > 0.$$

(The notation is the same as in Exercise 10.)

12. Consider the reactive-diffusive system

$$u_t = u_{xx} + (1 - u^2 - v^2)u, \quad v_t = v_{xx} + (1 - u^2 - v^2)v.$$

Show that the unit circle in uv space is an invariant set. *Hint:* Consider the set $\Sigma_\varepsilon = \{(u, v) : u^2 + v^2 \leq 1 + \varepsilon\}$ for $\varepsilon > 0$.

13. In combustion theory the solid-fuel model is

$$y_t = D y_{xx} - y r(T), \quad T_t = k T_{xx} + q y r(T),$$

where y is the mass fraction of the reactant, T is the temperature, and D , k , and q are positive constants. The reaction rate is $r(T) = \exp(-E/RT)$, where E and R are positive constants. Show that $\Sigma = \{(y, T) : 0 \leq y \leq 1, T \geq a > 0\}$ is an invariant set. *Hint:* First consider y in the range $-\varepsilon \leq y \leq 1$.

14. Consider the reaction-diffusion problem

$$u_t = u_{xx} - u + u^p, \quad 0 < x < \pi, \quad t > 0,$$

$$u(x, 0) = u_0(x) > 0, \quad 0 < x < \pi; \quad u(0, t) = u(\pi, t) = 0, \quad t > 0,$$

where $p > 1$. If

$$\int_0^\pi u_0(x) \sin x dx > 2^{p/(p-1)},$$

show that the solution blows up in finite time.

6.6 Pattern Formation

Morphogenesis describes how patterns and forms develop in animals and plants. The complex processes involved in morphogenesis are not completely understood. However, in a seminal paper Alan Turing (1952) developed the idea that chemical species present during the development process can react and diffuse in such a way to create patterns. The patterns are instabilities created by the presence of a diffusion mechanism, which is contrary to our established intuition of regarding diffusion as a stabilizing, smoothing process.

We begin by considering the general system of reaction–diffusion equations on the interval $0 < x < L$ given by

$$u_t = \alpha u_{xx} + f(u, v), \quad v_t = \beta v_{xx} + g(u, v), \quad (6.6.1)$$

with no-flux boundary conditions

$$u_x = v_x = 0 \quad \text{at } x = 0, L.$$

This system is called a *Turing system*. We are interested in the stability of a uniform steady state. We show that with special values of the parameters it is possible that small perturbations from the uniform state grow and produce patterns, or density variations, in spite of the presence of diffusion in the problem.

Suppose that $u = \bar{u}$, $v = \bar{v}$ is a constant equilibrium solution that must satisfy

$$f(\bar{u}, \bar{v}) = 0, \quad g(\bar{u}, \bar{v}) = 0.$$

Let U and V denote small perturbations from equilibrium and take

$$u = \bar{u} + U(x, t), \quad v = \bar{v} + V(x, t). \quad (6.6.2)$$

Substituting (6.6.2) into (6.6.1) and the boundary conditions, using Taylor's expansion on f and g , and then deleting the nonlinear terms, gives linearized perturbation equations for U and V that satisfy no-flux boundary conditions:

$$U_t = \alpha U_{xx} + f_u(\bar{u}, \bar{v})U + f_v(\bar{u}, \bar{v})V, \quad (6.6.3)$$

$$V_t = \beta V_{xx} + g_u(\bar{u}, \bar{v})U + g_v(\bar{u}, \bar{v})V. \quad (6.6.4)$$

In matrix notation, with

$$\mathbf{W} = \begin{pmatrix} U \\ V \end{pmatrix}, \quad D = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}, \quad J = \begin{pmatrix} f_u(\bar{u}, \bar{v}) & f_v(\bar{u}, \bar{v}) \\ g_u(\bar{u}, \bar{v}) & g_v(\bar{u}, \bar{v}) \end{pmatrix},$$

the perturbation equations (6.6.3)–(6.6.4) can be written as

$$\mathbf{W}_t = D\mathbf{W}_{xx} + J\mathbf{W}. \quad (6.6.5)$$

The boundary conditions are

$$\mathbf{W}_x = \mathbf{0} \text{ at } x = 0, L. \quad (6.6.6)$$

The system (6.6.5)–(6.6.6) is linear and can be solved by separation of variables. To this end assume

$$\mathbf{W}(x, t) = \phi(t)\mathbf{Y}(x).$$

Substituting into the system and separating the spatial and temporal parts gives

$$\phi'(t) = \sigma\phi(t), \quad (\text{giving } \phi(t) = e^{\sigma t})$$

and

$$D\mathbf{Y}'' + (J - \sigma I)\mathbf{Y} = \mathbf{0}, \quad \mathbf{Y}'(0) = \mathbf{Y}'(L) = \mathbf{0}, \quad (6.6.7)$$

where σ is the separation constant, which is the growth rate of the perturbations. The boundary value problem has solutions of the form

$$\mathbf{Y} = \mathbf{Y}_n = \mathbf{C}_n \cos \frac{n\pi x}{L}, \quad n = 0, 1, 2, \dots,$$

where \mathbf{C} is constant. Substituting into (6.6.7) gives

$$\left[J - \sigma I - \frac{n^2\pi^2}{L^2} D \right] \mathbf{C} = \mathbf{0}.$$

This system of algebraic equations will have a nontrivial solution when

$$\det \left(J - \sigma I - \frac{n^2\pi^2}{L^2} D \right) = 0. \quad (6.6.8)$$

This equation is a condition on the growth factor σ that must hold if there are nontrivial solutions. When expanded, (6.6.8) is a quadratic equation for the complex roots $\sigma = \sigma_n$, the growth rate of the n th mode. The roots σ_n depend on the diffusion constants α, β , the equilibrium solution \bar{u}, \bar{v} , the size of the medium L , and the wavelength $2L/n$ of the perturbation.

Modal solutions of the boundary values problem for the perturbations are given by

$$\mathbf{W}_n = \mathbf{C}_n e^{\sigma_n t} \cos \frac{n\pi x}{L}.$$

Therefore the general solution of (6.6.5)–(6.6.6), which is the general form of how perturbations evolve, is the sum of all the Fourier modes, or

$$\mathbf{W}(x, t) = \sum \mathbf{C}_n e^{\sigma_n t} \cos \frac{n\pi x}{L},$$

where the constants are determined by the initial perturbations. If one can find values of the parameters that make one of the roots σ_n positive or have a positive real part, then there is an unstable mode. If σ_n is negative or has a

negative real part for every n , then all modes decay and the uniform state is stable. In this latter case, patterns cannot form.

Example. (Chemotaxis) We now apply these ideas to a problem in cell aggregation, first introduced in Section 6.1. A *slime mold population* is a collection of unicellular amoeboid cells that feed on bacteria in the soil. When the food supply is plentiful, the bacteria are generally uniformly spatially distributed throughout the soil; but as the food supply becomes depleted and starvation begins, the amoeba start to secrete a chemical (cAMP) that acts as an attractant to the other amoeba and aggregation sites form. The rest of the story is even more interesting as the aggregation sites evolve into slugs that ultimately develop into a sporangiophores consisting of a stalk and head containing new spores. The spores are released, and the process begins anew. We are interested here in only the first part of this complicated problem, the onset of aggregation. We work in one spatial dimension.

We derived this model of chemotaxis in Section 6.1. To review, let $a = a(x, t)$ and $c = c(x, t)$ denote the density and concentration of the cellular amoeba and cAMP, respectively. The fundamental conservation laws are

$$a_t = \mu a_{xx} - \nu(ac_x)_x, \quad c_t = \delta c_{xx} + fa - kc. \quad (6.6.9)$$

Both a and c satisfy no-flux boundary conditions, specifically

$$a_x = c_x = 0, \quad x = 0, L,$$

which means that there is no escape from the medium. Notice that equations (6.6.9) do not quite form a Turing system; however, the analysis is the same.

There is a constant equilibrium solution $a = \bar{a}$, $c = \bar{c}$ provided

$$f\bar{a} = k\bar{c}.$$

In other words, the production of the chemical equals its degradation. This equilibrium state represents the spatially uniform state in the soil before aggregation begins. To determine the local stability of this state, we let

$$a = \bar{a} + A(x, t), \quad c = \bar{c} + C(x, t),$$

where A and C are small perturbations. Substituting these quantities into the system gives, after simplification, the perturbation equations

$$A_t = \mu A_{xx} - \nu((\bar{a} + A)C_x)_x, \quad C_t = \delta C_{xx} + fA - kC.$$

These equations are nonlinear because of the AC_x term in the amoeba equation. If we discard the nonlinear term on the assumption that the product of small terms is even smaller, then we obtain the linearized perturbation equations

$$A_t = \mu A_{xx} - \nu\bar{a}C_{xx}, \quad C_t = \delta C_{xx} + fA - kC. \quad (6.6.10)$$

Easily one sees that the perturbations satisfy no-flux boundary conditions.

Motivated by the previous discussion of the Turing system, we assume modal solutions of the form

$$A(x, t) = c_1 e^{\sigma t} \cos rx, \quad C(x, t) = c_2 e^{\sigma t} \cos rx, \quad (6.6.11)$$

where

$$r = \frac{n\pi}{L}, \quad n = 0, 1, 2, 3, \dots$$

The growth factor σ is to be determined, and c_1 and c_2 are constants. Notice the form of the solutions (6.6.11). The spatial part is bounded and periodic with frequency r and period $2\pi/r$, and the temporal part is exponential with growth factor σ , which may be a real or complex number. If σ is negative or has a negative real part, then the perturbation will decay and the equilibrium state will return (stable); if σ is positive or has positive real part, then the perturbations will grow and the equilibrium will be unstable. To obtain a condition on σ , we substitute (6.6.3) into (6.6.2) to obtain

$$(\sigma + \mu r^2)c_1 - \nu \bar{a}r^2c_2 = 0, \quad -fc_1 + (\sigma + k + \delta r^2)c_2 = 0,$$

which relate all the parameters. These are linear, homogeneous equations for c_1 and c_2 . For nontrivial solutions the determinant of the coefficient matrix must be zero, or

$$(\sigma + \mu r^2)(\sigma + k + \delta r^2) - f\nu \bar{a}r^2 = 0.$$

This equation relates the temporal growth factor σ , the spatial frequency r , and the other constants in the problem. Expanded out, this equation is quadratic in σ , specifically

$$\sigma^2 + \gamma_1\sigma + \gamma_2 = 0,$$

where

$$\gamma_1 = r^2(\mu + \gamma) + k > 0, \quad \gamma_2 = r^2[\mu(\delta r^2 + k) - f\nu \bar{a}].$$

The roots of this quadratic are

$$\sigma = \frac{1}{2}(-\gamma_1 \pm \sqrt{\gamma_1^2 - 4\gamma_2}).$$

Clearly, one of the roots is always negative or has a negative real part. The other root can have a positive or negative real part, depending on the value of the discriminant $\gamma_1^2 - 4\gamma_2$. We are interested in determining whether there are parameter choices that lead to an instability; therefore we want σ positive. Hence, γ_2 must be negative, or

$$\mu(\delta r^2 + k) < f\nu \bar{a}.$$

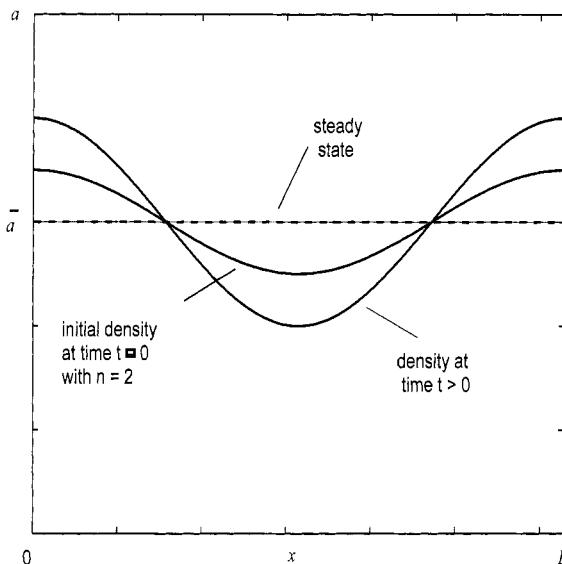


Figure 6.14 Plot showing the growing amoeba density at times $t = 0$ and $t > 0$ when the uniform state is unstable to local perturbations in the mode $n = 2$. The instability gives rise to two aggregation sites at the ends of the medium where the density is higher and is growing as time increases.

If this inequality holds, there is an unstable mode and perturbations will grow. Thus, for each value of n we obtain a frequency $r = r_n$ and a corresponding growth factor σ_n , and the n th mode grows when

$$\mu \left(\delta \frac{n^2 \pi^2}{L^2} + k \right) < f \nu \bar{a}. \quad (6.6.12)$$

We can ask what factors destabilize the uniform state in the amoeba-cAMP system and promote aggregation—that is, when is (6.6.12) likely to hold? We can list the factors that may make the left side of the inequality (6.6.12) smaller than the right side: low motility μ of the bacteria, low degradation rate k or large production rate f of cAMP, large chemotactic strength ν , large dimensions L of the medium, a small value of n (thus, low frequency perturbations are less stabilizing than high frequency perturbations), and decreasing the diffusion constant of the cAMP. Figure 6.14 shows time snapshots of the amoeba density for the mode $n = 2$ when it is unstable. The regions where the amplitude is high correspond to higher concentrations of amoeba, that is, regions of aggregation.

□

Example. (*Predator–Prey Model with Diffusion*) Consider the scaled Lotka–Volterra model of population dynamics with added diffusion of the two species:

$$u_t = Du_{xx} + u(1 - v), \quad (6.6.13)$$

$$v_t = Dv_{xx} + av(u - 1), \quad (6.6.14)$$

where $a > 0$ and D is the diffusion constant. The population densities are u and v , where u represents prey, and v represents predators. When diffusion is removed ($D = 0$), we obtain the Lotka–Volterra model for the spatially independent populations

$$u_t = u(1 - v), \quad (6.6.15)$$

$$v_t = av(u - 1). \quad (6.6.16)$$

Equations (6.6.15)–(6.6.16) have equilibria $(0, 0)$ and $(1, 1)$, and it is routine to show that $(0, 0)$ is a saddle point (the Jacobian of the linearization has real unequal eigenvalues) and $(1, 1)$ is a center for the linearized system (the Jacobian matrix has purely imaginary eigenvalues); in the latter case we can easily show that $(1, 1)$ is also a center for (6.6.15)–(6.6.16). What happens when there is diffusion? On an interval $x \in I = [0, L]$ with no-flux boundary conditions, namely

$$u_x = v_x = 0, \quad x = 0, L,$$

both populations tend to spatially uniform states for $t \rightarrow \infty$. The following argument is due to Murray (1975) and is based on the fact that there is a conservation law for the diffusionless system (6.6.15)–(6.6.16) given by

$$a(u - \ln u) + v - \ln v = \text{constant}.$$

(To see this, divide the two equations in (6.6.15)–(6.6.16), separate variables, and integrate.) Therefore, we define

$$s(x, t) = a(u - \ln u) + v - \ln v.$$

Note that s is constant when $D = 0$. A straightforward exercise with partial differentiation shows

$$s_t - Ds_{xx} = -aD \left(\frac{u_x}{u} \right)^2 - D \left(\frac{v_x}{v} \right)^2 \leq 0.$$

The conditions on s are

$$s_x = 0, \quad x = 0, L,$$

and

$$s(x, 0) = a(u_0(x) - \ln u_0(x)) + v_0(x) - \ln v_0(x) \equiv s_0(x).$$

Now define the energy

$$S(t) = \int_I s(x, t) dt.$$

Then, using integration by parts on the s_{xx} term and the no-flux boundary conditions:

$$\begin{aligned} S'(t) &= \int_I s_t dx \\ &= D \int_I \left((s_{xx} - a \left(\frac{u_x}{u} \right)^2 - \left(\frac{v_x}{v} \right)^2) \right) dx \\ &= -D \int_I \left(s_x^2 + a \left(\frac{u_x}{u} \right)^2 + D \left(\frac{v_x}{v} \right)^2 \right) dx \leq 0. \end{aligned}$$

Therefore $S(t)$ is nonincreasing. Moreover, it is bounded below because

$$S(t) = \int_I s(x, t) dx \geq \int_I (a+1) dx = L(a+1).$$

We can therefore conclude that $S(t)$ approaches a finite limit as $t \rightarrow \infty$. It follows that $S'(t)$ approaches zero as $t \rightarrow \infty$, which forces u_x and u_v to approach zero as $t \rightarrow \infty$. Therefore, the system (6.6.13)–(6.6.14) with no-flux conditions approaches a uniform state. This shows, for example, that no patterns can form in this model.

Example. (Critical Patch Size) For a population undergoing growth and dispersal, the critical patch size is the minimum size of the spatial domain needed for their survival. In the simplest case of linear growth, we take the population model

$$u_t = Du_{xx} + ru, \quad x \in (0, L), t > 0,$$

with $u(0, t) = u(L, t) = 0$ for all $t > 0$, and

$$u(x, 0) = f(x), \quad x \in [0, L].$$

This problem can be solved easily using separation of variables. Putting $u = \phi(t)y(x)$ and substituting into the PDE and boundary conditions gives

$$\frac{\phi'}{\phi} = \frac{Dy'' + ry}{y} = -\lambda, \quad y(0) = y(L) = 0.$$

Therefore $\phi(t) = e^{-\lambda t}$ and y satisfy the Sturm–Liouville problem

$$Dy'' + (r + \lambda)y = 0, \quad y(0) = y(L) = 0.$$

It is easy to see that the eigenvalues and eigenfunctions are

$$\lambda_n = -r + D \frac{n^2 \pi^2}{L^2}, \quad y_n = \sin \frac{n\pi x}{L}, \quad n = 1, 2, 3, \dots,$$

Using superposition we have

$$u(x, t) = \sum_{n=1}^{\infty} a_n \exp\left(rt - D \frac{n^2 \pi^2}{L^2} t\right) \sin \frac{n\pi x}{L},$$

where the a_n are determined by the initial condition and orthogonality. We get

$$a_n = \frac{2}{L} \int_0^L f(x) \sin n\pi x L dx.$$

Notice that the solution goes to zero if

$$r < \frac{Dn^2\pi^2}{L^2} \quad \text{for all } n = 1, 2, 3, \dots$$

But this inequality holds for all n if it holds for $n = 1$, or

$$r < \frac{D\pi^2}{L^2}.$$

Thus, for the population to survive, we need

$$r \geq \frac{D\pi^2}{L^2},$$

or

$$L \geq \pi \sqrt{\frac{D}{r}}.$$

The right side of this inequality is the critical domain size. What it means is that there is at least one Fourier mode that grows. Problems of this type become more interesting when the growth rate is nonlinear, or the spatial dimension is 2. \square

EXERCISES

1. Solve the linear system

$$u_t = u_{xx} + v, \quad v_t = v_{xx} - u, \quad x \in (0, L),$$

with no-flux boundary conditions on u and v at $x = 0, L$, and initial conditions

$$u(x, 0) = u_0(x), \quad v(x, 0) = v_0(x), \quad x \in (0, L).$$

Describe the long-time behavior of the system.

