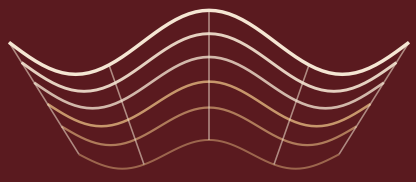
MATHEMATICAL PHYSICS

An Introduction to

the Linear Equations of Mathematical Physics

Boundary-Value Problems and Integral Transforms



GÖKSEL DAYLAN ESMER

İKSAD Publishing House

MATHEMATICAL PHYSICS

An Introduction to

the Linear Equations of Mathematical Physics

Boundary-Value Problems and Integral Transforms

GÖKSEL DAYLAN ESMER

İKSAD Publishing House

Copyright © 2026 by iksad publishing house
All rights reserved. No part of this publication may be reproduced, distributed or transmitted in any form or by
any means, including photocopying, recording or other electronic or mechanical methods, without the prior written permission of the publisher, except in the case of
brief quotations embodied in critical reviews and certain other noncommercial uses permitted by copyright law. Institution of Economic Development and Social Researches Publications®
(The Licence Number of Publicator: 2014/31220)
TÜRKİYE TR: +90 342 606 06 75
USA: +1 631 685 0 853
E mail: iksadyayinevi@gmail.com
www.iksadyayinevi.com

It is responsibility of the author to abide by the publishing ethics rules.
Iksad Publications – 2026©

ISBN: 978-625-378-659-5
Cover Design: Göksel Daylan ESMER
June / 2026
Ankara / Türkiye

Author:
Göksel Daylan Esmer
Department of Physics
Faculty of Science, Istanbul University
Vezneciler, 34134 Istanbul
Turkey

Mathematics Subject Classification (2020):
35-01, 35A08, 35J05, 35K05, 35L05, 35P10

© Göksel Daylan Esmer

This work is subject to copyright. All rights are reserved by the author, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The author and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the author nor the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Prof. Dr. Mehmet Faruk Oktay

I dedicate this book to the memory of my beloved cousin, ***Prof. Dr. Mehmet Faruk Oktay****, whom we recently lost to a relentless illness, and who touched many lives and left a profound mark upon them.*

Preface

The partial differential equations that arise in physics are, with very few exceptions, of the second order, and the analytical methods needed to solve them lie at the heart of mathematical physics. The aim of this book is to present those methods in a single connected exposition that begins, after a chapter on the analytical preliminaries, with the basic vocabulary of partial differential equations and ends with the finite integral transforms by which a wide class of boundary-value problems is reduced to a routine calculation.

The book is organised into thirteen chapters, the first of which collects the analytical preliminaries used throughout, and includes a total of some fifty worked exercises and seventy-five problems appended at the end of the chapters, together with a list of references and a subject index at the very end. Throughout the text I have tried to keep three pedagogical principles in mind. The first is to write down, rather than to suppress, the algebraic steps that connect one equation to the next, so that the reader is never asked to take a calculation on trust. The second is to insist, repeatedly, on the principle of superposition—introduced clearly in Chapter 2 and used as the guiding thread of the entire treatment. The third is to show how every special function of physics arises, in a natural way, as the eigenfunction of a Sturm–Liouville problem; for that reason the Sturm–Liouville theory is treated in full detail, and the subtle question of completeness of its eigenfunctions is presented in a deliberately pedagogical manner.

Chapter 1 gathers, without proof, the elements of vector analysis, ordinary differential equation theory, and Fourier-series theory on which the rest of the book draws; the reader who is at ease with this material may proceed directly to Chapter 2. Chapter 5, situated between the classification of second-order equations and the Cauchy problem, isolates and names the notions of boundary condition and well-posedness, in Hadamard’s sense, that organise the analytical chapters that follow. Chapters 11 and 12, placed after the bounded-domain methods of Chapters 7–10, develop the Green’s-function method for free-space and inhomogeneous problems and the variational/energy framework that unifies the analytic theory of the canonical equations. The final chapter, Chapter 13, sets the linear theory in a wider context, sketching the directions in which it is generalised by the introduction of nonlinearity, by modern

functional-analytic and probabilistic frameworks, and by the numerical methods of contemporary scientific computing.

Only *linear* partial differential equations are considered, and because the partial differential equations encountered in physics are most commonly of the second order, the book concentrates on second-order equations and the initial- and boundary-value problems associated with them. First-order equations and nonlinear equations are not treated; the closing chapter, Chapter 13, surveys briefly the directions in which the modern theory extends beyond these restrictions.