2. For the following reaction–diffusion system

$$u_t = u_{xx} + 3u + 13v, \quad v_t = 9v_{xx} - u - 3v, \quad x \in (0, \pi),$$

with no-flux boundary conditions, show that diffusion can destabilize an equilibrium state. Proceed by showing that the trivial solution is stable. Then do a stability analysis and determine which modes are unstable.

3. Extend the analysis of (6.6.13)–(6.6.14) to the problem of n dimensions, that is, to

$$\begin{aligned} u_t &= \Delta u_{xx} + u(1 - v), \\ v_t &= \Delta v_{xx} + av(u - 1), \end{aligned}$$

where $x \in \Omega$, a nice bounded region in \mathbb{R}^n , and

$$\frac{du}{dn} = \frac{dv}{dn} = 0, \quad x \in \partial\Omega.$$

4. A phytoplankton–zooplankton model is given by

$$\begin{aligned} P_t &= P_{xx} + rP(1 - P) - \frac{ZP^2}{a^2 + P^2}, \\ Z_t &= DZ_{xx} + \gamma Z \left(\frac{P^2}{a^2 + P^2} - \mu \right). \end{aligned}$$

Explain the origin of the terms in these equations. Assume that a positive equilibrium solution \bar{P} , \bar{Z} exists, and show that it is not possible to have diffusion-driven instabilities in this system.

5. Examine the stability of the equilibrium solution of the Turing system

$$u_t = Du_{xx} + 1 - u + u^2v, \quad v_t = v_{xx} + 2 - u^2v, \quad 0 < x < \pi,$$

under no-flux boundary conditions. Specifically, determine values of D for which various modes n are unstable.

6. Consider Fisher's equation with Dirichlet boundary conditions:

$$\begin{aligned} u_t &= u_{xx} + u(1 - u), \quad -\frac{\pi}{2} < x < \frac{\pi}{2}, \\ u &= 3 \quad \text{at} \quad x = \pm\frac{\pi}{2}. \end{aligned}$$

- (a) Show that $u_e(x) = 3/(1 + \cos x)$ is a nonconstant steady state solution.
- (b) Define perturbations $U(x, t)$ by the equation $u = u_e(x) + U(x, t)$, and find the linearized perturbation equation and boundary conditions for $U(x, t)$.
- (c) Assume a solution to the linearized equation of the form $U = e^{\sigma t}g(x)$, and show that g must satisfy

$$g'' + \frac{\cos x - 5}{1 + \cos x}g = \sigma g, \quad g = 0 \quad \text{at} \quad x = \pm\frac{\pi}{2}. \quad (6.6.17)$$

- (d) Show that if the perturbation equation has a nontrivial solution, then $\sigma < 0$, thereby showing local stability of the steady solution. Hint: Consider two cases, when g is positive and when g is negative on the interval, and then examine at the signs of g'' and the other terms in (5.41) at a maximum or minimum point.
7. In this extensive exercise we investigate a model developed by Gierer & Meinhardt (1972) to show how two-dimensional instabilities can be driven by diffusive effects. In two dimensions and time let $a = a(x, y, t)$ and $h = h(x, y, t)$ be concentrations of two chemicals (an activator and an inhibitor) in a fixed rectangular domain $0 \leq x \leq L_x$, $0 \leq y \leq L_y$. Both chemical species have natural decay, and both diffuse with unequal diffusion constants. Therefore, we consider the model

$$\begin{aligned} a_t &= c_1 \frac{a^2}{h} - \mu a + D_1 \Delta a, \\ h_t &= c_2 a^2 - \nu h + D_2 \Delta h, \end{aligned}$$

where the normal derivatives of a and h vanish on the boundary of the rectangle. The first terms on the right sides of both equations represent the chemical kinetics.

- (a) Show that the model may be nondimensionalized and written in the form

$$\begin{aligned} u_t &= \alpha \frac{u^2}{v} - u + \Delta u, \\ v_t &= \beta u^2 - \gamma v + D \Delta v, \end{aligned}$$

where α , β , γ , and D are appropriately chosen dimensionless constants. (Hint: Scale time by μ^{-1} , and scale the spatial variables by $\sqrt{D_1/\mu}$.) What is the scaled spatial domain $0 \leq x \leq L_1$, $0 \leq y \leq L_2$? Note that the x , y , and t in this system are scaled variables.

- (b) Show that a uniform equilibrium solution is

$$u^* = \frac{\alpha\gamma}{\beta}, \quad v^* = \frac{\alpha^2\gamma}{\beta}.$$

- (c) Assuming $u = u^* + \phi(x, y, t)$ and $v = v^* + \psi(x, y, t)$, where ϕ and ψ are small perturbations from equilibrium, show that the linearized system of perturbation equations is

$$\frac{\partial}{\partial t} \begin{pmatrix} \phi \\ \psi \end{pmatrix} = \begin{pmatrix} 1 & -\frac{1}{\alpha} \\ 2\alpha\gamma & -\alpha \end{pmatrix} \begin{pmatrix} \phi \\ \psi \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & D \end{pmatrix} \begin{pmatrix} \Delta\phi \\ \Delta\psi \end{pmatrix},$$

where the normal derivatives of ϕ and ψ vanish on the boundary of the scaled domain.

- (d) Assume modal solutions in (c) having the form

$$\begin{pmatrix} \phi \\ \psi \end{pmatrix} = \begin{pmatrix} A \\ B \end{pmatrix} e^{\sigma t} \cos \frac{m\pi x}{L_1} \cos \frac{n\pi y}{L_2},$$

where m and n are nonnegative integers, A and B are constants, and σ is the growth rate. Why can you do this? Show that nontrivial solutions of this form exist if, and only if

$$\det(\sigma I - M + q^2 N) = 0, \quad (6.6.18)$$

where

$$M = \begin{pmatrix} 1 & -\frac{1}{\alpha} \\ 2\alpha\gamma & -\alpha \end{pmatrix}, \quad N = \begin{pmatrix} 1 & 0 \\ 0 & D \end{pmatrix},$$

and

$$q^2 = \left(\frac{m\pi}{L_1} \right)^2 + \left(\frac{n\pi}{L_2} \right)^2.$$

- (e) Show that condition (6.6.18) reduces to

$$\sigma^2 + [-\text{tr } M + (N + 1)q^2]\sigma + Dq^4 - \text{tr } M q^2 + \det M = 0,$$

and conclude that $\text{Re } \sigma < 0$ if, and only if

$$\text{tr } M - (N + 1)q^2 < 0, \quad (6.6.19)$$

$$Dq^4 - \text{tr } M q^2 + \det M > 0. \quad (6.6.20)$$

- (f) Assume that the diffusionless steady state (occurring when $D_1 = D_2 = 0$) is stable, or $\text{tr } M < 0$ and $\det M > 0$. Show that (6.6.19) holds automatically and therefore, for instability, that equation (6.6.20) must be violated.
- (g) Show that there exist parameter values that violate condition (6.6.20). In fact, show that when

$$1 - \alpha D > \sqrt{2D \det M} > 0,$$

then there is a range of values of q that give this condition.
Hint: The left side of (6.6.20) is a concave down parabola in the variable q^2 . Find the minimum.)

- (h) Explain why the preceding calculations show that a diffusive instability arises in the Gierer–Meinhardt model.

8. In a fishing zone $0 \leq x \leq L$, with $x = 0$ representing a straight shoreline and $x = L$ an outer boundary, fish grow according to the logistics law and are harvested at a constant per capita rate h . Allowing for diffusion, the model for the population density is

$$u_t = Du_{xx} + r \left(1 - \frac{u}{K}\right) - hu.$$

At the outer boundary assume the fish population is zero (caused, e.g., by excessive fishing), and assume a no-flux boundary condition at the shoreline. Initially, take $u(x, 0) = u_0(x)$, and assume $r > E$. Determine conditions under which the uniformly zero equilibrium population is stable to small perturbations, and therefore the fish population goes extinct. In particular, to prevent the population from collapsing, show that one must have $L \geq \sqrt{D/(4r - 4h)}$.

Reference Notes. The literature on reaction–diffusion equations is extensive. Edelstein-Keshet (2005) is an outstanding introduction. Other good starting places are books by Murray (2002, 2003), Smoller (1994), Fife (1979), Britton (1986, 2003), Grindrod (1996), and Allen (2007). All have detailed bibliographies that serve as a guide to further reading. Articles on pattern formation include Turing (1952), Levin & Segel (1985), and the popular article by Murray (1988). Applications in mathematical physiology, which unfortunately is slighted in this text, can be found in Keener & Sneyd (1998). For information about chemically reacting systems, especially combustion phenomena, see the Reference Notes in Chapter 7.

Traveling waves are discussed in Volpert & Volpert (1994). An introductory paper in TWS in groundwater flow is van Duijn & Knabner (1992). DeMarsily (1986) is an excellent quantitative treatment of contaminant transport and groundwater flow.

Equilibrium Models

In this final chapter we address some problems in the theory of nonlinear elliptic partial differential equations. We consider only some representative examples, many of which are elementary; we do not lay out a general theory. One way to think about elliptic equations is that they are equilibrium problems born out of the time-asymptotic limit of reaction–diffusion equations. For example, steady-state solutions, if they exist, of the parabolic reaction–diffusion equation

$$u_t - \Delta u = f(u) \quad (7.0.1)$$

must satisfy the equilibrium equation

$$-\Delta u = f(u). \quad (7.0.2)$$

The latter semilinear equation is a steady-state equation in that time does not appear, and $u = u(x)$, a function of the spatial variable only. So one may think of the solutions $u(x, t)$ of (7.0.1) as evolving into the solutions $u = u(x)$ of (7.0.2) as t gets large. One of the major problems in the study of reaction–diffusion equations is to confirm this observation in special cases, and it leads to questions of existence of solutions of both (7.0.1) and (7.0.2) and also to questions of stability and bifurcation.

Equilibrium, or steady-state, equations such as (7.0.2) are elliptic, and they are fundamentally different from parabolic and hyperbolic equations. Initial conditions for elliptic equations are inappropriate; rather, only boundary conditions lead to well-posed elliptic problems. Thus, concepts such as characteristics, which are fundamental for hyperbolic problems, play no role in the

steady state. For elliptic equations we look for solutions on a spatial domain that match given conditions on the boundary of that domain; we do not regard, as in the case of hyperbolic or parabolic equations, initial or boundary data as being carried into the domain along, for example, characteristic curves.

In Section 7.1 a few basic elliptic models are reviewed. In Section 7.2 we present two of the fundamental tools in the study of nonlinear elliptic equations: the maximum principle and the basic existence theorem. Section 7.3 focuses on a nonlinear eigenvalue problem, and in Section 7.4 we address the question of the stability of equilibrium solutions to (7.0.1) and the related issue of the bifurcation of solutions of (7.0.2) when a parameter is present. Finally, we introduce nonlinear stability analysis based on normal modes.

The field of nonlinear elliptic boundary value problems is a vast area of intensive study, and our treatment is parsimonious when compared to evolution equations. The reader is invited to explore this field further by consulting one or more of the references cited in the Reference Notes at the end of the chapter.

7.1 Elliptic Models

In this section we review some classical elliptic models, observe how nonlinearities can appear, and focus in on some key questions that naturally arise.

Example. (*Laplace's Equation*) Let D be an open, bounded, connected region in \mathbb{R}^n with boundary ∂D , and let $u = u(x, t)$ denote a density in D at position $x = (x_1, x_2, \dots, x_n)$ at time t . Suppose that at time $t = 0$ we specify the density to be $u(x, 0) = u_0(x)$, and suppose that for all $t \geq 0$ we impose the boundary condition

$$u(x, t) = h(x), \quad x \in \partial D, \quad t > 0, \quad (7.1.1)$$

where h is given. The basic conservation law and Fick's law imply that the density evolves according to the diffusion equation

$$u_t - \Delta u = 0, \quad x \in D, \quad t > 0, \quad (7.1.2)$$

where we assume that the region is homogeneous and the diffusion constant is unity. Δ is the n -dimensional Laplacian. If we imagine that the system evolves for long enough time that the initial condition at $t = 0$ no longer affects changes in the system, then we reach an asymptotic state that is independent of time [i.e., $u = u(x)$] that satisfies

$$\Delta u = 0, \quad x \in D, \quad (7.1.3)$$

along with the boundary condition (7.1.1). The elliptic equation (7.1.3) is *Laplace's equation*, and the boundary value problem consisting of Laplace's equation (7.1.3) and the boundary condition (7.1.1) is the *Dirichlet problem* on the domain D . If the boundary condition (7.1.1) is replaced by a flux condition of the form

$$\frac{du}{dn} = g(x), \quad x \in \partial D, \quad t > 0, \quad (7.1.4)$$

where $du/dn = \text{grad } u \cdot \mathbf{n}$, and \mathbf{n} is the outward unit normal on ∂D , then the boundary value problem consisting of Laplace's equation (7.1.3) and the flux condition (7.1.4) is the *Neumann problem*. More generally, a boundary condition of the form

$$a(x)u + b(x)\frac{du}{dn} = c(x), \quad x \in \partial D, \quad t > 0 \quad (7.1.5)$$

is called a mixed condition (or a *Robin condition*).

The Dirichlet, Neumann, and Robin problems are three of the most important problems in analysis. Questions of existence, uniqueness, and how to construct solutions are key. \square

There is another issue with elliptic equations, namely the lack of continuous dependence of the solution on Cauchy data. As the next important example shows, the Cauchy problem is not well-posed for elliptic equations, in that small changes in the Cauchy data lead to arbitrarily large changes in the solution.

Example. (Hadamard's Example) Consider the Cauchy problem

$$\begin{aligned} u_{xx} + u_{yy} &= 0, \\ u(x, 0) &= 0, \quad u_y(x, 0) = 0. \end{aligned}$$

Clearly, $u = 0$ is a solution to this problem. Now change the Cauchy data to

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

which, for large n , represents only a small change in the data. It is straightforward to verify that the solution is given by

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

For large values of n the exponentially growing sinh function dominates the n^{-2} factor, and hence the solution deviates from the original solution $u = 0$ by an arbitrarily large amount. Thus a small change in the Cauchy data produces an arbitrarily large change in the solution. \square

Obtaining solutions to Laplace's equation on bounded domains D can sometimes be accomplished by separation of variables, provided the domain has a simple geometric form, for example, bounded by coordinate planes in a curvilinear coordinate system. On unbounded domains transform methods may apply. Conformal transformations may be applicable to both bounded and unbounded domains in two dimensions. Any text on complex variables can be consulted for conformal techniques.

Example. (*Poisson's Equation*) Returning to the discussion in the preceding example, if a time-independent source $f(x)$ is present in the domain D , then the density $u = u(x, t)$ is governed by the parabolic nonhomogeneous diffusion equation

$$u_t - k\Delta u = f(x), \quad x \in D, \quad t > 0.$$

Equilibrium solutions $u = u(x)$ are therefore governed by the elliptic equation

$$-\Delta u = f(x), \quad x \in D, \quad (7.1.6)$$

which is *Poisson's equation*. For boundary conditions on ∂D we can append either a Dirichlet, Neumann, or Robin condition. If the source depends on the u , we obtain a semilinear elliptic equation of the form

$$-\Delta u = f(x, u), \quad x \in D. \quad \square \quad (7.1.7)$$

Equation (7.1.7) is one of the basic nonlinear elliptic equations; when the nonlinearity occurs in the source term and not in the differential operator, the equation is in general called *semilinear*.

Example. (*Reaction-Diffusion Equations*) As we pointed out in the last chapter, there is considerable interest in nonlinear reactive-diffusive systems. These systems are governed, for example, by semilinear equations of the form

$$u_t - d_1\Delta u = f(u, v), \quad v_t - d_2\Delta v = g(u, v), \quad (7.1.8)$$

where u and v represent unknown density functions (populations, chemical concentrations, etc.) and d_1 and d_2 are diffusion constants. These equations provide qualitative models for a number of biological, chemical, and physical phenomena: predator-prey interactions, chemical reactions, combustion phenomena, morphogenesis, and nerve-pulse conduction, to mention only a few. We are often interested in the equilibrium situation where u and v are time-independent. In this case, $u = u(x)$ and $v = v(x)$ satisfy the elliptic system

$$-d_1\Delta u = f(u, v), \quad -d_2\Delta v = g(u, v), \quad (7.1.9)$$

subject to the boundary conditions imposed on the problem. We can think of the solutions of (7.1.9) as the asymptotic limit of the solutions of (7.1.8), provided that solutions exist. \square

The interpretation above leads to several interesting questions. For example, what kinds of initial data imposed on (7.1.8) cause its solutions to evolve into the stationary solutions of the equilibrium problem (7.1.9)? This question becomes even more exciting when one discovers that (7.1.9) may admit multiple equilibrium solutions. So, we can liken equilibrium solutions of (7.1.9) to critical points for plane autonomous systems of ODEs and then ask similar stability questions for the stationary solutions of the PDEs (7.1.9). For example, is a given stationary solution of (7.1.9) asymptotically stable, that is, does it attract all solutions of (7.1.8) that are sufficiently close to it (measured in some norm) at time $t = 0$? These are important questions in nonlinear analysis associated with reaction–diffusion equations and the resulting elliptic systems that arise in a time-asymptotic limit.