I am deeply grateful to Professor Şehsuvar Zebitay for his guidance and for many valuable discussions over the years. I am indebted, too, to the authors of the works collected in the references at the end of the book, which were most extensively consulted during its preparation, and I record here my admiration, respect, and gratitude to those authors—most notably R. Courant and D. Hilbert, P. R. Garabedian, N. S. Koshlyakov, M. M. Smirnov and E. B. Gliner, P. M. Morse and H. Feshbach, T. Myint-U, I. N. Sneddon, and A. Sommerfeld. I also wish to thank my students Dr. Aytül Filiz and Saliha Türkmen for their valuable technical support.

I am convinced that, with all its imperfections, the book will be useful to its readers and—perhaps more importantly—will encourage the writing of much better ones.

Göksel Daylan Esmer

List of Symbols

The following symbols and notations are used throughout the book.

Sets and spaces

$\mathbb{R}$, $\mathbb{R}^n$	Real numbers; n-dimensional Euclidean space
$\mathbb{C}$	Complex numbers
$\Omega \subset \mathbb{R}^n$	Bounded spatial domain
$\partial\Omega$	Boundary of Ω
$\bar{\Omega}$	Closure of Ω
$L^2(\Omega)$	Lebesgue space of square-integrable functions
$H^1(\Omega)$, $H_0^1(\Omega)$	Sobolev spaces

Differential operators

$\partial/\partial x_i$, u_{x_i}	Partial derivative with respect to x_i
∇, grad	Gradient
div, $\nabla\cdot$	Divergence
rot, $\nabla\times$	Curl
∇^2	Laplacian
$\Box^2$	d'Alembertian
$\partial/\partial n$	Outward normal derivative on $\partial\Omega$
∇^4	Biharmonic operator

Special functions and constants

$J_n(x)$, $Y_n(x)$	Bessel functions of the first and second kind
$P_n(x)$	Legendre polynomials

$P_n^m(x)$	Associated Legendre polynomials
$\delta(x)$, $\delta(\mathbf{r})$	Dirac delta
δ_{ij}	Kronecker delta
$H(x)$	Heaviside step function
$\Gamma(x)$	Gamma function
$\operatorname{erf}(x)$	Error function

Geometric quantities

$\mathbf{r} = (x, y, z)$	Position vector in $\mathbb{R}^3$
$r = \lvert\mathbf{r}\rvert$	Magnitude of position vector
(ρ, φ, z)	Cylindrical coordinates
(r, θ, φ)	Spherical coordinates
$\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$	Cartesian unit vectors
$\hat{\mathbf{n}}$	Outward unit normal on $\partial\Omega$
dV, dS	Volume and surface elements
ω_n	Surface area of unit sphere in $\mathbb{R}^n$

Equation- and problem-specific symbols

c	Wave speed
α, χ	Thermal diffusivity, conductivity
λ, λ_n	Separation constant; eigenvalue
φ_n	Eigenfunction of a Sturm–Liouville problem
$W(y_1, y_2)$	Wronskian determinant
$P(x)$, $Q(x)$, $R(x)$	Coefficients of a Sturm–Liouville equation
$E(\mathbf{r})$, $E(\mathbf{r}, t)$	Fundamental solution
$G(\mathbf{r}, \mathbf{r}')$	Green's function
$K(\mathbf{r}, \mathbf{r}'; t)$	Heat kernel

Variational symbols

$J[u]$, $S[u]$	Functionals
δJ, δI	First variation of a functional
$D[u]$	Dirichlet energy
$E(t)$, $E_T(t)$	Total energy; thermal energy
K_{ij}, F_i	Stiffness matrix; load vector

Other notation

$f * g$	Convolution of f and g
$\hat{f}(\xi)$	Fourier transform of f
$\bar{f}, \tilde{f}, \hat{f}$	Iterated finite integral transforms
$F_s[\cdot], F_c[\cdot]$	Finite Fourier sine, cosine transforms
$\langle \cdot, \cdot \rangle$	Inner product (occasionally)
$f \equiv g$	f identically equal to g
$f \propto g$	f proportional to g

Conventions. Vectors are typeset in boldface or arrowed. Equations are numbered by section: equation $(N.k.\ell)$ is the ℓ-th numbered equation of Section k of Chapter N. Cross-references to definitions, theorems, exercises, and problems carry the chapter number throughout the book; the abbreviation "Sect." is used for section references. Cross-references to equations of the same chapter omit the chapter prefix when no ambiguity arises.

Contents

1 Mathematical Preliminaries 5
1.1 Vector Analysis 5
1.2 Linear Ordinary Differential Equations 7
1.3 Trigonometric Fourier Series 9
1.4 The Special Functions of Mathematical Physics 10
Notational Conventions 12

2 General Properties of Partial Differential Equations 13
2.1 Fundamental Concepts 13
2.2 Linear Differential Operators and the Principle of Superposition 17
Problems 21

3 Mathematical Models and Partial Differential Equations 23
3.1 Transverse Vibrations of an Elastic String 23
3.2 The Vibrating Membrane 25
3.3 Waves in an Elastic Medium 26
3.4 Heat Conduction in Solids 30
3.5 The Gravitational Potential 31
3.6 Sound Waves Propagating in a Gas 33

4 Classification of Second-Order Partial Differential Equations 35
4.1 The Case of Two Independent Variables 35
4.2 Classification in More Than Two Independent Variables 43
4.3 Worked Exercises in Reduction to Canonical Form 45
Problems 52

5 Boundary Conditions and Well-Posedness 53
5.1 Auxiliary Conditions 53
5.2 Hadamard's Criteria of Well-Posedness 54
5.3 Well-Posed Problems Type by Type 55
Problems 57

6 The Initial-Value Problem for Partial Differential Equations 61
6.1 The Cauchy Problem 61
6.2 Initial-Value Problem for the Homogeneous Wave Equation 65
6.3 Initial-Value Problem for the Inhomogeneous Wave Equation: Method of Characteristics 67
6.4 Solution of the Cauchy Problem for Hyperbolic Equations: Riemann's Method 70
Problems 79

7 The Method of Separation of Variables 81
7.1 Solvability Conditions 81
7.2 The Dirichlet Problem for the Vibrating String 85
Problems 100

8 The Finite Fourier Transform 103
8.1 Lagrange's Method of Variation of Parameters 103
8.2 The Finite Fourier Transform Method 106
Worked Exercises 107
Exercise VI.1 107
Exercise VI.2 109
Exercise VI.3 110
Exercise VI.4 111
Exercise VI.5 113
Exercise VI.6 114
Exercise VI.7 115
Problems 117

9 The Sturm–Liouville Problem 121
9.1 The General Equations of Mathematical Physics and Their Separation 121
9.2 The Sturm–Liouville Problem and Its Eigenfunctions 124
9.3 The Variational Method 127
9.4 Asymptotic Behaviour of the Sturm–Liouville Eigenvalues 129
9.5 Completeness of the Eigenfunctions of the Sturm–Liouville Problem . 134
9.6 Orthogonality Relations for the Bessel and Legendre Functions 134
Problems 150

10 Finite Integral Transforms 153
10.1 Determination of the Integral Transform 153

11 Green's Functions and Fundamental Solutions 177
11.1 The Dirac Delta 177
11.2 Fundamental Solutions 180
11.3 Green's Functions on Bounded Domains 182
11.4 The Spectral Representation 185
Worked Exercises 186
Problems 188

12 Variational Methods and Energy Estimates 191
12.1 Functionals and the Euler–Lagrange Equation 191
12.2 The Dirichlet Principle 192
12.3 Hamilton's Principle and the Wave Equation 193
12.4 Energy Methods for the Heat Equation 195
12.5 The Rayleigh–Ritz Method 196
Worked Exercises 198
Problems 200

13 Beyond the Linear Theory: An Outlook 203
13.1 Nonlinear Partial Differential Equations 203
13.2 Higher-Order Equations and Systems 205
13.3 Probabilistic and Modern Functional-Analytic Perspectives 206
13.4 Numerical Methods 207
13.5 Closing Remarks 208
References 209

Index 211

List of Figures

1.1 The standard families of special functions arising in mathematical physics. **Left**: the Bessel functions of the first kind $J_0(x)$ and $J_1(x)$, both regular at the origin and oscillatory beyond their first zero (the Bessel functions of the second kind Y_n, having a logarithmic singularity at $x = 0$, are not shown). **Right**: the first four Legendre polynomials $P_n(x)$ on $[-1,1]$, normalised so that $P_n(1) = 1$; the polynomial P_n has exactly n zeros in the open interval $(-1,1)$. 11

3.1 Element AB of the displaced string. The arrows indicate the tensions T exerted on the element by the rest of the string at A and B, directed tangentially outward; α and β are the corresponding angles with the horizontal at x and $x+\Delta x$. 24

3.2 Surface element $\Delta S \simeq \Delta x \Delta y$ of the displaced membrane. Each Δx-edge carries an outward tangential tension of magnitude $T\Delta x$, and each Δy-edge a tension $T\Delta y$. The angles α, β are made by the two Δx-edges with the horizontal (x,y)-plane, and γ, δ by the two Δy-edges. 25

3.3 Volume element $\Delta V = \Delta x \Delta y \Delta z$ of the elastic medium. On every face whose outward normal is $+\hat{e}_i$ act three stress components τ_{ij} (one normal, two shear); the figure shows all three on the visible $+\hat{y}$-face, and only the normal stresses τ_{xx}, τ_{zz} on the other two visible faces, to avoid clutter. 27