Another important task is to investigate bifurcation, or branching, phenomena in systems such as (7.1.8) and (7.1.9). For example, the system (7.1.8) may contain a parameter such that below a critical value there is a single, asymptotically stable equilibrium solution of (7.1.9); then, as the parameter increases beyond the critical value, additional equilibria appear and the previous asymptotically stable solution may become unstable. The study of multiple equilibria and their stability properties, as functions of parameters in the system, lies at the foundation of bifurcation theory.

Yet another class of problems associated with elliptic equations are nonlinear eigenvalue problems. These are discussed in Section 7.4. A classic linear eigenvalue problem is introduced in the next example.

Example. (Schrödinger Equation) In the doctrine of classical mechanics the position of a particle of mass m moving under the influence of a potential $V(x)$, $x \in \mathbb{R}^3$, is determined precisely by solving the differential equations of motion given by Newton's second law, subject to initial conditions. After the turn of the twentieth century it was recognized that the classical theory did not apply in all cases, and quantum mechanics was developed. In quantum theory a statistical interpretation is advanced where all that can be known about a particle's location is the probability of it being in some region of space. The information about a quantum mechanical system is contained in a state function $\psi(x, t)$, called the wavefunction, where $|\psi|^2$ represents the probability density; that is,

$$\int_D |\psi|^2 dx$$

represents the probability of finding the particle in the volume D . It is a fundamental postulate of quantum mechanic that the wave function satisfies the time-dependent Schrödinger equation

$$i\hbar\psi_t = \left[-\hbar^2 \frac{\Delta}{2m} + V(x) \right] \psi, \quad (7.1.10)$$

where Δ is the three-dimensional Laplacian and $\hbar = 2\pi\hbar$ is Planck's constant. If we assume that $\psi(x, t) = T(t)\Psi(x)$, then (7.1.10) separates into the two equations

$$T' = -\frac{iE}{\hbar}T \quad \text{and} \quad \Delta\Psi + \frac{2m}{\hbar^2}[E - V(x)]\Psi = 0,$$

where E is a separation constant. The equation for Ψ is called the *time-independent Schrödinger equation*, and it is elliptic. When boundary conditions are imposed, the time-independent problem is an eigenvalue problem where the eigenvalues E are the possible energy levels of the particle. \square

EXERCISES

1. The divergence theorem in \mathbb{R}^n is contained in the statement

$$\int_D D_k u(x) dx = \int_{\partial D} u(x) n_k(x) dS, \quad D_k = \frac{\partial}{\partial x_k},$$

where D is an open, bounded region in \mathbb{R}^n with smooth boundary ∂D , n_k is the k th component of the outward unit normal on ∂D , and u is a scalar function with $u \in C(\overline{D}) \cap C^1(D)$.

- (a) Use the divergence theorem to derive the integration-by-parts formula

$$\int_D v D_k u \, dx = \int_{\partial D} u v n_k \, dS - \int_D u D_k v \, dx.$$

- (b) Use the integration-by-parts formula to derive Green's identities:

$$\begin{aligned} \int_D v \Delta u \, dx &= \int_{\partial D} v \frac{du}{dn} \, dS - \int_D \operatorname{grad} u \cdot \operatorname{grad} v \, dx, \\ \int_D v \Delta u \, dx &= \int_D u \Delta v \, dx + \int_{\partial D} \left(v \frac{du}{dn} - u \frac{dv}{dn} \right) \, dS, \end{aligned}$$

where Δ is the n -dimensional Laplacian, $\operatorname{grad} = (D_1, \dots, D_n)$ is the gradient operator, and du/dn is the normal derivative given by

$$\frac{du}{dn} = \operatorname{grad} u \cdot \mathbf{n}$$

where \mathbf{n} is the normal vector.

- (c) Prove that classical solutions in $C(\overline{D}) \cap C^2(D)$ to the Dirichlet problem

$$\begin{aligned}-\Delta u &= g(x), \quad x \in D, \\ u &= f(x), \quad x \in \partial D,\end{aligned}$$

are unique, provided that they exist. Assume $f, g \in C(\overline{D})$.

2. Use an energy method to show that the parabolic problem

$$\begin{aligned}u_t - \Delta u &= 0, \quad x \in D, \quad t > 0, \\ u(x, t) &= 0, \quad x \in \partial D, \quad t > 0, \\ u(x, 0) &= 0, \quad x \in D,\end{aligned}$$

has only the trivial solution $u = 0$ on $D \times \mathbb{R}^+$. Hint: Consider $E(t) = \int_D u(x, t)^2 dx$ and use Exercise 1(b).

3. Determine all solutions of Laplace's equation $\Delta u = 0$ of the form $u = \psi(r)$, $r = |x - \xi|$, where $x = (x_1, \dots, x_n)$, where $\xi = (\xi_1, \dots, \xi_n)$ is a fixed point in \mathbb{R}^n , and where r is the Euclidean distance from x to ξ . In particular, show

$$\psi(r) = A + C \ln r \quad \text{if } n = 2; \quad \psi(r) = A + \frac{Cr^{2-n}}{2-n} \quad \text{if } n > 2,$$

where A and C are constants.

4. Show that the PDE

$$u_{xx} + x^2 u_{yy} = y u_y$$

is elliptic. Determine characteristic coordinates and reduce the equation to canonical form.

5. Consider the initial value problem for the Schrödinger equation for a free particle

$$\begin{aligned}u_t &= iu_{xx}, \quad x \in \mathbb{R}, \quad t > 0, \\ u(x, 0) &= u_0(x), \quad x \in \mathbb{R},\end{aligned}$$

where u_0 is continuous and square integrable on \mathbb{R} . Prove that if a solution $u(x, t)$ exists and vanishes, along with all of its derivatives, at infinity, then

$$\int_{\mathbb{R}} |u(x, t)|^2 dx = \int_{\mathbb{R}} |u_0(x)|^2 dx \quad \text{for all } t > 0.$$

(This scaled form of the Schrödinger equation, without the accompanying constants, is often discussed in mathematics literature.)

6. Use the solution to the initial value problem for the heat equation (see Section 5.1) to determine a formal solution to the initial value problem in Exercise 5. Compare and contrast the two kernels in the integral representations of the solutions to these two initial value problems.
7. The Helmholtz equation in \mathbb{R}^3 is

$$\Delta u + cu = 0,$$

where c is a positive constant and Δ is the three-dimensional Laplacian. Find all solutions of the form $u(x) = \Psi(|x - \xi|)$, where ξ is a fixed point in \mathbb{R}^3 .

8. (*Korteweg-deVries and Schrödinger*) There is a close connection between potentials in the Schrödinger equation and solutions to the KdV equation, and this exercise explores one of them. Let $u = u(x, t)$ be a solution to the KdV equation

$$u_t + u_{xxx} + 6uu_x = 0,$$

and suppose that u is chosen as the potential in the time-independent Schrödinger equation

$$-(\psi_{xx} + u\psi) = \lambda\psi, \quad (7.1.11)$$

where t appears as a parameter in λ and ψ . Assume that u and its derivatives approach zero as $|x| \rightarrow \infty$, and assume

$$\int_{\mathbb{R}} |\psi|^2 dx = 1.$$

Show that if λ is an eigenvalue in (7.1.11) with corresponding eigenfunction ψ , then, in fact, λ is independent of t . *Hint:* Substitute u from (7.1.11) into the KdV equation and show that

$$\lambda'(t)\psi^2 + (\psi h_x - h\psi_x)_x = 0,$$

for some appropriately chosen function h .

7.2 Theoretical Results

Another feature of elliptic problems is that a solution must satisfy a condition requiring that it take on its maximum or minimum value on the boundary of the domain over which the problem is defined. In this section we present two of the fundamental tools used in the analysis of elliptic problems, the maximum principle and the basic existence theorem.

7.2.1 Maximum Principle

In the sequel we assume that D is an open, bounded, connected subset of \mathbb{R}^n , and we denote the points of D by $x = (x_1, \dots, x_n)$. The symbol \overline{D} will denote the *closure* of D , that is, $\overline{D} = D \cup \partial D$, where ∂D is the boundary of D . We will not deal with questions of regularity of the boundary; generally, we assume that the boundary of D is sufficiently smooth to give validity to our results. For elliptic equations it is not much more difficult to work in n -dimensional space, and then our results will be valid when specialized to either \mathbb{R}^2 or \mathbb{R}^3 . The Laplacian in \mathbb{R}^n is the second-order differential operator Δ defined by

$$\Delta u = u_{x_1 x_1} + \cdots + u_{x_n x_n}$$

We say that a function $u \in C^2(D)$ is *subharmonic* (*superharmonic*) in D if $-\Delta u \leq 0$ in D (correspondingly, $-\Delta u \geq 0$ in D). A function $u \in C^2(D)$ that satisfies Laplace's equation $\Delta u = 0$ in D is said to be *harmonic* in D .

The maximum principle, in its simplest form, requires that a subharmonic function on D that is continuous in \overline{D} assume its maximum value on the boundary ∂D . Correspondingly, the minimum principle insists that a superharmonic function assume its minimum value on ∂D . It is clear that these results should be expected if one restricts the analysis to $n = 1$, where D is an interval in R^1 and u is a function on D . In this case the Laplacian is just the second derivative operator and $u'' \geq 0$ means that u is concave up, and $u'' \leq 0$ means that u is concave down. Respectively, such functions have their maximum and their minimum at the endpoints (assuming that they are not constant functions). In fact, the one-dimensional case is a convenient mnemonic device. Therefore, we have the following weak form of the maximum principle.

Theorem. (*Weak Maximum Principle*) If $u \in C(\overline{D}) \cap C^2(D)$ and $\Delta u \geq 0$ on D , then $\max_{\overline{D}} u = \max_{\partial D} u$.

Because u is a continuous function on a compact (closed and bounded) set, the function u must assume its maximum somewhere on \overline{D} . The theorem states that the maximum must occur at least somewhere on ∂D , but it does not preclude the maximum also occurring at an interior point. A stronger version of the maximum principle, stated later, will preclude this, unless the function is constant. Now the proof of the weak maximum principle (this proof is given in most books on PDEs). First, we observe that if $\Delta u > 0$ on D (strict inequality), then a maximum cannot occur in D ; at such a point $u_{x_i x_i} \leq 0$ for all i , which violates $\Delta u > 0$ at that point. Next consider the case $\Delta u \geq 0$ on D . For $\varepsilon > 0$ we define the auxiliary function $v = u + \varepsilon|x|^2$ for which $\Delta v > 0$. From the

strict inequality case, we know that v cannot assume its maximum in D . Then $\max_{\overline{D}} v = \max_{\partial D} v$. Consequently

$$\max_{\overline{D}} u \leq \max_{\overline{D}} (u + \varepsilon |x|^2) = \max_{\partial D} (u + \varepsilon |x|^2) \leq \max_{\partial D} u + \varepsilon \max_{\partial D} |x|^2.$$

Because ε is arbitrary, $\max_{\overline{D}} u \leq \max_{\partial D} u$. The opposite inequality is automatically true, and therefore the theorem is proved. \square

Replacing u by $-u$ in the statement of the theorem gives a weak form of the minimum principle, as follows.

Theorem. (Weak Minimum Principle) If $u \in C(\overline{D}) \cap C^2(D)$ and $\Delta u \leq 0$ on D , then $\min_{\overline{D}} u = \min_{\partial D} u$.

It is obvious that if u is harmonic in D and continuous on \overline{D} , then u must assume both its maximum and minimum on ∂D . In particular, if $u = 0$ on ∂D and $\Delta u = 0$ in D , then u must vanish identically in D . Therefore, a harmonic function on D that is continuous in \overline{D} is uniquely determined by its values on the boundary.

We now state, without proof, the strong version of the maximum principle. A proof can be found in Protter & Weinberger (1967), which is an excellent reference for maximum principles. We formulate the theorem for more general elliptic operators than the Laplacian. To this end, let us consider an operator L defined by

$$Lu = \sum_i \sum_j a_{ij}(x)u_{x_i x_j} + \sum_j b_j(x)u_{x_j}, \quad (7.2.1)$$

where the coefficients a_{ij} and b_j are continuously differentiable functions on D that are continuous on \overline{D} , and where

$$\sum_i \sum_j a_{ij}(x)\xi_i \xi_j \geq \mu \sum_j \xi_j^2 \quad (7.2.2)$$

for all x in D , for some positive real number μ . An operator L satisfying these conditions is called *uniformly elliptic* in D . Of course, the Laplacian Δ is uniformly elliptic. Then, one version of the strong form of the maximum principle can be stated as follows.

Theorem. (Strong Maximum Principle) Let $u \in C(\overline{D}) \cap C^2(D)$, and assume that $Lu + c(x)u \geq 0$ for $x \in D$, where L is uniformly elliptic in D , and where the function c is continuous, bounded, and $c \leq 0$ on D . If u attains a nonnegative maximum M in D , then $u = M$ for all $x \in D$.

One can also draw some conclusions about the normal derivative of u at the boundary.

Theorem. Under the same hypotheses, assume that $u \leq M$ in D , $u = M$ at a boundary point x_0 , and $M \geq 0$. If x_0 lies on a boundary of some sphere inside D , then $du/dn > 0$ at x_0 , provided that u is not constant in D .

If the boundary of D is smooth, then an interior ball, or sphere, can always be placed at x_0 . It can be shown that we need not restrict the conclusion of the theorem to the normal derivative; more generally, the directional derivative of u in any outward direction is positive.

The corresponding minimum principle may be stated as follows. The proof comes from replacing u by $-u$ in the weak maximum principle, and we leave it as a simple exercise.

Theorem. (Strong Minimum Principle) Let $u \in C(\overline{D}) \cap C^2(D)$, and assume that $Lu + c(x)u \leq 0$ for $x \in D$, where L is uniformly elliptic in D , and where c is continuous, bounded, and $c \leq 0$ on D . If u attains a nonpositive minimum m in D , then $u = m$ for all $x \in D$.

Once we have maximum and minimum principles, it is straightforward to obtain comparison theorems. For example, we have the following theorem.

Theorem Let $u, v \in C(\overline{D}) \cap C^2(D)$ and assume that $Lu \geq 0$ and $Lv = 0$ in D , where L is a uniformly elliptic operator given by (7.2.1). If $u \leq v$ on ∂D , then $u \leq v$ for all $x \in D$.

For the proof, let $w = v - u$. Then $w \geq 0$ on ∂D and $Lw \leq 0$ on D . Now, by way of contradiction, assume that $w(x_0) < 0$ for some x_0 in D . Then w has a negative minimum in D , and therefore by the strong minimum principle the function w must be constant and negative in D . But this contradicts the fact that w is nonnegative on ∂D . Hence $w \geq 0$ on D , giving the result. Our assumption that D is a bounded set is important in this theorem; it is not true, in general, for unbounded domains. \square

The Exercises contain several applications of the maximum and minimum principles.

7.2.2 Existence Theorem

Let us consider the semilinear boundary value problem

$$-Lu = f(x, u), \quad x \in D, \tag{7.2.3}$$

$$u = h(x), \quad x \in \partial D, \tag{7.2.4}$$

where L is the uniformly elliptic operator defined by (7.2.1). The function h in the Dirichlet boundary condition, and the function f , defined in $D \times \mathbb{R}$, are both assumed to be smooth (class C^1) on their respective domains. A minus sign appears in front of the operator L on the left side of (7.2.3) so that subsequent inequalities will have a symmetric form. (In fact, the maximum and minimum principles stated above are often presented with negative signs on the operator and the inequalities reversed.)

Another fundamental tool in the analysis of elliptic equations and eigenvalue problems for elliptic operators is the notion of upper and lower solutions to (7.2.4)–(7.2.4), and the fact that if such solutions exist, a classical solution can be sandwiched in between. In the next section, where we discuss eigenvalue problems, we will experience the power of this method to prove existence results. A function \bar{u} in $C(\overline{D}) \cap C^2(D)$ is called an *upper solution* to (7.2.3)–(7.2.4) if

$$-L\bar{u} \geq f(x, \bar{u}), \quad x \in D, \quad (7.2.5)$$

$$\bar{u} \geq h(x), \quad x \in \partial D. \quad (7.2.6)$$

A *lower solution* $\underline{u}(x)$ is defined similarly, with the two inequalities reversed. The fundamental theorem can now be stated.

Theorem. (Existence) If the boundary value problem (7.2.3)–(7.2.4) has a lower solution $\underline{u}(x)$ and an upper solution $\bar{u}(x)$ with $\underline{u}(x) \leq \bar{u}(x)$ in D , then the boundary value problem (7.2.3)–(7.2.4) has a solution $u(x)$ with the property that $\underline{u}(x) \leq u(x) \leq \bar{u}(x)$ in D .

Accessible proofs can be found in Sattinger (1973) and Smoller (1994). One method of proof uses the idea of monotone iteration schemes; that is, monotone sequences of iterates bounded by the upper and lower solutions are constructed and can be shown to converge a solution from both above and below. There are other proofs using topological, or degree, methods. Upper and lower solutions were introduced in the late 1960s. The proof of the general existence theorem using monotone iteration methods and upper and lower solutions was given by H. Amann in 1971 and D. Sattinger in 1972. The result is valid for boundary conditions more general than (7.2.4). Generalizations have been offered in other directions as well. The reader can consult the complete bibliography in Pao (1992) for specific references to these and other treatments of the subject.