3.4 Geometry of the gravitational interaction between m at the origin and m' at the tip of the position vector $\mathbf{r}$ of magnitude r. The attractive force $\mathbf{K}$ acts on m' along the line joining the two masses, anti-parallel to $\mathbf{r}$. 31

4.1 The three types of second-order linear PDEs in two independent variables, distinguished by the geometry of their characteristic curves. In the hyperbolic case ($\Delta = B^2 - 4AC > 0$) two distinct real families of characteristics cross transversally, and information propagates along them. In the parabolic case ($\Delta = 0$) the two families coalesce into a single one, along which information propagates instantaneously. In the elliptic case ($\Delta < 0$) no real characteristics exist; no propagation phenomenon is associated, and the natural problem is the boundary-value problem of Chapter 5. The distinction governs which auxiliary conditions yield well-posed problems. 40

5.1 Hadamard's example (5.2.1): the solution profiles $u_n(x, 0.5)$ at the fixed altitude $y = 0.5$ for $n = 2$, 6, 10. The Cauchy data on $y = 0$ tend uniformly to zero in n, yet at any fixed positive y the solution amplitude grows exponentially. The third Hadamard criterion (continuous dependence on the data) fails: an arbitrarily small perturbation of the boundary data produces an arbitrarily large change in the solution. 56

6.1 The Cauchy data curve C_0 through (x_0, y_0) with local basis vectors $\hat{e}_1$ (tangent) and $\hat{e}_2$ (normal). 62
6.2 Tangent $\hat{t}$ and outward normal $\hat{n}$ on C_0, with φ measured from the x-axis. 63
6.3 Domain-of-dependence triangle $P_0P_1P_2$ for the one-dimensional wave equation; B_1, B_2 are characteristic sides and B_0 lies on the initial line $y = 0$. 68
6.4 Domain of dependence in Exercise 6.3. 69
6.5 Riemann's geometry: an initial-data curve Λ, the point $P(\alpha, \beta)$, and the characteristic lines $\overline{PQ}$ and $\overline{PR}$ enclosing the region B. 71
6.6 When the characteristic $y = \beta$ through P meets Λ at two points Q, S, the Cauchy problem is overdetermined. 73
6.7 In the characteristic (ξ, η)-plane, the initial line $t = 0$ corresponds to the diagonal $\xi = \eta$. 74
6.8 Characteristic triangle PQR in the (ξ, η)-plane for the telegrapher's equation. 75
6.9 Region B in Exercise 6.6, with characteristic diagonals $\eta = \xi$ (image of $x = 0$) and $\eta = -\xi$ (image of $y = 0$). 77

7.1 The first four eigenmodes of the vibrating string, $X_n(x) = \sin(n\pi x/\ell)$, drawn at the instant of maximum displacement. The end-points $x = 0$ and $x = \ell$ are clamped (filled dots); the dashed curves show the same modes at the opposite phase of the oscillation. Mode n has $n - 1$ interior nodes (small black dots), oscillates with the angular frequency $\omega_n = n\pi\alpha/\ell$, and is the simplest physical realisation of an eigenfunction of the Sturm–Liouville problem (7.2.7). The general solution of the vibrating string is the linear superposition of these modes, weighted by the Fourier coefficients of the initial data. 87

7.2 Initial triangular profile of the plucked string in Exercise V.1. 88

9.1 Annular plate of inner radius a and outer radius b. 146
9.2 Heavy chain of length ℓ suspended at O, with transverse displacement u. 152

10.1 Hanging string of length ℓ in the equilibrium and a deflected configuration. 158
10.2 Slab of width L between $x = 0$ and $x = L$. 174

11.1 The Gaussian family $\eta_\varepsilon(x)$ of (11.1.3), plotted for three successively smaller values of ε. As $\varepsilon \to 0$ each profile narrows and grows, its total area remaining equal to 1. The family converges, in the distributional sense of (11.1.4), to the Dirac delta. 178
11.2 The backward light cone of the observation point $(\mathbf{r},t)$ in three-dimensional spacetime (drawn here with one spatial dimension suppressed). The light-cone surface is given by $|\mathbf{r}-\mathbf{r}_0| = c\,(t-t_0)$. By the Huygens principle, the value of $u(\mathbf{r},t)$ in formula (11.2.11) is determined by the values of the source f on the cone's surface alone—only those source events whose signal has had exactly enough time to reach $(\mathbf{r},t)$ at the speed c. The dashed segment from $(\mathbf{r}',t')$ to $(\mathbf{r},t)$ has length $c\,(t-t')$. 183
11.3 The method of images for the upper half-space with Dirichlet boundary condition on $\{z = 0\}$. A unit positive source placed at $\mathbf{r}' = (x',y',z_0)$ inside Ω is paired with an opposite "image" source at the reflected point $\mathbf{r}'_* = (x',y',-z_0)$. For any point $\mathbf{r}$ on the boundary plane $z = 0$ the two distances $|\mathbf{r}-\mathbf{r}'|$ and $|\mathbf{r}-\mathbf{r}'_*|$ coincide, so the two fundamental solutions in (11.3.5) cancel exactly: the homogeneous Dirichlet boundary condition is enforced by symmetry. 184

12.1 Comparison of the Rayleigh–Ritz trial function $\varphi(r) = 1 - r^2$ (blue) and the exact fundamental Dirichlet eigenfunction $J_0(j_{0,1}\,r)$ (red) on the unit disc, both normalised to the value 1 at the centre. The two profiles are visually almost indistinguishable; the small difference accounts for the 4% discrepancy between the Rayleigh-quotient estimate $\lambda_1 \le 6$ and the exact value $\lambda_1 = j_{0,1}^2 \approx 5.78$. 200

Introduction

The vast majority of the equations used to describe physical phenomena are differential equations, and these are most often partial differential equations (PDEs). The PDEs of mathematical physics can, in their majority, be reduced either to a single *master equation*,

$$\nabla^2\Psi(\mathbf{r},t) \,+\, \beta(\mathbf{r},t)\,\Psi(\mathbf{r},t) \,=\, A(\mathbf{r},t) \,+\, B(t)\,\frac{\partial\Psi(\mathbf{r},t)}{\partial t} \,+\, C(t)\,\frac{\partial^2\Psi(\mathbf{r},t)}{\partial t^2}, \quad (0.0.1)$$

or to one of its special cases. For example:

1. **Telegraphers' Equation.** With $A = 0$, $B = \text{const.}$, $C = \text{const.}$, and $\beta = \text{const.}$ in equation (0.0.1), one obtains the equation encountered in electrodynamics,

$$\nabla^2\Psi + \beta\,\Psi \,=\, B\,\frac{\partial\Psi}{\partial t} + C\,\frac{\partial^2\Psi}{\partial t^2}.$$

2. **Laplace's Equation.** For $A = B = C = \beta = 0$, equation (0.0.1) reduces to

$$\nabla^2\Psi \,=\, 0,$$

   which arises in electrostatics, magnetostatics, hydrodynamics, heat conduction, and the classical theory of gravitation.

3. **Poisson's Equation.** When $B = C = \beta = 0$, the master equation gives the inhomogeneous counterpart of Laplace's equation, the inhomogeneity being a source term:

$$\nabla^2\Psi \,=\, A(\mathbf{r},t).$$

4. **Wave Equation.** With $A = B = 0$, $\beta = \text{const.}$, and $C = \text{const.}$, one obtains

$$\nabla^2\Psi + \beta\,\Psi \,=\, C\,\frac{\partial^2\Psi}{\partial t^2},$$

   encountered in continuum mechanics, acoustics, sound waves, and electromagnetic theory.

5. **Diffusion Equation.** For $A = C = 0$, $\beta = \text{const.}$, and $B = \text{const.}$, the master equation becomes
$$\nabla^2\Psi + \beta\,\Psi \;=\; B\,\frac{\partial\Psi}{\partial t},$$
which appears in the theory of heat, the kinetic theory of gases, and the neutron transport theory.

6. **Helmholtz Equation.** The choice $A = B = C = 0$ with $\beta = \text{const.}$ yields the time-independent diffusion equation
$$\nabla^2\Psi + \beta\,\Psi \;=\; 0,$$
which occurs in classical mechanics, quantum mechanics, and the diffusion theory of neutrons.

7. **Heat-Conduction Equation.** With $\beta = C = 0$, denoting by χ the thermal-conductivity coefficient and by $\alpha \equiv \beta^{-1}$ the thermal-diffusivity coefficient, and writing $A(\mathbf{r},t) \equiv -\dfrac{1}{\chi}\,k(\mathbf{r},t)$ for the source term, one arrives at
$$\nabla^2\Psi + \frac{1}{\chi}\,k(\mathbf{r},t) \;=\; \frac{1}{\alpha}\,\frac{\partial\Psi}{\partial t}.$$

8. **Schrödinger Wave Equation.** Setting $A = C = 0$, $B = \text{const.}$, and writing the term in β as one proportional to a potential function,
$$\beta(\mathbf{r},t) \;\equiv\; -\frac{2m}{\hbar^2}\,U(\mathbf{r},t),$$
one obtains
$$-\frac{\hbar^2}{2m}\,\nabla^2\Psi + U(\mathbf{r},t)\,\Psi \;=\; i\hbar\,\frac{\partial\Psi}{\partial t}.$$

All of the above are particular cases of the master equation (0.0.1). Whenever the scalar function $\Psi(\mathbf{r},t)$ is replaced by a vector-valued function, the corresponding vector equations are obtained in an entirely analogous fashion.