Example. Let D be the square $0 < x, y < a$ in \mathbb{R}^2 and consider the boundary value problem (BVP)

$$-\Delta u = u(1 - u) \quad x \in D, \quad (7.2.7)$$

$$u = 0 \quad x \in \partial D, \quad (7.2.8)$$

where Δ is the two-dimensional Laplacian. We use the existence theorem to prove that a solution exists to this BVP, provided that the square D is large enough. The reader should be alert to the notation change, where we are using x and y in lieu of x_1 and x_2 . First, define $\bar{u}(x)$ as the solution to the ordinary differential equation $-\bar{u}'' = \bar{u}(1 - \bar{u})$ on $0 < x < a$, with $\bar{u}(0) = \bar{u}(a) = 0$. Then $\bar{u}(x) \geq 0$ on ∂D and $-\Delta \bar{u} = -\bar{u}'' = \bar{u}(1 - \bar{u}) \geq 0$. Therefore, $\bar{u}(x)$ is an upper solution of the BVP (7.2.7)–(7.2.8). To find a lower solution we proceed as follows. First, consider the eigenvalue problem

$$\begin{aligned}-\Delta u &= \lambda u && \text{in } D, \\ u &= 0 && \text{on } \partial D.\end{aligned}$$

The eigenvalues are given by $\lambda = \pi^2(n^2 + m^2)/a^2$ for $m, n = 1, 2, 3, \dots$. Thus the smallest eigenvalue is $\lambda_1 = 2\pi^2/a^2$, and an associated eigenfunction is given by $\phi(x, y) = \sin(\pi x/a) \sin(\pi y/a)$, which is positive on D . Then $-\Delta \phi = \lambda_1 \phi$ and $\phi = 0$ on ∂D . Now choose a such that $\lambda_1 < 1$, and take $\underline{u} = \varepsilon \phi$, where ε is a positive constant to be selected later. Then

$$\Delta \underline{u} + \underline{u}(1 - \underline{u}) = \varepsilon \phi(-\lambda_1 + 1 - \varepsilon \phi) \geq 0,$$

provided that ε is chosen small enough. Therefore, \underline{u} is a lower solution. By the existence theorem there is a solution u to (7.2.7)–(7.2.8) with the property that $\varepsilon \phi(x, u) \leq u(x, y) \leq \bar{u}(x)$ for all (x, y) in D , provided that a is large enough. \square

A similar existence theorem using upper and lower solutions holds for parabolic equations as well. For example,

$$u_t - \Delta u = f(x, u), \quad x \in D, \quad 0 < t < T, \quad (7.2.9)$$

$$u(x, t) = g(x, t), \quad x \in \partial D, \quad 0 < t < T, \quad (7.2.10)$$

$$u(x, 0) = u_0(x), \quad x \in D. \quad (7.2.11)$$

We assume that f is a smooth function of its arguments and that the initial and boundary conditions are continuous. We say that $\underline{u}(x, t)$ is a lower solution of (7.2.9)–(7.2.11) if (7.2.9)–(7.2.11) hold with the inequality \leq rather than with equality. In the same way, $\bar{u}(x, t)$ is an upper solution if (7.2.9)–(7.2.11) hold with the inequality \geq . Both \underline{u} and \bar{u} are assumed to be twice continuously differentiable on $\Omega_T = D \times (0, T)$ and continuous on Ω_T . Then we state, without proof, the following existence theorem.

Theorem. Let \underline{u} and \bar{u} be lower and upper solutions of (7.2.9)–(7.2.11) with $\underline{u}(x, t) \leq \bar{u}(x, t)$ on Ω_T . Then the initial boundary value problem (7.2.9)–(7.2.11) has a solution $u(x, t)$ with $\underline{u}(x, t) \leq u(x, t) \leq \bar{u}(x, t)$ on Ω_T that is continuous on $\bar{\Omega}_T$ and class $C^2(\Omega_T)$.

EXERCISES

1. Let $u \in C(\bar{D}) \cap C^2(D)$ be a solution of $\Delta u + \sum_{k=1}^n b_k(x)u_x + c(x)u = 0$ on D , where $c(x) < 0$ in D . Prove that $u = 0$ on ∂D implies that $u = 0$ in D .
Hint: Show that $\min u \geq 0$ and $\max u \leq 0$.
2. Use the maximum principle to prove that a harmonic function on D and continuous in \bar{D} is uniquely determined by its values on ∂D .
3. What can be deduced about solutions to the nonlinear Dirichlet problem

$$\begin{aligned}\Delta u &= u^2, & x \in D, \\ u(x) &= 0, & x \in \partial D?\end{aligned}$$
4. Let $D = \{(x, y) | 0 < x, y < \pi\}$ be a domain in \mathbb{R}^2 , and consider the Dirichlet problem

$$\begin{aligned}u_{xx} + u_{yy} + 2u &= 0 & \text{in } D, \\ u &= 0 & \text{on } \partial D.\end{aligned}$$
 Where does the maximum of u occur?

5. Let

$$u_t = u_{xx} + e^u, \quad x > 0, t > 0,$$

with $u = 0$ on $x = 0$ and $t = 0$. Show that $-\ln(1 - t)$ is an upper solution and $-\ln(1 - t + h(x, t))$ is a lower solution provided $h_t = h_{xx}$ with $h = t$ on $x = 0$ and $h = 0$ on $t = 0$. Show that $u(x, 1) \rightarrow +\infty$ as $x \rightarrow +\infty$.

6. Let Ω in \mathbb{R}^n denote the unbounded domain $|x| > 1$, and let $u \in C^2(\bar{\Omega})$, $\lim_{x \rightarrow \infty} u(x) = 0$, and $\Delta u = 0$ in Ω . Prove the $\max_{\bar{\Omega}} |u| = \max_{\partial\Omega} |u|$.

7.3 Eigenvalue Problems

7.3.1 Linear Eigenvalue Problems

The reader is familiar with eigenvalue problems occurring in matrix theory (the algebraic eigenvalue problem) and in ordinary differential equations (Sturm–Liouville problems). In Section 7.1 we mentioned how an eigenvalue problem arises naturally for the Schrödinger equation in quantum mechanics. In this section we examine eigenvalue problems for PDEs with the goal of obtaining results about eigenvalues for nonlinear equations.

One of the most important problems for linear partial differential operators is the eigenvalue problem for the negative Laplacian on a given open, bounded, connected domain D in \mathbb{R}^n . The problem is to determine values of λ for which the BVP

$$-\Delta u = \lambda u, \quad x \in D, \quad (7.3.1)$$

$$u = 0, \quad x \in \partial D, \quad (7.3.2)$$

has a nontrivial solution. To review, the values of λ for which nontrivial solutions exist are called *eigenvalues*, and the corresponding solutions are called *eigenfunctions*. The set of eigenvalues is called the *spectrum* of $-\Delta$. It is also of interest to replace the Dirichlet boundary condition (7.3.2) with a Neumann condition on the normal derivative, $du/dn = 0$ on ∂D . Generally, eigenvalues and the corresponding eigenfunctions depend on the boundary condition, the operator, and the domain D .

Example. In addition to quantum theory, the physical origin of eigenvalue problems for PDEs comes out of the study of vibration problems. For example, let D represent an elastic membrane in two dimensions. Under special assumptions, the vertical displacement $u(x, t)$ of the membrane from equilibrium at position $x = (x_1, x_2)$ at time t is governed by the two-dimensional wave equation

$$u_{tt} - a \Delta u = 0, \quad x \in D, \quad t > 0, \quad (7.3.3)$$

where $a > 0$ is a physical constant. If the boundary of the membrane is held fixed (pinned), the boundary condition is

$$u = 0, \quad x \in \partial D, \quad t > 0. \quad (7.3.4)$$

Of interest in many engineering applications are solutions of the form

$$u(x, t) = U(x)e^{i\omega t}. \quad (7.3.5)$$

Such solutions are oscillatory in time with frequency ω . The shape of the solution is $U(x)$. If (7.3.3)–(7.3.4) admit solution of the form (7.3.5), ω is called a *natural frequency* and $U(x)$ is called a *normal mode*. To find such solutions we substitute (7.3.5) into (7.3.3) and (7.3.4) to obtain

$$-\Delta U = \lambda U, \quad x \in D,$$

$$U = 0, \quad x \in \partial D,$$

where $\lambda \equiv \omega^2/a$. Therefore, an eigenvalue problem must be solved to determine the natural frequencies and normal modes of vibrating membranes. \square

A lot is known about the eigenvalues and eigenfunctions for $-\Delta$, the negative Laplacian. For special regions such as rectangles and balls, explicit formulas are known. The classic book by Courant & Hilbert (1953) is one of the best references for properties of eigenvalues of linear partial differential operators. For the sequel we require only a few facts that are collected together in the following theorem on the *principal* (smallest) eigenvalue and its monotonicity. The domain D over which the problem is defined is assumed to be an open, bounded region with a sufficiently smooth boundary.

Theorem. For the eigenvalue problem (7.3.1)–(7.3.2) on the domain D , there are infinitely many eigenvalues λ_n , $n = 1, 2, 3, \dots$ that can be arranged in a sequence $0 < \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \dots$ with $\lim \lambda_n = +\infty$. The eigenfunction $\phi(x)$ associated with the *principal eigenvalue* λ_1 can be chosen to be positive on D . Moreover, if D^* is a subdomain of D , the principal eigenvalue λ_1^* for the eigenvalue problem (7.3.1)–(7.3.2) on D^* has the property that $\lambda_1^* \geq \lambda_1$.

This theorem extends to more general operators and boundary conditions. For example, if L is a uniformly elliptic operator on D , we can consider the eigenvalue problem

$$\begin{aligned} -Lu + c(x)u &= \lambda r(x)u, \quad x \in D, \\ a \frac{du}{dn} + bu &= 0, \quad x \in \partial D, \end{aligned}$$

where $a = 1$ and $b \geq 0$ or $a = 0$ and $b = 1$, and where $c, r \in C(D)$, $c > 0$ and $r > 0$. The principal eigenvalue λ_1 is real, nonnegative, and has multiplicity 1; the corresponding principal eigenfunction can be taken to be positive on D .

In the theorem, the set of eigenfunctions $\phi_n(x)$ form an orthonormal basis for $L_2(D)$, and therefore any function $f \in L_2(D)$ can be expanded in its generalized Fourier series

$$f(x) = \sum f_n \phi_n(x),$$

where f_n are the Fourier coefficients

$$f_n = \int_D f(x) \phi_n(x) dx.$$

This is the fact that makes the separation-of-variables method work for PDEs involving the Laplacian on bounded domains.

7.3.2 Nonlinear Eigenvalue Problems

We begin with an example of a semilinear problem that has its origins in combustion phenomena. Our analysis of this problem follows that of Bebernes and Eberly (1989); see also Pao (1992).

Example. Consider the *solid-fuel model* in several dimensions

$$u_t - \Delta u = \delta e^u, \quad x \in D, \quad t > 0,$$

with initial and boundary conditions

$$\begin{aligned} u(x, 0) &= u_0(x), \quad x \in D, \\ u(x, t) &= 0, \quad x \in \partial D, \quad t > 0. \end{aligned}$$

In the combustion model u is a dimensionless, perturbed temperature variable. In the steady-state case we have the nonlinear boundary value problem

$$-\Delta u = \delta e^u, \quad x \in D, \tag{7.3.6}$$

$$u = 0, \quad x \in \partial D, \tag{7.3.7}$$

which is called the *Gelfand problem*. An important problem in combustion theory is to determine values of the parameter δ , which is an eigenvalue, for which the problem has a positive solution. \square

With the Gelfand problem as motivation, we are led to consider a wide class of nonlinear eigenvalue problems of the form

$$-\Delta u = \lambda f(x, u), \quad x \in D, \tag{7.3.8}$$

$$u = 0, \quad x \in \partial D. \tag{7.3.9}$$

We assume that f is continuous on $D \times \mathbb{R}$ and $f(x, u) \geq 0$. Notice that the linear eigenvalue problem for $-\Delta$ is not a special case of (7.3.8)–(7.3.9) because $f(x, u) = u$ does not satisfy the positivity requirement. Because we are interested in nonnegative solutions, we define the *spectrum* σ as the set of all real numbers λ for which a nonnegative solution of (7.3.8)–(7.3.9) exists. It is easy to show that if there is a positive number λ_1 in σ , then σ must also contain the interval $(0, \lambda_1)$. So for any λ in the interval $0 \leq \lambda \leq \lambda_1$, there is a nonnegative solution to (7.3.8)–(7.3.9). (Clearly, $0 \in \sigma$.) Consequently, the nonlinear problem (7.3.8)–(7.3.9) will not have a discrete spectrum as we have come to expect from linear boundary value problems. We record this as our first result.

Theorem. If $\lambda_1 > 0$ and $\lambda_1 \in \sigma$, then $[0, \lambda_1] \subseteq \sigma$.

The proof uses the fundamental existence theorem on upper and lower solutions for elliptic equations stated in Section 7.2. Let $0 < \lambda < \lambda_1$ be fixed. Then $\underline{u}(x) = 0$ is a lower solution to (7.3.8)–(7.3.9) because f is nonnegative. Now let $\bar{u}(x)$ be the nonnegative solution of (7.3.8)–(7.3.9) corresponding to λ_1 . It is clear that

$$-\Delta\bar{u} = \lambda_1 f(x, \bar{u}) \geq \lambda f(x, \bar{u}), \quad x \in D,$$

and therefore \bar{u} is an upper solution to (7.3.8)–(7.3.9). By the existence theorem there is a solution $u(x)$ to (7.3.8)–(7.3.9) satisfying $0 \leq u(x) \leq \bar{u}(x)$, $x \in D$, thereby completing the proof. \square

Now we prove an important result on the nonexistence of eigenvalues beyond a certain value, provided that f is bounded below by a linear function of u .

Theorem. Suppose that

$$f(x, u) \geq h(x) + r(x)u, \quad (x, u) \in D \times [0, \infty),$$

where h and r are continuous on D , and $h, r > 0$ on D . Then the eigenvalue problem (7.3.8)–(7.3.9) has no nonnegative solutions for any $\lambda \geq \lambda_1(r)$, where $\lambda_1(r)$ is the principal eigenvalue of the problem

$$-\Delta u = \lambda r(x)u, \quad x \in D; \quad u = 0, \quad x \in \partial D. \quad (7.3.10)$$

Again the proof uses the idea of upper and lower solutions. By way of contradiction, assume that $\bar{u}(x)$ is a nonnegative solution of (7.3.8)–(7.3.9) with $\lambda \geq \lambda_1(r)$. Then it is easy to see that \bar{u} is an upper solution of the BVP

$$-\Delta u = \lambda(h(x) + r(x)u), \quad x \in D; \quad u = 0, \quad x \in \partial D. \quad (7.3.11)$$

Moreover, it is obvious that $\underline{u} = 0$ is a lower solution of (7.3.11). By the existence theorem on upper and lower solutions there is a solution $u(x)$ of (7.3.11) with the property that $0 \leq u(x) \leq \bar{u}(x)$ for $x \in D$. Moreover, by the maximum principle we must have $u(x) > 0$ on D [because $-\Delta u > \lambda r(x)u \geq 0$ on D , and $u = 0$ on ∂D].

Now let ϕ be the positive, principal eigenfunction of (7.3.10) corresponding to the principal eigenvalue $\lambda_1(r)$. By Green's identity, we have

$$0 = \int_D (u\Delta\phi - \phi\Delta u) dx = \int_D [\phi(\lambda h + \lambda r u) - \lambda_1(r)r\phi u] dx.$$

Therefore

$$[\lambda_1(r) - \lambda] \int_D r u \phi dx = \lambda \int_D \phi h dx > 0,$$

which implies $\lambda_1(r) > \lambda$, a contradiction. Consequently, no nonnegative solution of (7.3.8)–(7.3.9) can exist for $\lambda \geq \lambda_1(r)$, completing the proof. \square

Example. Consider the Gelfand problem

$$\begin{aligned}-\Delta u &= \lambda e^u, & x \in D, \\ u &= 0, & x \in \partial D.\end{aligned}$$

Here we have $f(x, u) = e^u \geq u + 1$ for $u \geq 0$, so the last theorem applies with $h = r = 1$. We conclude that the Gelfand problem has no nonnegative solutions for any $\lambda \geq \lambda_1$, where λ_1 is the principal eigenvalue of the problem $-\Delta u = \lambda u$, $x \in D$, with $u = 0$ on ∂D . To compute λ_1 see Exercise 3. \square

Of course, the main question is the existence of a positive eigenvalue of (7.3.8)–(7.3.9).

Theorem. If $f(x, u) > 0$ for $x \in D$ and $u \geq 0$, then there exists a real number $\lambda^* > 0$ such that the eigenvalue problem (7.3.8)–(7.3.9) has a positive solution on D for $\lambda = \lambda^*$.

The strong maximum principle shows that a necessary condition for a positive solution is $\lambda \geq 0$. For, let $\lambda < 0$. Then $-\Delta u = \lambda f(x, u) < 0$ on D and $u \leq 0$ on D . By the maximum principle, $u \leq 0$ in D , a contradiction. To complete the proof we need to show that there is a positive solution for some $\lambda^* > 0$. Again the idea of upper and lower solutions is essential. We have already shown that $\underline{u} = 0$ is a lower solution for any $\lambda^* > 0$. Now let \bar{u} be the solution to

$$\begin{aligned}-\Delta u &= 1, & x \in D, \\ u &= 0, & x \in \partial D.\end{aligned}$$

By the maximum principle we have $\bar{u} > 0$ on D . For λ^* sufficiently small

$$1 = -\Delta \bar{u} \geq \lambda^* f(x, \bar{u}), \quad x \in D.$$

Therefore, \bar{u} is an upper solution to (7.3.8)–(7.3.9) for $\lambda = \lambda^*$ (note that f is bounded). By the existence theorem there is a solution $u(x)$ of (7.3.8)–(7.3.9), with $\lambda = \lambda^*$, satisfying the condition $0 \leq u(x) \leq \bar{u}(x)$, $x \in D$. Again by the maximum principle we must have $u(x) > 0$. Note that u cannot be an identically zero solution because $f(x, u) > 0$.