Within the framework of the special theory of relativity, generalisations of some of these equations to four-dimensional Minkowski space—defined by the imaginary time coordinate $ict \equiv x_4$—are encountered. The Minkowski-space counterpart of the Laplacian is the d'Alembertian operator,
$$\Box^2 \;\equiv\; \nabla^2 + \frac{\partial^2}{\partial x_4^2} \;=\; \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} - \frac{1}{c^2}\,\frac{\partial^2}{\partial t^2},$$
through which one may write the time-dependent wave equation as
$$\Box^2\Psi \;=\; 0,$$
the generalised Poisson equation as

$$\Box^2 \Psi = A(\mathbf{r},t),$$

and the Klein–Gordon equation as

$$\Box^2 \Psi = \mu^2 \Psi.$$

These last three constitute further examples of the equations under discussion.

The partial differential equations of mathematical physics enumerated above are all of the second order and are linear. The famous Maxwell equations and Dirac equation of mathematical physics, however, consist of four *first*-order partial differential equations. Nevertheless, by elimination of one of the two unknown functions they contain, a single second-order PDE can be obtained for the remaining one. As is evident, the partial differential equations encountered in physics are predominantly of the second order; consequently, the present book will be primarily concerned with second-order linear PDEs.

The fact that these equations are linear, as will be demonstrated in Chapter 2, makes it possible to develop standard methods for the solution of the equations themselves and of the initial-value problems associated with them. By contrast, no general, exact solution methods are available for nonlinear equations—a state of affairs typified by the basic equations of fluid mechanics. Although the rigorous formulation of the physical phenomena described by the equations listed above generally leads to nonlinear equations, under suitable conditions (as in the examples cited) these reduce to linear ones. Accordingly, the present book is restricted to the presentation of methods for solving initial-value problems for linear partial differential equations, and only such examples will be discussed.

Chapter 1
Mathematical Preliminaries

Abstract This chapter fixes notation and collects, without proof, the analytical tools used throughout the book: vector calculus (gradient, divergence, curl, Laplacian; Gauss's and Stokes's theorems; the forms of ∇^2 in cylindrical and spherical coordinates), the elements of linear ordinary differential equation theory (existence and uniqueness, the Wronskian, the constant-coefficient equation), and the trigonometric Fourier series. The reader who is comfortable with these topics may proceed directly to Chapter 2.

1.1 Vector Analysis

Throughout the book we shall use $\mathbf{r} = (x, y, z)$ for the position vector in three-dimensional Euclidean space, $r = |\mathbf{r}|$ for its length, and $\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$ for the unit vectors along the Cartesian axes. The fundamental differential operator is the *nabla* (or del) operator,

$$\nabla \equiv \hat{\mathbf{i}} \frac{\partial}{\partial x} + \hat{\mathbf{j}} \frac{\partial}{\partial y} + \hat{\mathbf{k}} \frac{\partial}{\partial z}, \tag{1.1.1}$$

which acts on scalar and vector fields in three distinct ways.

Gradient. Applied to a scalar field $f(x, y, z)$, the operator ∇ produces the vector field

$$\operatorname{grad} f \equiv \nabla f = \frac{\partial f}{\partial x} \hat{\mathbf{i}} + \frac{\partial f}{\partial y} \hat{\mathbf{j}} + \frac{\partial f}{\partial z} \hat{\mathbf{k}}, \tag{1.1.2}$$

called the *gradient* of f. At each point of space grad f points in the direction of steepest increase of f, and its magnitude equals that rate of increase.

Divergence. For a vector field $\mathbf{A} = A_x \hat{\mathbf{i}} + A_y \hat{\mathbf{j}} + A_z \hat{\mathbf{k}}$, the scalar field

$$\operatorname{div} \mathbf{A} \equiv \nabla \cdot \mathbf{A} = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \tag{1.1.3}$$

is called the *divergence* of $\mathbf{A}$. Physically, $\operatorname{div}\mathbf{A}$ measures the net outflow of $\mathbf{A}$ per unit volume at the point in question.

Curl. The *curl* (or rotation) of a vector field $\mathbf{A}$ is the vector field

$$\operatorname{rot}\mathbf{A} \;\equiv\; \nabla\times\mathbf{A} = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ A_x & A_y & A_z \end{vmatrix}. \tag{1.1.4}$$

A vector field with $\operatorname{rot}\mathbf{A} \equiv \mathbf{0}$ throughout a simply connected region is called *irrotational*; it then admits a scalar potential Φ such that $\mathbf{A} = -\operatorname{grad}\Phi$.