EXERCISES

1. Consider the eigenvalue problem (7.3.8)–(7.3.9), where $f(x, u) \in C^2$, $f(x, 0) > 0$, $f_u(x, 0) > 0$, and $f_{uu}(x, u) \geq 0$, for all $x \in D$ and $u \geq 0$. Given $\lambda > \lambda_1(f_u(x, 0))$, prove that there are no nonnegative solutions.

2. Consider the nonlinear boundary value problem

$$\begin{aligned}\Delta u + \mu u &= u^3, & x \in D, \\ u &= 0, & x \in \partial D.\end{aligned}$$

Prove that if μ is sufficiently small, then $u = 0$ is the only solution. *Hint:* Take $\mu < \lambda_1$, where λ_1 is the principal eigenvalue of $-\Delta$ on D .

3. Let D be the open ball $x^2 + y^2 < 1$ in the plane. Find a value of λ_1 for which the eigenvalue problem

$$-(u_{xx} + u_{yy}) = \lambda e^u, \quad x \in D; \quad u = 0, \quad x \in \partial D$$

has no nonnegative solution for $\lambda \geq \lambda_1$.

4. In two dimensions, find the eigenvalues and eigenfunctions of $-\Delta$ with Dirichlet boundary conditions on the domain $0 \leq x \leq a$, $0 \leq y \leq b$. Answer the same question when the domain is a disk of radius R .

7.4 Stability and Bifurcation

Physical systems governed by nonlinear partial differential equations suggest many interesting mathematical questions. One of these, which we now consider, is the existence of equilibrium solutions and how those solutions depend on parameters that occur in the problem. For example, equilibrium solutions may exist for some values of a parameter, yet not for other values. More interestingly, as a parameter is varied, there may be a critical value where an equilibrium solution may lose its stability properties and no longer exist as a physical possibility, even though it may continue to exist in a mathematical sense. At the same critical value, division may occur and additional equilibria may suddenly appear. The study of problems of this type lies in the domain of stability and bifurcation theory. We first introduce some problems in the realm of ordinary differential equations, after which we study related questions for PDEs.

7.4.1 Ordinary Differential Equations

We motivate the key ideas with an example in ODEs.

Example. Consider an animal population whose growth is determined by a logistics growth law, while at the same time the population is harvested at a

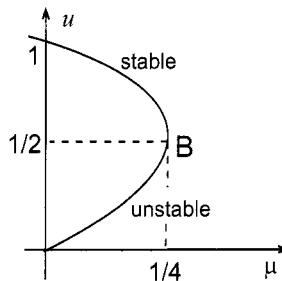


Figure 7.1 Bifurcation diagram showing the locus (7.4.3), a graph of the constant, equilibrium solutions versus the bifurcation parameter μ . The point B is a bifurcation point where the stability changes; the states on the upper branch are stable, and the states on the lower branch are unstable.

constant rate. If $U = U(\tau)$ is the population at time τ , then the governing evolution equation is

$$U' = rU \left(1 - \frac{U}{K}\right) - h, \quad \tau > 0, \quad (7.4.1)$$

where r is the growth rate, K the carrying capacity, and h the constant harvesting rate, with $h > 0$. Equation (7.4.1) can be nondimensionalized by introducing the scaled variables

$$u = \frac{U}{K}, \quad t = r\tau.$$

Then, in dimensionless variables, (7.4.1) becomes

$$\frac{du}{dt} = u(1 - u) - \mu, \quad (7.4.2)$$

where μ is dimensionless harvesting parameter defined by $\mu = h/rK$. Now, treating μ as control parameter, we ask whether equilibrium populations exist; that is, whether there are any constant solutions of (7.4.2). These are easily found by setting $du/dt = 0$. In the present cases, the equilibrium populations are values of u that satisfy

$$u = \frac{1 \pm \sqrt{1 - 4\mu}}{2}. \quad (7.4.3)$$

Therefore, if the harvesting parameter μ exceeds $\frac{1}{4}$, there are no equilibrium populations; if $0 < \mu < \frac{1}{4}$, there are two possible equilibrium states. We sketch the locus (7.4.3) on a μ - u -coordinate system to indicate the dependence of equilibrium solutions on the parameter μ . Such a graph is shown in Figure 7.1 and is called a *bifurcation diagram*; the harvesting parameter μ is called the

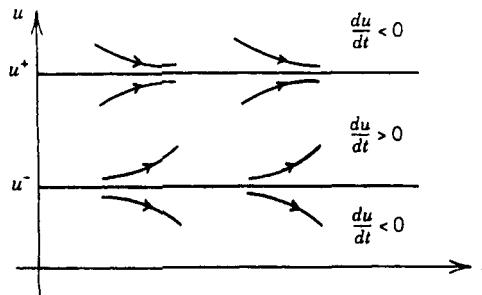


Figure 7.2 Diagram showing the direction of the integral curves of (7.4.2).

bifurcation parameter. At the critical value $\mu = \frac{1}{4}$, where equilibrium solutions appear, the point on the bifurcation diagram is called a *bifurcation point*. We can physically interpret the result in the following manner. If μ is too large, that is, if there is too much harvesting, then (7.4.2) shows that du/dt is negative the population dies out. As the harvesting is decreased, there is a critical value of μ below which two sizes of equilibrium populations can exist; in these two populations there is a balance between growth and harvesting. We can now ask whether nature prefers one of those equilibrium populations over the other. This question is at the heart of the concept of stability. If the system is in one of these equilibrium states and a small perturbation is imposed, say, by introducing a few more animals, does the system go out of balance, or does it return to the original equilibrium? This question can be resolved easily by examining the direction field of the differential equation (7.4.2) in the case $0 < \mu < \frac{1}{4}$. Letting u^+ and u^- denote the two equilibrium populations defined in (7.4.3), with the plus and minus signs, respectively, we observe that

$$\begin{aligned}\frac{du}{dt} &< 0 & \text{if } u > u^+, \\ \frac{du}{dt} &> 0 & \text{if } u^- < u < u^+, \\ \frac{du}{dt} &< 0 & \text{if } 0 < u < u^-. \end{aligned}$$

Therefore, the solutions of (7.4.2) must behave as in Figure 7.2. So u^+ must be an attractor, or a stable equilibrium, and u^- must be a repeller, or unstable equilibrium. Therefore, nature prefers the larger equilibrium population. Therefore the upper branch of the equilibrium curve shown on the bifurcation diagram in Figure 7.1 is stable, while the lower branch is unstable. \square

The types of questions addressed above form the basis of bifurcation and stability theory. Many problems in the physical and natural sciences have these

ingredients, and it is of interest to determine how equilibrium solutions depend on parameters occurring in the problem, and the stability of these equilibrium states.

This example leads us to consider the general ordinary differential equation

$$\frac{du}{dt} = f(\mu, u), \quad (7.4.4)$$

where f is a given function of u and a bifurcation parameter μ . Equilibrium solutions u of (7.4.4) satisfy the relation

$$f(\mu, u) = 0. \quad (7.4.5)$$

A graph of the locus (7.4.5) in a μu plane is called a *bifurcation diagram*. Generally, the locus may be quite complicated, with many branches and several intersections. A heuristic argument determines if a given point (μ, u_e) on the locus represents a stable or unstable equilibrium. Let $u(t) = u_e + w(t)$, where $w(t)$ represents a small perturbation from the constant equilibrium state u_e . Note that u_e depends on μ , but to keep the notation uncluttered we suppress this dependence in the symbolism. Substituting into (7.4.4) and expanding in a Taylor series give

$$\begin{aligned} \frac{dw}{dt} &= f(\mu, u_e + w) \\ &= f(\mu, u_e) + f_u(\mu, u_e)w + O(w^2) \\ &= f_u(\mu, u_e)w + O(w^2). \end{aligned}$$

Ignoring the higher-order term $O(w^2)$, the perturbation w satisfies the linearized equation

$$\frac{dw}{dt} = aw, \quad a \equiv f_u(\mu, u_e),$$

which has solution

$$w(t) = \text{const} \cdot e^{at}.$$

Therefore, if $a > 0$, the perturbation w grows and u_e is unstable; if $a < 0$ the perturbation w decays to zero and u_e is stable. This argument contains the seeds of a rigorous proof that the sign of the quantity $f_u(\mu, u_e)$ determines the nature of the stability of an equilibrium state u_e for fairly general conditions on the function f . We refer the reader to texts on ODEs for a precise formulation and proof of a general result.

Example. In equation (7.4.2), $f(\mu, u) = u(1 - u) - \mu$. Therefore, $f_u(\mu, u) = 1 - 2u$. If u is an equilibrium solution with $u < \frac{1}{2}$, then $f_u > 0$ and u is unstable; if $u > \frac{1}{2}$, then $f_u < 0$ and u is stable. This is consistent with the previous conclusion that any solution on the upper branch of the locus shown

in Figure 7.1 is stable, while any equilibrium state on the lower branch is unstable. \square

These ideas easily extend to systems of ordinary differential equations and are discussed in most texts in differential equations.

7.4.2 Partial Differential Equations

The archetypical reaction–diffusion equation is the semilinear equation

$$u_t - \Delta u = f(\mu, u), \quad x \in D, \quad t > 0, \quad (7.4.6)$$

with boundary and initial conditions given by

$$u(x, t) = 0, \quad t > 0, \quad x \in \partial D; \quad u(x, 0) = u_0(x), \quad x \in D, \quad (7.4.7)$$

where f is a given function and μ is a real parameter. Equilibrium solutions are defined by solutions of the elliptic problem

$$-\Delta u = f(\mu, u), \quad x \in D, \quad (7.4.8)$$

$$u = 0, \quad x \in \partial D. \quad (7.4.9)$$

From the discussion in Section 7.3, where we considered eigenvalue problems and f had the form $f(\mu, u) = \mu g(u)$, it is clear that this elliptic problem may have solutions for some values of μ and yet no solution for other values of μ . Suppose that there is an equilibrium solution $u_e(x)$ of (7.4.8)–(7.4.9) for some value of the parameter μ . In what sense is this equilibrium stable or unstable? There are several definitions of stability for reaction diffusion equations, but here we focus on one version of Liapunov stability. The fundamental question is whether a time-dependent solution of the reaction diffusion problem (7.4.6)–(7.4.7) converges to an equilibrium solution of (7.4.8)–(7.4.9) as t approaches infinity. Whether or not we get convergence depends on the set of initial functions $u_0(x)$ in the initial condition in (7.4.7). Evidently, if we start at $t = 0$ too far away from the equilibrium solution, then we may not obtain convergence. Thus stability of u_e may be local, requiring that u_0 be sufficiently close to u_e . We formulate the following definition, which is similar to the definition in ordinary differential equations.

Definition. An equilibrium solution $u_e(x)$ of (7.4.8)–(7.4.9) is *stable* if for any $\varepsilon > 0$ there is a $\delta > 0$ for which

$$|u_0(x) - u_e(x)| < \delta \quad \text{for } x \in D$$

implies that

$$|u(x, t) - u_e(x)| < \varepsilon \quad \text{for } x \in D, t > 0.$$

In addition, if

$$\lim_{t \rightarrow \infty} |u(x, t) - u_e(x)| = 0 \quad \text{for } x \in D,$$

then u_e is *asymptotically stable*. If u_e is not stable, it is *unstable*.

The set of initial conditions u_0 for which u_e is stable is called the *domain of attraction* of u_e . If the domain of attraction contains all initial functions, then u_e is called *globally stable*, or a *global attractor*.

Example. Consider the elliptic problem

$$-\Delta u = \mu u - u^3, \quad x \in D, \tag{7.4.10}$$

$$u = 0, \quad x \in \partial D, \tag{7.4.11}$$

which we regard as a time-asymptotic limit of a reaction-diffusion system. It is natural to ask whether this problem has solutions. We shall show that if $0 < \mu < \lambda_1$, where λ_1 is the principal eigenvalue of the operator $-\Delta$ on D , no nontrivial solutions exist. By way of contradiction, assume that a nontrivial solution u of (7.4.10)–(7.4.11) exists. Then there is a domain $D' \subseteq D$ such that u satisfies the boundary value problem

$$-\Delta u = \mu u - u^3, \quad x \in D', \quad u = 0, \quad x \in \partial D', \tag{7.4.12}$$

and u is of one sign on D' . Let λ'_1 be the principal eigenvalue of $-\Delta$ on D' with its corresponding eigenfunction $\phi(x)$ strictly positive on D' . Thus

$$-\Delta \phi = \lambda'_1 \phi, \quad x \in D'; \quad \phi = 0, \quad x \in \partial D'. \tag{7.4.13}$$

Note that $\lambda'_1 \geq \lambda_1$ because $D' \subseteq D$. Using Green's identity and equations (7.4.12) and (7.4.13), we have

$$\begin{aligned} 0 &= \int_{D'} (u \Delta \phi - \phi \Delta u) dx \\ &= \int_{D'} [-u \lambda'_1 \phi + \phi (\mu u - u^3)] dx \\ &= \int_{D'} u \phi (\mu - \lambda'_1 - u^2) dx. \end{aligned}$$

But the last integral is strictly of one sign because u is of one sign, ϕ is positive, and $\mu < \lambda_1 \leq \lambda'_1$. This a contradiction, and no nontrivial solution can exist.

□

Example. Next consider the reaction-diffusion problem

$$u_t - \Delta u = \mu u - u^3, \quad x \in D, \quad t > 0, \quad (7.4.14)$$

$$u(x, 0) = u_0(x), \quad x \in D, \quad (7.4.15)$$

$$u(x, t) = 0, \quad x \in \partial D, \quad t > 0, \quad (7.4.16)$$

where u_0 is a bounded, continuous function. From an earlier example this problem has only one equilibrium solution, namely, $u_e = 0$, for $0 < \mu < \lambda_1$, where λ_1 is the principal eigenvalue of the negative Laplacian on D . We show that $u_e = 0$ is asymptotically stable in the sense that there exist positive constants a and C for which

$$|u(x, t)| \leq C e^{-at} \sup_{x \in D} |u_0(x)| \quad (7.4.17)$$

for all $t > 0$ and $x \in D$. To begin, pick a region $D' \supseteq D$ such that the principal eigenvalue λ'_1 of $-\Delta$ on D' satisfies the inequality $\mu < \lambda'_1 < \lambda_1$. Choose the eigenfunction $\psi(x)$ corresponding to λ'_1 to be positive and normalized to satisfy the condition $\sup_{D'} |\psi(x)| = 1$. Then

$$-\Delta \psi = \lambda'_1 \psi, \quad x \in D' \quad \text{with} \quad \psi = 0, \quad x \in \partial D'. \quad (7.4.18)$$

Now set

$$u(x, t) = w(x, t)\psi(x)e^{-at}$$

for $x \in D$, $t > 0$, where a is to be selected later. Note that ψ is positive on D . From (7.4.14) and (7.4.18) we can show that w satisfies the PDE

$$\Delta w + \frac{2\operatorname{grad} \psi \cdot \operatorname{grad} w}{\psi} + (a + \mu - \lambda'_1 - w^2 \psi^2 e^{-2at})w - w_t = 0, \quad (7.4.19)$$

for $x \in D$ and $t > 0$. This equation has the form

$$\Delta w + \sum b_j(x)w_{x_j} + cw - w_t = 0,$$

where

$$c = a + \mu - \lambda'_1 - w^2 \psi^2 e^{-2at}.$$

Equation (7.4.19) has the form for which the maximum principle for parabolic operators applies, provided that $c < 0$ (see Section 6.4). But we can ensure that $c < 0$ by choosing a sufficiently small, because $\mu < \lambda'_1$. The auxiliary conditions on w are

$$w(x, t) = 0, \quad x \in \partial D, \quad t > 0; \quad w(x, 0) = \frac{u_0(x)}{\psi(x)}, \quad x \in D.$$

Therefore, by the maximum principle,

$$|w(x, t)| \leq \sup_{x \in D} \left| \frac{u_0(x)}{\psi(x)} \right|, \quad x \in D, \quad t > 0. \quad (7.4.20)$$

Selecting the constant C to be the quantity on the right side of (7.4.20) then gives the result (7.4.17), showing that the zero equilibrium solution of (7.4.14)–(7.4.16) is asymptotically stable, provided that $\mu < \lambda_1$, for arbitrary initial conditions u_0 . \square

We are left to wonder about what happens in the reaction–diffusion system (7.4.14)–(7.4.16) when the parameter μ exceeds λ_1 . Clearly, $u = 0$ is still an equilibrium solution. The following example shows that two additional equilibria appear. The method relies on the fundamental existence theorem on upper and lower solutions stated in Section 7.2.

Example. Consider (7.4.10)–(7.4.11) with $\mu > \lambda_1$, where λ_1 is the principal eigenvalue of $-\Delta$ on domain D . We show that there exist at last two nontrivial solutions, one positive and one negative. Let $u = \alpha\phi$, where α is a constant to be selected later and ϕ is the positive eigenfunction associated with the eigenvalue λ_1 . Then

$$-\Delta\phi = \lambda_1\phi, \quad x \in D; \quad \phi = 0, \quad x \in \partial D.$$

Therefore,

$$\begin{aligned} -\Delta u - \mu u + u^3 &= -\alpha\Delta\phi - \mu\alpha\phi + \alpha^3\phi^3, \\ &= \alpha\phi(\lambda_1 - \mu + \alpha^2\phi^2). \end{aligned}$$

Now choose $|\alpha|$ sufficiently small to force $\lambda_1 - \mu + \alpha^2\phi^2 < 0$. Consequently, if $\alpha > 0$, then $\alpha\phi$ is a positive lower solution, and if $\alpha < 0$, then $\alpha\phi$ is a negative upper solution of (7.4.10)–(7.4.11). It remains to determine a positive upper solution and a negative lower solution. Consider a domain D' containing D , and let $\psi(x)$ be the positive eigenfunction of the principal eigenvalue λ'_1 of the negative Laplacian on D' . That is,

$$-\Delta\psi = \lambda'_1\psi, \quad x \in D'; \quad \psi = 0, \quad x \in \partial D'.$$

We then have $\lambda'_1 \leq \lambda_1 < \mu$. Now take $u = \beta\psi$, which yields

$$-\Delta u - \mu u + u^3 = \beta\psi(\lambda'_1 - \mu + \beta^2\psi^2),$$

and consider this equation on D , where $\psi > 0$. If $|\beta|$ is chosen large enough and $\beta > 0$, then $\beta\psi$ is a positive upper solution that exceeds $\alpha\phi$, $\alpha > 0$. Similarly, if $|\beta|$ is chosen large enough and $\beta < 0$, then $\beta\psi$ is a negative lower solution that

does not exceed $\alpha\phi$, $\alpha < 0$. We conclude that there must be a positive solution of (7.4.10)–(7.4.11) and a negative solution of (7.4.10)–(7.4.11). \square

We regard (7.4.10)–(7.4.11) as a bifurcation problem in the following sense. As the parameter μ increases from small positive values, there is a critical value $\mu = \lambda_1$ where the equilibrium solution $u = 0$ branches, or bifurcates, into three equilibrium solutions ($u = 0$, and the positive and negative solution). We represent this phenomenon schematically on a bifurcation diagram as in Figure 7.3. This type of bifurcation is called a *pitchfork bifurcation* because of the shape of the locus. It can be shown that for $\mu > \lambda_1$ there is only one positive and one negative solution, so there are no additional branches to the diagram. See, for example, Sattinger (1973), from where these examples are adapted. Stakgold (1971) contains an elementary discussion of these ideas.

Example. We address the question of stability of the zero solution in the equation above, in one spatial dimension. Another standard method, called *linearized stability analysis*, is applied. The question is: Do small perturbations at the initial time, near the zero solution distribution, decay to zero or grow without bound as t gets large? Let us consider the problem

$$u_t - u_{xx} = \mu u - u^3, \quad 0 < x < \pi, \quad t > 0, \quad (7.4.21)$$

$$u(0, t) = u(\pi, t) = 0, \quad t > 0, \quad (7.4.22)$$

$$u(x, 0) = \varepsilon u_0(x), \quad 0 < x < \pi. \quad (7.4.23)$$

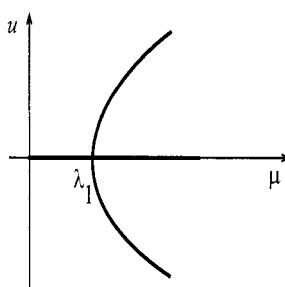


Figure 7.3 Bifurcation diagram indicating the existence of an equilibrium solution for values of the parameter μ . For $\mu < \lambda_1$, only a trivial solution exists; for $\mu > \lambda_1$, there are three equilibrium solutions, two nontrivial. The vertical axis could represent any feature of an equilibrium solution, for example, its maximum value (upper branch) or its minimum value (lower branch).

We note that $u = 0$ is an asymptotic solution to the problem for all $\mu > 0$. We introduced an initial condition of the form (7.4.23), where ε is a small parameter, to study the evolution of a small initial perturbation in the problem. From prior results the zero solution is stable for $\mu < \lambda_1$, where λ_1 is the principal eigenvalue of $-d^2/dx^2$ on the interval $0 < x < \pi$. It is easy to determine the eigenvalues and eigenfunctions of the problem

$$-u''(x) = \lambda u(x), \quad 0 < x < \pi; \quad u(0) = u(\pi) = 0.$$

They are

$$\lambda_n = n^2, \quad u_n(x) = \sin nx, \quad n = 1, 2, 3, \dots$$

Therefore, the principal eigenvalue is $\lambda_1 = 1$. The standard stability argument proceeds by assuming a solution of (7.4.21)–(7.4.23) of the form $u(x, t) = \varepsilon v(x, t)$, where v is an order 1 function. One can regard this form as the zero solution plus a small perturbation represented by εv , $\varepsilon \ll 1$. If this form of u is substituted into the problem (7.4.21)–(7.4.23), then

$$v_t - v_{xx} = \mu v - \varepsilon^2 v^3, \quad 0 < x < \pi, t > 0, \quad (7.4.24)$$

$$v(0, t) = v(\pi, t) = 0, \quad t > 0, \quad (7.4.25)$$

$$v(x, 0) = u_0(x), \quad 0 < x < \pi. \quad (7.4.26)$$

Because we are unable to solve this nonlinear equation exactly, we use an approximation and carry out a linear stability analysis. That is, we assume that the nonlinear term in (7.4.24) is small and negligible, which gives the linearized equation

$$v_t - v_{xx} = \mu v. \quad (7.4.27)$$

If v remains bounded, then the assumption that εv is small remains valid and it is consistent to drop the nonlinear term. In this case we say that the zero solution is (linearly) stable to small perturbations. If v grows without bound, the assumption that the nonlinear term in (7.4.24) is small, with the subsequent linearization, is invalid. In this case the zero solution is (linearly) unstable to small perturbations. In the latter case the full nonlinear equations must be analyzed to predict the correct evolutionary behavior of the small perturbation.

Equation (7.4.27), along with the boundary and initial conditions (7.4.25)–(7.4.26), can be solved by the method of separation of variables. Let $v(x, t) = T(t)X(x)$, and substitute into (7.4.27) and the boundary conditions (7.4.25). In the usual way the PDE separates and we obtain a boundary value problem (a Sturm–Liouville problem) for X , namely

$$X'' + \lambda X = 0, \quad 0 < x < \pi, \quad X(0) = X(\pi) = 0, \quad (7.4.28)$$

and an equation for T , namely

$$T' = (\mu - \lambda)T. \quad (7.4.29)$$

Here λ is the separation constant. The boundary value problem (7.4.28) has eigenvalues and eigenfunctions given by

$$\lambda_n = n^2, \quad X_n(x) = \sin nx, \quad n = 1, 2, 3, \dots,$$

and the T equation has solutions

$$T_n(t) = e^{(\mu-n^2)t}, \quad n = 1, 2, 3, \dots.$$

Therefore, we have determined infinitely many solutions of the PDE (7.4.27) and the boundary conditions (7.4.25) of the form

$$v_n(x, t) = T_n(t)X_n(x) = e^{(\mu-n^2)t} \sin nx, \quad n = 1, 2, 3, \dots.$$

These are the Fourier modes, which we can superimpose to find a linear combination that meets the initial condition. Hence, we take

$$v(x, t) = \sum_n a_n e^{(\mu-n^2)t} \sin nx. \quad (7.4.30)$$

Applying the initial condition (7.4.26) yields

$$u_0(x) = \sum_n a_n \sin nx,$$

which is the Fourier sine series for $u_0(x)$. The Fourier coefficients are

$$a_n = \frac{2}{\pi} \int_0^\pi u_0(x) \sin nx dx. \quad (7.4.31)$$

Consequently, the solution of the boundary value problem (7.4.25)–(7.4.27) is given by (7.4.30), with the a_n given by (7.4.31). In summary,

$$v(x, t) = a_1 e^{(\mu-1)t} \sin x + a_2 e^{(\mu-4)t} \sin 2x + a_3 e^{(\mu-9)t} \sin 3x + \dots. \quad (7.4.32)$$

It is clear from (7.4.32) that if $\mu < 1$, then all the exponential terms decay and $v(x, t)$ approaches zero for large t . Thus, the zero solution is asymptotically stable to small perturbations. If $\mu > 1$, at least one of the terms (Fourier modes) in (7.4.32) grows exponentially and the perturbation grows without bound; the zero solution is (linearly) unstable. \square

The stability analysis in this example was based on linearization. In the case $\mu > 1$ the linearized equation (7.4.27) is no longer an accurate approximation of (7.4.24) because, as the solution (7.4.32) shows, the assumption that v^3 is small is no longer valid. Consequently, the solution (7.4.32) is no longer an accurate approximation of the solution to the nonlinear problem in this case, and we have inherent inconsistencies. Therefore, let us see if we can deal with the nonlinear problem directly to get more information in the unstable regime.

Example. Consider the nonlinear boundary value problem (7.4.24) with auxiliary conditions (7.4.25) and (7.4.26), where the bifurcation parameter μ is just slightly larger than unity. We introduce another standard technique by performing a normal-mode analysis for the nonlinear equation. We assume that the boundary value problem (7.4.24)–(7.4.26) has a solution of the form

$$v(x, t) = \sum_n y_n(t) \sin nx, \quad (7.4.33)$$

where the $y_n(t)$ are to be determined. We are motivated by the linearized problem and the fact that the eigenfunctions of the spatial problem are $\sin nx$. By Fourier analysis we know that the $y_n(t)$ are given by the Fourier coefficients

$$y_n(t) = \frac{2}{\pi} \int_0^\pi v(x, t) \sin nx dx. \quad (7.4.34)$$

Multiplying equation (7.4.24) by $\sin nx$ and integrating from 0 to π gives

$$\begin{aligned} & \int_0^\pi v_t \sin nx dx - \int_0^\pi v_{xx} \sin nx dx \\ &= \mu \int_0^\pi v \sin nx dx - \varepsilon^2 \int_0^\pi v^3 \sin nx dx. \end{aligned} \quad (7.4.35)$$

Then, using (7.4.33), we get

$$\begin{aligned} & \int_0^\pi \left(\sum_m y'_m \sin mx \right) \sin nx dx + \int_0^\pi \left(\sum_m m^2 y_m \sin mx \right) \sin nx dx \\ &= \mu \int_0^\pi \sum_m (y_m \sin mx) \sin nx dx - \varepsilon^2 \int_0^\pi v^3 \sin nx dx. \end{aligned} \quad (7.4.36)$$

Using the fact that the eigenfunctions $\sin nx$ are orthogonal on $[0, \pi]$, that is,

$$\int_0^\pi \sin mx \sin nx dx = \begin{cases} 0, & \text{if } n \neq m \\ \frac{\pi}{2}, & \text{if } n = m \end{cases}$$

equation (7.4.36) can be written as

$$y'_n + (n^2 - \mu)y_n = -\frac{2\varepsilon^2}{\pi} \int_0^\pi v^3 \sin nx dx, \quad n = 1, 2, 3, \dots \quad (7.4.37)$$

Now let us manipulate the nonlinear term on the right side of (7.4.37) by substituting for v . We have

$$\begin{aligned}\int_0^\pi v^3 \sin nx dx &= \int_0^\pi \left(\sum_k y_k \sin kx \right)^3 \sin nx dx \\ &= \int_0^\pi \left(\sum_{j,k,m} y_j y_k y_m \sin jx \sin kx \sin mx \right) \sin nx dx \\ &= \sum_{j,k,m} a_{jkmn} y_j y_k y_m,\end{aligned}$$

where

$$a_{jkmn} = \int_0^\pi \sin jx \sin kx \sin mx \sin nx dx.$$

Therefore, (7.4.37) becomes

$$y'_n + (n^2 - \mu)y_n = -\frac{2\varepsilon^2}{\pi} \sum a_{jkmn} y_j y_k y_m, \quad n = 1, 2, 3, \dots, \quad (7.4.38)$$

and we have derived a coupled system of ODEs for the coefficients $y_n(t)$. This system, because of the nonlinearity, is poorly coupled because each equation contains all of the y_n . Initial conditions supplement (7.4.38) in the form

$$y_n(0) = \frac{2}{\pi} \int_0^\pi u_0(x) \sin nx dx, \quad (7.4.39)$$

which comes from (7.4.34) and (7.4.36). We cannot solve (7.4.38)–(7.4.39), but we can advantageously apply the assumption that μ is slightly larger than 1 to obtain the qualitative behavior of the solution in the asymptotic limit of large t . First we remark that if the terms of order ε^2 are neglected in (7.4.38), then the equations decouple and the only growing solution is $y_1(t)$; the remaining y_n , for $n = 2, 3, \dots$, still decay because $n^2 - \mu > 0$. Therefore, for a first approximation of the solution we take

$$y_n(t) = y_n(0)e^{-(n^2-\mu)t} \quad \text{for } n = 2, 3, \dots$$

Then, for $n = 1$, in equation (7.4.38) we retain only the term involving y_1 on the right side to obtain

$$y'_1 + (1 - \mu)y_1 = -\frac{2\varepsilon^2 a_{1111}}{\pi} y_1^3 = -\frac{3\varepsilon^2}{4} y_1^3, \quad (7.4.40)$$

where we have used the fact that $a_{1111} = \int_0^\pi \sin^4 x dx = 3\pi/8$. We recognize (7.4.40) as a Bernoulli equation, which can be reduced to a linear equation

via the transformation $y_1 = w^{-1/2}$. A routine calculation shows that w satisfies

$$w' - 2(1 - \mu)w = \frac{3\varepsilon^2}{2},$$

which has solution

$$w(t) = Ce^{2(1-\mu)t} - \frac{3\varepsilon^2}{4(1-\mu)}.$$

Thus y_1 is given to leading order by

$$y_1(t) = \frac{1}{\sqrt{w(t)}}, \quad (7.4.41)$$

and the constant C can be determined from the initial condition to be

$$C = \frac{3\varepsilon^2}{4(1-\mu)} + y_1^{-2}(0).$$

For large t we observe that $w(t) \rightarrow 3\varepsilon^2/[4(\mu - 1)]$, and therefore

$$y_1(t) \sim \sqrt{\frac{4(\mu - 1)}{3\varepsilon^2}}, \quad t \gg 1. \quad (7.4.42)$$

Consequently,

$$u(x, t) \sim \sqrt{\frac{4(\mu - 1)}{3}} \sin x, \quad t \gg 1 \quad (7.4.43)$$

as the leading order approximation to (7.4.21)–(7.4.23) for large t . In conclusion, the linearized analysis in the foregoing example predicted, incorrectly, that $u(x, t)$ grows exponentially without bound if $\mu > 1$; the refined analysis shows, in fact, that u approaches a nonzero steady state given by (7.4.43) in the case when μ is slightly larger than unity. \square

This nonlinear, normal-mode stability analysis applies to many types of nonlinear problems. The general issue of stability of equilibrium solutions to reaction-diffusion equations is an intriguing and well-developed area of study, and additional sources are given in the references.

Example. (Critical Domain Size) We end the section with an example from ecology on the critical patch size for animals experiencing both growth and diffusion. Does growth dominate, or does diffusion cause the population to go extinct, and how does it depend on the size of the patch? We model the problem with Fisher's equation and no-flux boundary conditions. We have

$$u_t - Du_{xx} = ru \left(1 - \frac{u}{K}\right), \quad (7.4.44)$$

$$u \left(\pm \frac{L}{2}, t\right) = 0 \quad (7.4.45)$$

Let us reduce to dimensionless form by scaling time by $1/r$, distance by $L/2$, and population density by K . Then we have the problem

$$\begin{aligned} u_t - \delta u_{xx} &= u(1-u), \\ u(\pm 1, t) &= 0 \end{aligned}$$

where

$$\delta = \frac{4D}{rL^2}. \quad (7.4.46)$$

The steady state $u = u(x)$ satisfies

$$\begin{aligned} \delta u'' + u(1-u) &= 0, \\ u(\pm 1) &= 0. \end{aligned}$$

We now go to the phase plane

$$\begin{aligned} u' &= v, \\ v' &= -\frac{1}{\delta}u(1-u). \end{aligned}$$

We can get integral curves by dividing the two equations to obtain

$$\frac{dv}{du} = -\frac{u(1-u)}{\delta v},$$

which separates to

$$v dv = -\frac{1}{\delta}u(1-u) du.$$

It is convenient to introduce the function

$$F(u) = \frac{1}{2}u^2 - \frac{1}{3}u^3,$$

which is the antiderivative of $u(1-u)$. Then the first integrals, or orbits, of the system are

$$\frac{1}{2}v^2 = -\frac{1}{\delta}F(u) + C, \quad (7.4.47)$$

where K is an arbitrary constant.

Before evaluating the constant C , let us picture the phase plane. The critical points are $(0, 0)$ and $(1, 0)$ and the Jacobian is

$$J(u, v) = \begin{pmatrix} 0 & 1 \\ -(1/\delta)(1-2u) & 0 \end{pmatrix}.$$

The matrix $J(1, 0)$ has two real eigenvalues of opposite sign, and so $(1, 0)$ is a saddle point. The eigenvalues of $J(0, 0)$ are purely imaginary, and so the linearized system has a center at the origin. By symmetry, the nonlinear system has a center as well. Figure 7.4 shows the form of the phase diagram. It is clear

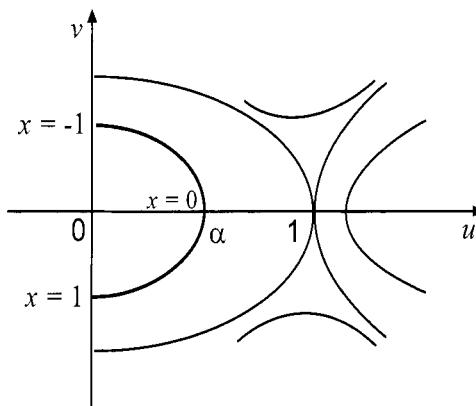


Figure 7.4 The uv phase plane showing a center at $(0, 0)$ and a saddle at $(1, 0)$. The value α , occurring at $x = 0$ is the maximum value of the solution curve.

that if there is a positive equilibrium solution $u = u(x)$ to the boundary value problem, then it is symmetric about the origin $x = 0$, and its maximum value is $u(0) = \alpha < 1$. The figure shows what the solution must look like in the phase plane. It still must be shown that there is an orbit that satisfies the zero boundary conditions when $x = \pm 1$.