Laplacian. The two operations div and grad may be composed to give the *Laplacian* of a scalar field,

$$\nabla^2 f \;\equiv\; \operatorname{div}(\operatorname{grad} f) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}. \tag{1.1.5}$$

This operator is the central object of the present book: every partial differential equation considered from Chapter 3 onwards contains ∇^2 either as its entire spatial part or as a component of it. By convention, $\nabla^2\mathbf{A}$ denotes the vector field whose Cartesian components are the Laplacians of the components of $\mathbf{A}$.

The Laplacian in curvilinear coordinates. Many of the boundary-value problems of mathematical physics are most naturally formulated in coordinates adapted to the geometry of the domain. The two most important such systems are the cylindrical coordinates (ρ,φ,z), related to the Cartesian coordinates by

$$x = \rho\cos\varphi, \qquad y = \rho\sin\varphi, \qquad z = z,$$

and the spherical coordinates (r,θ,φ), related to the Cartesian coordinates by

$$x = r\sin\theta\cos\varphi, \qquad y = r\sin\theta\sin\varphi, \qquad z = r\cos\theta.$$

A direct computation, which we omit, gives the Laplacian in the form

$$\nabla^2 f = \frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\,\frac{\partial f}{\partial\rho}\right) + \frac{1}{\rho^2}\frac{\partial^2 f}{\partial\varphi^2} + \frac{\partial^2 f}{\partial z^2} \qquad \text{(cylindrical)}, \tag{1.1.6}$$

$$\nabla^2 f = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\,\frac{\partial f}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\,\frac{\partial f}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 f}{\partial\varphi^2} \qquad \text{(spherical)}. \tag{1.1.7}$$

These formulas will be invoked repeatedly in Chapters 7–10, where the separation of variables in cylindrical and spherical geometries gives rise, respectively, to the Bessel and Legendre equations.

The integral theorems. Two integral identities are used throughout the derivations of Chapter 3. Let V be a bounded region of three-dimensional space with sufficiently smooth boundary ∂V, oriented by the outward unit normal $\hat{\mathbf{n}}$. Then *Gauss's theorem*

(also called the divergence theorem) asserts that, for every continuously differentiable vector field $\mathbf{A}$,

$$\int_V \operatorname{div}\mathbf{A}\, dV = \oint_{\partial V} \mathbf{A}\cdot\hat{\mathbf{n}}\, dS. \tag{1.1.8}$$

The corresponding statement for the curl is *Stokes's theorem*: if S is a smooth oriented surface with boundary curve ∂S oriented coherently with S, and $\mathbf{A}$ a continuously differentiable vector field, then

$$\int_S (\operatorname{rot}\mathbf{A})\cdot\hat{\mathbf{n}}\, dS = \oint_{\partial S} \mathbf{A}\cdot d\mathbf{l}. \tag{1.1.9}$$

A specialisation of Gauss's theorem to a scalar field f in the form $\mathbf{A} = \operatorname{grad} f$ produces *Green's first identity*, which underlies several of the uniqueness and existence arguments encountered in the sequel:

$$\int_V \left(g\,\nabla^2 f + \operatorname{grad} f\cdot\operatorname{grad} g \right) dV = \oint_{\partial V} g\,\frac{\partial f}{\partial n}\, dS, \tag{1.1.10}$$

where $\partial f/\partial n \equiv \hat{\mathbf{n}}\cdot\operatorname{grad} f$ is the outward normal derivative of f.

1.2 Linear Ordinary Differential Equations

The method of separation of variables, developed in Chapter 7, reduces a partial differential equation to a system of ordinary differential equations. The Sturm–Liouville theory of Chapter 9 and the eigenfunction expansions of Chapters 8 and 10 both rest upon a small body of classical results on linear ODEs, which we summarise here without proof.

The general second-order linear ODE. A second-order linear ordinary differential equation in the unknown $y(x)$ has the standard form

$$y'' + p(x)\,y' + q(x)\,y = r(x), \tag{1.2.1}$$

in which p, q, r are continuous functions on an interval $I \subseteq \mathbb{R}$. The equation is called *homogeneous* if $r \equiv 0$ and *inhomogeneous* otherwise. The basic existence-and-uniqueness result is the following.

Theorem 1.1 (Existence and uniqueness) *Let p, q, r be continuous on the interval I, and let $x_0 \in I$ and $y_0, y_1 \in \mathbb{R}$ be given. Then the initial-value problem*

$$y'' + p(x)\,y' + q(x)\,y = r(x), \qquad y(x_0) = y_0, \quad y'(x_0) = y_1,$$

admits a unique twice-continuously differentiable solution on I.