We can evaluate the constant C in (7.4.47) by noting that $u = \alpha$ at $v = 0$. Thus $C = (1/\delta)F(\alpha)$, and the orbits are given by

$$v^2 = \frac{2}{\delta}(F(\alpha) - F(u)).$$

Now we go to u and x coordinates and rewrite the last expression as

$$\frac{du}{dx} = \pm \sqrt{\frac{2}{\delta}} \sqrt{F(\alpha) - F(u)}.$$

Integrating over the bottom half of the curve, from $x = 0$ to $x = 1$, and applying the fact that the derivative is negative, we get

$$\int_{\alpha}^0 \frac{du}{\sqrt{F(\alpha) - F(u)}} = -\sqrt{\frac{2}{\delta}}.$$

In the integral on the left we make the substitution $w = u/\alpha$, and we replace δ by (7.4.46) to obtain

$$L = \Psi(\alpha) \equiv \sqrt{\frac{2D}{r}} \int_0^1 \frac{\alpha dw}{\sqrt{F(\alpha) - F(\alpha w)}}. \quad (7.4.48)$$

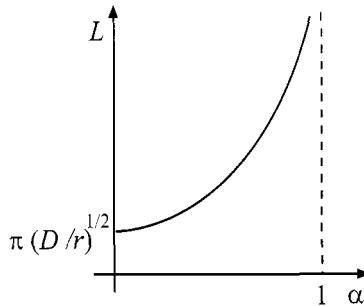


Figure 7.5 Plot showing the domain size L versus the α , the maximum value of the equilibrium solution curve. The curve crosses the L axis at $\pi\sqrt{D/r}$.

The integral function $\Psi(\alpha)$ defines a mapping that associates with each α (the maximum of u) a value of the domain size L . It has the following properties:

1. $\Psi(\alpha)$ is positive on $(0, 1)$ and $\Psi'(\alpha) > 0$.
2. $\Psi(\alpha) \rightarrow \infty$ as $\alpha \rightarrow 1$.
3. $\Psi(\alpha) \rightarrow \pi\sqrt{D/r}$ as $\alpha \rightarrow 0$.

The proof of property 1 is left as an exercise. Property 2 follows from the fact that the the improper integral (7.4.48) diverges in the limit as $\alpha \rightarrow 1$. Property 3 can be verified by first noting

$$F(\alpha) - F(\alpha w) = \frac{1}{2}\alpha^2 (1 - w^2 - \frac{2}{3}\alpha(1 - w^3)),$$

and thus, for small α , using the following binomial expansion:

$$\begin{aligned} \frac{\alpha}{\sqrt{F(\alpha) - F(\alpha w)}} &= \frac{\alpha}{(1/\sqrt{2})\alpha} \left(1 - w^2 - \frac{2}{3}\alpha(1 - w^3)\right)^{(-1/2)} \\ &= \sqrt{2}(1 - w^2)^{-1/2} + O(\alpha). \end{aligned}$$

Therefore, as $\alpha \rightarrow 0$, we obtain

$$\Psi(\alpha) \rightarrow \sqrt{\frac{2D}{r}}\sqrt{2} \int_0^1 (1 - w^2)^{-1/2} dw = 2\sqrt{\frac{D}{r}}\frac{\pi}{2} = \pi\sqrt{\frac{D}{r}}.$$

Figure 7.5 shows a plot of $L = \Psi(\alpha)$. For any $L > \pi\sqrt{D/r}$ there is a unique value of the maximum α for which the problem has a solution. If $L < \pi\sqrt{D/r}$ then there is no solution. \square

Example. (*Stability of the Equilibrium*) Let us denote the nonzero equilibrium found in the last example by $u = u_0(x)$. To investigate its stability we let $\phi(x, t)$

be a small perturbation and take $u = u_0(x) + \phi(x, t)$. Substituting into the PDE and boundary condition (7.4.44)–(7.4.45) gives, after ignoring the nonlinear terms,

$$\begin{aligned}\phi_t &= D\phi_{xx} + r \left(1 - 2\frac{u_0}{K}\right) \phi, \quad -\frac{L}{2} < x < \frac{L}{2}, \quad t > 0, \\ \phi\left(-\frac{L}{2}, t\right) &= 0, \quad \phi\left(\frac{L}{2}, t\right) = 0, \quad t > 0.\end{aligned}$$

We can solve this linear PDE by the method of separation of variables. Let $\phi(x, t) = T(t)y(x)$ and substitute into the PDE to obtain

$$\frac{T'}{T} = \frac{Dy'' + r(1 - 2u_0/K)y}{y} = -\lambda,$$

where $-\lambda$ is the separation constant. Then we get $T(t) = Ce^{-\lambda t}$ and a Sturm–Liouville problem for the spatial part y :

$$Dy'' + r \left(1 - 2\frac{u_0}{K}\right) y = -\lambda y, \quad y\left(-\frac{L}{2}\right) = 0, \quad y\left(\frac{L}{2}\right) = 0. \quad (7.4.49)$$

Our goal is to show that the eigenvalues λ of (7.4.49) are strictly positive, which will prove that perturbations decay over time and the steady solution is stable. Fix λ to be the smallest eigenvalue and $y(x)$ a corresponding eigenfunction. From Sturm–Liouville theory we know that y is of one sign on the interval $-L/2 < x < L/2$.

The demonstration proceeds much like an energy argument. We multiply the differential equation in (7.4.49) by u_0 and integrate over the interval to obtain

$$D \int_{-L/2}^{L/2} u_0 y'' dx + r \int_{-L/2}^{L/2} y u_0 \left(1 - 2\frac{u_0}{K}\right) dx = -\lambda \int_{-L/2}^{L/2} u_0 y dx.$$

Integrating the first term by parts and using the fact that u_0 vanishes on the boundary gives

$$-D \int_{-L/2}^{L/2} u'_0 y' dx + r \int_{-L/2}^{L/2} y u_0 \left(1 - 2\frac{u_0}{K}\right) dx = -\lambda \int_{-L/2}^{L/2} u_0 y dx. \quad (7.4.50)$$

Now, the steady-steady solution u_0 satisfies the equilibrium equation

$$Du_0'' + ru_0 \left(1 - \frac{u_0}{K}\right) = 0.$$

We multiply this equation by y and integrate the same way as above to obtain

$$-D \int_{-L/2}^{L/2} u'_0 y' dx + r \int_{-L/2}^{L/2} y u_0 \left(1 - \frac{u_0}{K}\right) dx = 0. \quad (7.4.51)$$

Subtracting equation (7.4.51) from (7.4.50) and solving for λ gives

$$\lambda = \frac{r}{K} \frac{\int_{-L/2}^{L/2} u_0^2 y \, dx}{\int_{-L/2}^{L/2} u_0 y \, dx}.$$

Because y is of one sign in the interval, we conclude that the smallest eigenvalue is positive; therefore all the eigenvalues are positive and the equilibrium solution u_0 is stable to small perturbations.

Figure 7.6 summarizes the stability properties of the two equilibrium solutions, u_0 and the trivial equilibrium 0. \square

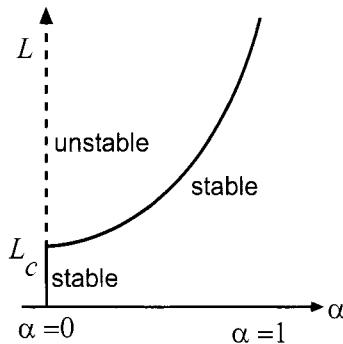


Figure 7.6 If $L < L_c = \pi\sqrt{D/r}$, the only equilibrium solution is the trivial one, with $\alpha = 0$. As L increases there is a bifurcation at L_c into two solutions, with a new nonzero solution appearing. For $L > L_c$, the zero solution becomes unstable while the nonzero solution is stable.

EXERCISES

1. Consider the linear problem

$$\begin{aligned} u_t &= u_{xx} + u, \quad x \in (0, L), \\ u(0) &= u(L) = 0, \end{aligned}$$

where L is a bifurcation parameter. Show that a bifurcation of the equilibrium solution occurs at $L = \pi$ and determine the stability for the cases $L < \pi$ and $L > \pi$.

2. Consider the nonlinear elliptic problem

$$\begin{aligned} -\Delta u &= f(u), \quad x \in D, \\ u &= 0, \quad x \in \partial D. \end{aligned}$$

where the function $f(u) - \lambda_1 u$ does not vanish identically on any open interval $(0, u^*)$, where λ_1 is the principal eigenvalue of $-\Delta$ on the domain D .

- Prove that if this problem has a positive solution, then necessarily $f(u) - \lambda_1 u$ changes sign for $u < 0$. *Hint:* Compute the integral $\int_D (f(u) - \lambda_1 u) \phi \, dx$, using Green's identity, where ϕ is the positive eigenfunction corresponding to λ_1 .
- Given $f(u) - \lambda_1 u < 0$ for $u > 0$, prove that a solution cannot be positive at any point in D .

- Consider the nonlinear elliptic problem

$$\begin{aligned} -\Delta u + f(u) &= \mu u, & x \in D, \\ u &= 0, & x \in \partial D, \end{aligned}$$

where $f \in C^2(\mathbb{R})$, $f(0) = f'(0) = 0$, where $f(u)/u$ is strictly increasing (decreasing) for $u > 0$ ($u < 0$), and where $\lim_{|u| \rightarrow \infty} f(u)/u = +\infty$. Let λ_1 be as in Exercise 2. Prove the following:

- If $0 < \mu < \lambda_1$, the problem has only the trivial solution.
- If $\mu > \lambda_1$, there exists at least one positive nontrivial solution.
- If $\mu > \lambda_1$, the problem has a unique positive solution.

[See Hernandez (1986) or Stakgold & Payne (1973).]

- Consider the nonlinear initial-boundary value problem

$$\begin{aligned} u_t - u_{xx} &= \mu u(1 - u), & 0 < x < \pi, & t > 0, \\ u(0, t) &= u(\pi, t) = 0, & t > 0, \\ u(x, 0) &= \varepsilon u_0(x), & 0 < x < \pi. \end{aligned}$$

Use a linear stability analysis to investigate the stability of the zero solution, and determine a critical value μ_c of the parameter μ above which the zero solution is unstable. For μ only slightly larger than μ_c , perform a nonlinear stability analysis to determine the long-time behavior of the solution.

- Consider the pure initial value problem

$$u_t - u_{xx} = \mu u(1 - u), \quad x \in \mathbb{R}, \quad t > 0,$$

on an infinite domain. Both $u = 0$ and $u = 1$ are asymptotic solutions. Use linear stability analysis to determine the stability of each of these solutions.

(Hint: First take $u = \varepsilon v$; then take $u = 1 + \varepsilon v$. Use the superposition principle and the Fourier integral theorem to solve the linearized problems.) Solve the nonlinear problem exactly when the initial condition is given by $u(x, 0) = \varepsilon$, $x \in \mathbb{R}$.

6. Consider the nonlinear BVP for $u = u(x, y, t)$:

$$\begin{aligned} u_t &= \frac{1}{R}(u_{xx} + u_{yy}) - u_{xxxx} - u_{xx}u_y, \quad x \in \mathbb{R}, 0 < y < 1, t > 0, \\ u(x, 0, t) &= 0, \quad u(x, 1, t) = 1, \quad x \in \mathbb{R}, t > 0. \end{aligned}$$

- (a) Show that $u = y$ is an equilibrium solution.
- (b) Letting $u = y + U(x, y, t)$, where U is a small perturbation, find the linearized perturbation equation for U .
- (c) Letting $U = e^{\sigma t}e^{ikx}F(y)$, show that

$$F'' + (k^2 R - \sigma R - k^4 R - k^2)F = 0, \quad F(0) = F(1) = 0.$$

Hence, show that solutions exist of the form

$$F_n(y) = \sin n\pi y$$

exist provided

$$\sigma = \sigma_n = k^2 - k^4 - \frac{1}{R}k^2 + \frac{n^2\pi^2}{R}, \quad n = 1, 2, 3, \dots$$

- (d) Given

$$R > \frac{k^2 + n^2\pi^2}{k^2 - k^4},$$

prove that the n th mode grows without bound, and therefore that $u = y$ is unstable.

- (a) Find the equilibrium solution of

$$u_t = u_{xx} - uu_x, \quad 0 < x < L, \quad u(0) = 0, \quad u(L) = 1.$$

7. Show that the trivial solution to (7.4.44)–(7.4.45) is stable if $L < \pi\sqrt{D/r}$ and unstable if $L > \pi\sqrt{D/r}$.

8. Consider

$$u_t = u_{xx} + f(u), \quad 0 < x < L,$$

with no-flux boundary conditions at $x = 0$ and $x = L$. Let $f(u_e) = 0$ and determine conditions for which u_e is locally stable.

9. (*Insect Swarms*) In one dimension, consider a swarm of insects that advects to the right or the left depending on whether the density is higher to the right of the center of the distribution or to the left of center, respectively. The model is

$$u_t = Du_{xx} - (c(x, t)u)_x,$$

where

$$c(t, x) = \int_x^\infty u(y, t)dy - \int_{-\infty}^x u(y, t)dy.$$

- (a) Let $v(x, t) = \int_{-\infty}^x u(y, t)dy$ and $U(t) = \int_{\mathbb{R}} u(y, t)dy$ and show that

$$v_t = Dv_{xx} - (U - 2v)v_x,$$

with $v \rightarrow 0$ as $x \rightarrow -\infty$, and $v \rightarrow M$ as $x \rightarrow +\infty$, where M is constant.

- (b) Find the equilibrium solution to the problem in part (b).
(c) Find the equilibrium solution for u , and sketch its graph.

10. On \mathbb{R}^2 consider

$$u_t = \Delta u, \quad |x|, |y| < 1, \quad t > 0,$$

with Robin data

$$\frac{du}{dn} + au = 0, \quad |y| = 1; \quad \frac{du}{dn} - bu = 0, \quad |x| = 1$$

with $a, b > 0$. Show that the trivial solution is unstable.

Reference Notes. There are many outstanding general treatments, as well as books treating special topics, on elliptic equations. For general study, one can consult Gilbarg & Trudinger (1983), Courant & Hilbert (1953, 1962), and Smoller (1983). Linear eigenvalue problems are discussed in Courant & Hilbert (1953). The classic book by Protter & Weinberger (1967) has a complete treatment of the maximum principle for elliptic and parabolic equations; see also Smoller (1994). The book by Pao (1992) contains a thorough treatment of nonlinear elliptic problems and monotone iteration methods. Bifurcation problems are discussed in many places; we cite the monograph by Sattinger (1973) and the review article by Stakgold (1971). The review by Hernandez (1986) on elliptic problems is recommended, as well as the account in the monograph by Bebernes & Eberly (1989), which is focused on combustion. Zauderer (2006) can be consulted for a treatment of the nonlinear stability analysis based on normal modes, and asymptotic methods in general.

References

- Abramowitz, M. & Stegun, I. A. 1964. *Handbook of Mathematical Functions*, National Bureau of Standards, Washington, DC.
- Allen, L. J. S. 2007. *An Introduction to Mathematical Biology*, Pearson Education, Inc., Upper Saddle River, NJ.
- Ames, W. F. 1965. *Nonlinear Partial Differential Equations in Engineering*, Vol. 1, Academic Press, New York.
- Ames, W. F. 1972. *Nonlinear Partial Differential Equations in Engineering*, Vol. 2, Academic Press, New York.
- Aronson, D. G. 1986. The porous medium equation, in *Nonlinear Diffusion Problems*, Lecture Notes in Mathematics, No. 1224, 1–46, Springer-Verlag, New York.
- Bebernes, J. & Eberly, D. 1989. *Mathematical Problems from Combustion Theory*, Springer-Verlag, New York.
- Bender, C. M. & Orszag, S. A. 1978. *Advanced Mathematical Methods for Scientists and Engineers*, McGraw-Hill, New York.
- Bhatnagar, P. L. 1979. *Nonlinear Waves in One-Dimensional Dispersive Systems*, Oxford University Press, Oxford, UK.
- Birkhoff, G. & Rota, G.-C. 1965. *Ordinary Differential Equations*, Blaisdell, Waltham, MA.
- Bischoff, K. B., Himmelstein, K. J., Dedrick, R. L., & Zaharko, D. S. 1973. Pharmacokinetics and cell population growth models in cancer chemotherapy, *Chem. Eng. Med. Biol. (Advances in Chemistry Series)* **118**, 47–64.

An Introduction to Nonlinear Partial Differential Equations, Second Edition.
By J. David Logan
Copyright © 2008 John Wiley & Sons, Inc.

- Brauer, F & Castillo-Chavez, C. 2001. *Mathematical Models in Population Biology and Epidemiology*, Springer-Verlag, New York.
- Britton, N. F. 1986. *Reaction-Diffusion Equations and Their Applications to Biology*, Academic Press, San Diego.
- Britton, N. F. 2003. *Essential Mathematical Biology*, Springer-Verlag, New York.
- Canosa, J. 1973. On a nonlinear diffusion equation describing population growth, *IBM J. Res. Devel.* **17**, 307–313.
- Churchill, R. V. 1969. *Fourier Series and Boundary Value Problems*, 2nd ed., McGraw-Hill, New York.
- Cohn, S. & Logan, J. D. 1995. Mathematical analysis of a reactive-diffusive model of the dispersal of a chemical tracer with nonlinear convection, *Math. Models Meth. Appl. Sci.* **5**(1), 29–46.
- Courant, R. & Friedrichs, K. O. 1948. *Supersonic Flow and Shock Waves*, Wiley-Interscience, New York (reprinted by Springer-Verlag, New York, 1976).
- Courant, R. & Hilbert, D. 1953. *Methods of Mathematical Physics*, Vol. 1, Wiley-Interscience, New York.
- Courant, R. & Hilbert, D. 1962. *Methods of Mathematical Physics*, Vol. 2, Wiley-Interscience, New York.
- Cushing, J. M. 1994. *An Introduction to Structured Population Dynamics*, Society for Industrial and Applied Mathematics, Philadelphia.
- Debnath, L. 1997. *Nonlinear Partial Differential Equations for Scientists and Engineers*, Birkhauser, Boston.
- de Marsily, G. 1986. *Quantitative Hydrogeology*, Academic Press, San Diego.
- Drazin, P. G. & Johnson, R. S. 1989. *Solitons: An Introduction*, Cambridge University Press, Cambridge, UK.
- Dresner, L. 1983. *Similarity Solutions of Nonlinear Partial Differential Equations*, Pitman Books, London.
- Dunbar, S. 1983. Traveling wave solutions of the Lotka-Volterra equations, *J. Math. Biol.* **17**, 11–32.
- Edelstein-Keshet, L. 2005. *Mathematical Models in Biology*, SIAM (Soc. Industr. Appl. Math.), Philadelphia.
- Evans, L. C. 1998. *Partial Differential Equations*, Am. Math. Soc., Providence, RI.

- Fickett, W. & Davis, W. C. 1979. *Detonation*, University of California Press, Berkeley.
- Fife, P. C. 1979. *Mathematical Aspects of Reacting and Diffusing Systems*, Springer-Verlag, New York.
- Friedman, A. 1964. *Partial Differential Equations of Parabolic Type*, Prentice Hall, Englewood Cliffs, NJ.
- Gilbarg, D. & Trudinger, N. S. 1983. *Elliptic Partial Differential Equations of Second Order*, 2nd ed., Springer-Verlag, New York.
- Gierer, A. & Meinhardt, H. 1972. A theory of biological pattern formation, *Kybernetik* **12**, 30–39.
- Grindrod, P. 1996. *The Theory of Reaction-Diffusion Equations: Patterns and Waves*, 2nd ed., Oxford University Press, Oxford, UK.
- Guenther, R. B. & Lee, J. W. 1988. *Partial Differential Equations of Mathematical Physics and Integral Equations*, Prentice Hall, Englewood Cliffs, NJ (reprinted by Dover Publications, 1996).
- Hartman, P. 1964. *Ordinary Differential Equations*, Wiley, New York.
- Hernandez, J. 1986. Qualitative methods for nonlinear differential equations, in *Nonlinear Diffusion Problems*, Lecture Notes in Mathematics, No. 1224, Springer-Verlag, New York, pp. 47–118.
- Hirsch, M. W., Smale, S., & Devaney, R. L. 2004. *Differential Equations, Dynamical Systems, and an Introduction to Chaos*, Elsevier, New York.
- John, F. 1982. *Partial Differential Equations*, 4th ed., Springer-Verlag, New York.
- Kapila, A. K. 1983. *Asymptotic Treatment of Chemically Reacting Systems*, Pitman, Boston.
- Keener, J. & Sneyd, J. 1998. *Mathematical Physiology*, Springer-Verlag, New York.
- Kevorkian, J. 1990. *Partial Differential Equations*, Brooks/Cole, Pacific Grove, CA.
- Kevorkian, J. & Cole, J. D. 1981. *Perturbation Methods in Applied Mathematics*, Springer-Verlag, New York.
- Kolmogorov, A., Petrovsky, I., & Piscounoff, N. 1937. Study of the diffusion equation with growth of the quantity of matter and its application to a biology problem, in *Dynamics of Curved Fronts*, P. Pelce, ed., Academic

- Press, San Diego, 1988 [originally published in *Bull. Univ. Etat Moscou Ser. Int. A* **1** (1937)].
- Kot, M. 2001. *Elements of Mathematical Ecology*, Cambridge University Press, Cambridge, UK.
- Kriess, H. O. & Lorenz, J. 1989. *Initial-Boundary Value Problems and the Navier-Stokes Equations*, Academic Press, New York.
- Kurganov, K. & Rosenau, P. 1997. Effects of a saturating dissipation in Burgers-type equations, *Commun. Pure Appl. Math.* **50**, 753–761.
- Landau, L. D. & Lifshitz, E. M. 1987. *Fluid Mechanics*, 2nd ed., Pergamon Press, Oxford, UK.
- Lax, P. 1973. *Hyperbolic Systems of Conservation Laws and the Mathematical Theory of Shock Waves*, Society for Industrial and Applied Mathematics, Philadelphia.
- LeVeque, R. J. 1992. *Numerical Methods for Conservation Laws*, Birkhäuser, Boston.
- Levin, S. A. & Segel, L. A. 1985. Pattern generation in space and aspect, *SIAM Rev.* **27**, 45–67.
- Lighthill, J. 1978. *Waves in Fluids*, Cambridge University Press, Cambridge, UK.
- Lin, C. C. & Segel, L. A. 1974. *Mathematics Applied to Deterministic Problems in the Natural sciences*, Macmillan, New York (reprinted by the Society for Industrial and Applied Mathematics, Philadelphia).
- Logan, J. D. 1977. *Invariant Variational Principles*, Academic Press, New York.
- Logan, J. D. 1987. *Applied Mathematics: A Contemporary Approach*, Wiley-Interscience, New York.
- Logan, J. D. 1989. Forced response of a linear hyperbolic system, *Appl. Anal.* **33**, 255–266.
- Logan, J. D., 1994. *Introduction to Nonlinear Partial Differential Equations*, Wiley-Interscience, New York.
- Logan, J. D. 2001. *Transport Modeling in Hydrogeochemical Systems*, Springer-Verlag, New York.
- Logan, J. D. 2003. Nonlocal advection problems, *Int. J. Math. Edu. Sci. & Technol.* **34**(2), 271–277.

- Logan, J. D. 2004. *Applied Partial Differential Equations*, 2nd ed., Springer-Verlag, New York.
- Logan, J. D. 2006a. *Applied Mathematics*, 3rd ed., Wiley-Interscience, New York.
- Logan, J. D. 2006b. *A First Course in Differential Equations*, Springer-Verlag, New York.
- Logan, J. D. 2008. Phenologically-structured predator-prey dynamics with temperature dependence, *Bull. Math. Biol.* **70**(1), 1–20.
- Logan, J. D. & Dunbar, S. R. 1992. Traveling waves in model reacting flows with reversible kinetics, *IMA J. Appl. Math.* **49**, 103–121.
- Logan, J. D., Joern, A., & Wolesensky, W. 2002. Location, time, and temperature dependence of digestion in simple animal tracts, *J. Theor. Biol.* **216**, 5–18.
- Logan, J. D. & Shores, T. S. 1993. On a system of nonlinear hyperbolic conservation laws with sources, *Math. Models Meth. Appl. Sci.* **3**(3), 341–358.
- Majda, A. J. 1986. *Compressible Fluid Flow and Systems of Conservation Laws in Several Space Variables*, Springer-Verlag, New York.
- McOwen, R. C. 2003. *Partial Differential Equations*, 2nd ed., Pearson, Upper Saddle River, NJ.
- Metz, J. A. J., De Roos, A. M., & Van den Bosch, F. 1988. Population models incorporating physiological structure: A quick survey of the basic concepts and an application to size-structured population Dynamics in waterfleas, in *Size-Structured Populations*, B. Ebermann & L. Persson, eds., Springer-Verlag, New York, pp 106–126.
- Murray, J. D. 1975. Nonexistence of wave solutions for the class of reaction-diffusion equations given by the Volterra interacting population equations with diffusion, *J. Theor. Biol.* **52**, 459–469.
- Murray, J. D. 1977. *Lectures on Nonlinear-Differential-Equation Models in Biology*, Oxford University Press, Oxford, UK.
- Murray, J. D. 1984. *Asymptotic Analysis*, Springer-Verlag, New York.
- Murray, J. D. 1988. How the leopard gets its spots, *Sci. Am.* **258**, 80–87.
- Murray, J. D. 2002. *Mathematical Biology I: An Introduction*, 3rd ed., Springer-Verlag, New York.
- Murray, J. D. 2003. *Mathematical Biology II: Spatial Models and Biomedical Applications*, 3rd ed., Springer-Verlag, New York.

- Noble, J. V. 1974. Geographical and temporal development of plagues, *Nature* **250**, 726–729.
- Okubo, A. & Levin, S. A. 2001. *Diffusion and Ecological Problems*, 2nd ed., Springer-Verlag, New York.
- Pao, C. V. 1992. *Nonlinear Elliptic and Parabolic Equations*, Plenum Press, New York.
- Polking, J. 2004. *dfield7.m* and *pplane7.m*, <http://www.rice.edu/~polking>.
- Protter, M. H., & Weinberger, H. 1967. *Maximum Principles in Differential Equations*, Prentice-Hall, Englewood Cliffs, NJ.
- Renardy, M. & Rogers, R. C. 2004. *An Introduction to Partial Differential Equations*, 2nd ed., Springer-Verlag, New York.
- Rhee, H.-K., Aris, R. & Amundson, N. R. 2001. *First-Order Partial Differential Equations*, Vols. I, II, Dover Publications, Mineola, NY.
- Rogers, C. & Ames, W. F. 1989. *Nonlinear Boundary Value Problems in Science and Engineering*, Academic Press, New York.
- Rozdestvenskii, B. L. & Janenko, N. N. 1983. *Systems of Quasilinear Equations and Their Applications to Gas Dynamics*, American Mathematical Society, Providence, RI (translated from the 1968 Russian edition by J. R. Schulenberger).
- Sattinger, D. 1973. *Topics in Stability and Bifurcation Theory*, Lecture Notes in Mathematics, No. 309, Springer-Verlag, New York.
- Shigesada, N. & Kawasaki, K. 1997. *Biological Invasions: Theory and Practice*, Oxford University Press, Oxford, UK.
- Smoller, J., 1994. *Shock Waves and Reaction-Diffusion Equations*, 2nd ed., Springer-Verlag, New York.
- Stakgold, I., 1971. Branching of solutions on nonlinear equations, *SIAM Rev.* **13**, 289–332.
- Stakgold, I. & Payne, L. E. 1973. Nonlinear problems in nuclear reactor analysis, in *Nonlinear Problems in the Physical Sciences and Biology*, Lecture Notes in Mathematics, No. 322, Springer-Verlag, New York, pp. 298–307.
- Stakgold, I., 1998. *Green's Functions and Boundary Value Problems*, 2nd ed., Wiley-Interscience, New York.
- Strauss, A. 1992. *Partial Differential Equations*, John Wiley and Sons, New York.

- Strichartz, R. 1994. *A Guide to Distribution Theory and Fourier Transforms*, CRC Press, Boca Raton, FL.
- Strogatz, S. H. 1994. *Nonlinear Dynamics and Chaos*, Addison-Wesley, Reading, MA.
- Turing, A. M. 1952. The chemical basis of morphogenesis, *Proc. Roy. Soc. London* **B237**, 37–72.
- van Duijn, C. J. & Knabner, P. 1992. Traveling waves in the transport of reactive solutes through porous media: Adsorption and binary ion exchange, *Transport in Porous Media* **8**, 167–194.
- Volpert, A. I. & Volpert, V. A. 1994. *Traveling Wave Solutions of Parabolic Systems*, American Mathematical Society, Providence, RI.
- Whitham, G. B., 1974. *Linear and Nonlinear Waves*, Wiley-Interscience, New York.
- Zauderer, E., 2006. *Partial Differential Equations of Applied Mathematics*, 3rd ed., Wiley-Interscience, New York.

Index

- acoustics, 168
- advection, 29
- age structure, 97
- Allee model, 242
- Arrhenius law, 272
- Banach space, 302
- barotropic flow, 166, 177
- bifurcation diagram, 365
- biological invasion, 278
- bioremediation, 291
- Birkhoff, G. D., 217
- blowup, 323
- bore, 187
- boundary condition, 4
 - Dirichlet, 41, 347
 - Neumann, 41, 347
- breaking time, 120, 127
- Burgers' equation, derivation, 202
- Cauchy problem, 35, 91
- characteristic coordinates, 6
- characteristic curve, 6
- characteristic diagram, 63
- characteristic parallelogram, 17
- characteristic system, 88
- characteristics, 6, 63, 87, 171
 - derivatives along, 82
- chemical reactor, 136
- chemotaxis, 274, 335
- chemotherapy, 103
- Chezy law, 140
- classification, 5, 170
- combustion, 271
- conservation law
 - higher dimensions, 23
 - integral form, 22
 - one-dimensional, 20
- constitutive relation, 25
- contaminant transport, 284
- contraction mapping, 296
- convex, quasi-, 329
- convolution, 216
- critical domain size, 377
- critical patch size, 339
- critical point, 257
- D'Alembert's formula, 4
- Darcy's law, 31
- dispersion relation, 51
- eigenvalue problem, 358
- elliptic equation, 5
- energy balance, 166, 272
- entropy condition, 145
- equal area principle, 149
- equation
 - advection, 29, 37
 - backward diffusion, 44
 - Barenblatt, 216
 - Born-Infeld, 198
 - Burgers, 28, 205
 - Burgers', 245, 250
 - diffusion, 26, 37, 210, 220
 - diffusion in higher dimensions, 29
 - diffusive, 52

An Introduction to Nonlinear Partial Differential Equations, Second Edition.

By J. David Logan

Copyright © 2008 John Wiley & Sons, Inc.

- dispersive, 52
 eikonal, 92, 93
 elliptic, 7
 Euler–Darboux, 19
 Euler–Lotka, 99
 Fisher, 28, 32, 235, 341
 groundwater, 31
 Hamilton–Jacobi, 93
 heat, 27
 Helmholtz, 93, 352
 hyperbolic, 6
 inviscid Burgers', 29, 33, 38, 252
 Korteweg–deVries, 47
 Laplace, 347
 McKendrick–Von Foerster, 96
 Nagumo, 242
 Navier–Stokes, 331
 nonlinear first-order, 68
 nonlinear Schrödinger, 58
 parabolic, 7
 Poisson, 348
 porous media, 32
 porous medium, 34, 231
 quasi-linear, 8
 quasilinear, 73
 reaction–diffusion, 28, 269
 renewal, 99
 Schrödinger, 52, 350
 shallow-water, 163, 169
 wave, 4, 42
 erf function, 43
- Fick's law, 26, 30, 210, 212
 first integral, 76
 first-order equations, 86
 fixed-point iteration, 294
 flood waves, 140
 Fourier integral theorem, 16
 Fourier method, 52
 Fourier series, 14
 Fourier's heat conduction law, 27
 fundamental solution, diffusion equation, 210
- gas dynamics, 164, 177
 Gelfand problem, 330, 361, 363
 Gierer–Meinhardt model, 342
 Gorter–Mellink law, 226
 Green's identity, 350
- Hölder's inequality, 319
 Hadamard's example, 347
 harmonic, 353
 heat conduction, 225
- heat equation, 27
 hyperbolic equation, 5
 hyperbolic system, 171
- integration by parts, 350
 invariance of a PDE, 218
 invariant set, 326
- Jacobian matrix, 263
 jump condition, 83, 117, 144, 180
- kernel, diffusion, 210
- Lipschitz condition, 94, 294, 297, 300
 Lotka–Volterra, 338
 Lotka–Volterra model, 101, 106, 271
- mass-structured population, 105
 maximum principle
 strong, 313, 354
 weak, 311, 353
- moments, method of, 102
 momentum balance, 165
- non-Fickian flux, 225
 nonlinear dispersion, 54
 nonlinear equation, 9
 normed linear space, 301
 nullcline, 258
- Ohmic heating, 276
 orbit
 heteroclinic, 264
 homoclinic, 264
- Ornstein–Uhlenbeck process, 223
- parabolic equation, 5
 pattern formation, 333
 phase diagram, 258
 phase plane phenomena, 258
 Picard iteration, 296
 plane wave, 50
 Poincaré inequalities, 319
 porous medium, 146
 principal eigenvalue, 360
 pulse, 46
- random walk, 213
 Rankine–Hugoniot condition, 117
 rarefaction wave, 119
 region of influence, 215
 Riemann invariant, 175, 176
 Riemann problem, 114
 Rosenzweig–MacArthur model, 107
- self-similar solution, 218

- Semenov problem, 273
separation of variables, 11
separatrix, 260
shock fitting, 121, 152
shock speed, 118, 180
shock structure, 247
shock wave, 180
signaling problem, 40, 66
similarity method, 217
simple wave, 187
size-dependent predation, 106
Sobolev inequality, 331
solid-fuel model, 272, 361
soliton, 49
solution
 classical, 2, 141
 complete integral, 91
 general, 76
 weak, 142
stability, 258, 368
Sturm–Liouville problem, 13
subharmonic, 353
support of a function, 141
- theorem
 Buckingham’s Pi, 32
 Banach fixed-point, 303
 comparison, 314
 divergence, 350
 existence, 356
 Fourier integral, 53
 Haar, 94
- Laplace’s, 250
local existence, 304
Poincaré–Bendixson, 258
principal eigenvalue, 360
uniqueness, 94
tidal bore, 197
traffic flow, 132
transformation
 Cole–Hopf, 33, 248
 hodograph, 18, 192
 Legendre, 18
 Liouville–Green, 241
 similarity, 218
 stretching, 217
traveling wave, 45
 Fisher’s equation, 235
 SIR model, 278
Turing system, 333
- uniformly elliptic, 354
- vibrating membrane, 359
- wave number, 50
wavefront expansion, 193
wavefront solution, 45
weakly nonlinear, 85
weakly nonlinear model, 204
well-posed problem, 36
Whitham’s rule, 154
- Young’s inequality, 318

An Introduction to Nonlinear Partial Differential Equations, Second Edition.
By J. David Logan
Copyright © 2008 John Wiley & Sons, Inc.

PURE AND APPLIED MATHEMATICS

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

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 Categories
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
BASENER—Topology and Its Applications
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
BRIDGER—Real Analysis: A Constructive Approach
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, GARCÍA, 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

*Now available in a lower priced paperback edition in the Wiley Classics Library.

†Now available in paperback.

- FATICONI—The Mathematics of Infinity: A Guide to Great Ideas
- FOLLAND—Real Analysis: Modern Techniques and Their Applications
- 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 and Its Applications, Second Edition
- LOGAN—An Introduction to Nonlinear Partial Differential Equations, Second Edition
- MARKLEY—Principles of Differential Equations
- MORRISON—Functional Analysis: An Introduction to Banach Space Theory
- NAYFEH—Perturbation Methods
- NAYFEH and MOOK—Nonlinear Oscillations
- O'NEIL—Beginning Partial Differential Equations, Second Edition
- 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

*Now available in a lower priced paperback edition in the Wiley Classics Library.

[†]Now available in paperback.

- SHICK—Topology: Point-Set and Geometric
*SIEGEL—Topics in Complex Function Theory
 Volume 1—Elliptic Functions and Uniformization Theory
 Volume 2—Automorphic Functions and Abelian Integrals
 Volume 3—Abelian Functions and Modular Functions of Several Variables
SMITH and ROMANOWSKA—Post-Modern Algebra
ŠOLIN—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
STANOVYEVITCH—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.