The space of solutions of the homogeneous equation $y'' + p\,y' + q\,y = 0$ is therefore a two-dimensional real vector space: once any pair of linearly independent solutions

y_1, y_2 has been exhibited, every solution can be written $y = c_1 y_1 + c_2 y_2$ for unique constants c_1, c_2.

The Wronskian. Linear independence of two solutions is conveniently tested by means of the *Wronskian determinant*

$$W(y_1, y_2)(x) \;\equiv\; \begin{vmatrix} y_1(x) & y_2(x) \\ y_1'(x) & y_2'(x) \end{vmatrix} = y_1 y_2' - y_2 y_1'. \tag{1.2.2}$$

A direct calculation, using the homogeneous equation to eliminate the second derivatives, yields *Abel's identity*

$$W(x) = W(x_0) \exp\left(-\int_{x_0}^{x} p(s)\, ds \right), \tag{1.2.3}$$

from which it follows that W either vanishes identically on I or nowhere on I. Consequently y_1 and y_2 are linearly independent on I if and only if $W(y_1, y_2)(x_0) \neq 0$ at any single point $x_0 \in I$.

Constant-coefficient equations. When the coefficients p, q in (1.2.1) are constants the equation is solved explicitly. Substituting $y = e^{\lambda x}$ into the homogeneous equation $y'' + py' + qy = 0$ leads to the *characteristic equation*

$$\lambda^2 + p\lambda + q = 0, \tag{1.2.4}$$

whose roots λ_1, λ_2 determine the general solution according to the following case analysis:

1. if $\lambda_1 \neq \lambda_2$ are both real, the general solution is $y = c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x}$;
2. if $\lambda_1 = \lambda_2 \equiv \lambda$ is a repeated real root, the general solution is $y = (c_1 + c_2 x)\, e^{\lambda x}$;
3. if $\lambda_{1,2} = \alpha \pm i\beta$ are complex conjugates with $\beta \neq 0$, the general solution is $y = e^{\alpha x}\big(c_1 \cos \beta x + c_2 \sin \beta x\big)$.

The case $p = 0$ deserves separate mention, since it occurs almost without exception in the chapters that follow. The equation

$$y'' + \omega^2 y = 0 \qquad (\omega > 0) \tag{1.2.5}$$

has, in view of case 3 above, the general solution $y = c_1 \cos \omega x + c_2 \sin \omega x$, while the equation

$$y'' - \omega^2 y = 0 \qquad (\omega > 0) \tag{1.2.6}$$

has, in view of case 1, the general solution $y = c_1 e^{\omega x} + c_2 e^{-\omega x}$, equivalently $y = c_1 \cosh \omega x + c_2 \sinh \omega x$. The borderline case $y'' = 0$ is solved by $y = c_1 + c_2 x$.

1.3 Trigonometric Fourier Series

The eigenfunction expansions central to Chapters 7 and 8 have, in the simplest cases, the form of a trigonometric Fourier series, with which the reader is presumed to be familiar. We collect here the basic facts.

Let f be a function defined on the interval $[-\ell,\ell]$, piecewise continuous together with its first derivative, and extended periodically to all of $\mathbb{R}$ with period 2ℓ. The *Fourier series* of f is the trigonometric series

$$\frac{a_0}{2} + \sum_{n=1}^{\infty}\left[a_n\cos\frac{n\pi x}{\ell} + b_n\sin\frac{n\pi x}{\ell}\right], \tag{1.3.1}$$

whose coefficients are obtained from the orthogonality relations

$$\int_{-\ell}^{\ell}\cos\frac{m\pi x}{\ell}\cos\frac{n\pi x}{\ell}\,dx = \ell\,\delta_{mn}, \quad \int_{-\ell}^{\ell}\sin\frac{m\pi x}{\ell}\sin\frac{n\pi x}{\ell}\,dx = \ell\,\delta_{mn},$$

$$\int_{-\ell}^{\ell}\cos\frac{m\pi x}{\ell}\sin\frac{n\pi x}{\ell}\,dx = 0$$

$(m,n = 1,2,\ldots)$ by means of the explicit formulas

$$a_n = \frac{1}{\ell}\int_{-\ell}^{\ell} f(x)\cos\frac{n\pi x}{\ell}\,dx, \quad b_n = \frac{1}{\ell}\int_{-\ell}^{\ell} f(x)\sin\frac{n\pi x}{\ell}\,dx. \tag{1.3.2}$$

The classical pointwise convergence theorem of Dirichlet asserts that, under the hypotheses stated above, the series (1.3.1) converges at every point $x\in\mathbb{R}$ to

$$\tfrac{1}{2}\big[f(x^+)+f(x^-)\big],$$

which equals $f(x)$ wherever f is continuous.

When the function is defined only on the half-interval $[0,\ell]$ and is extended either as an odd or as an even function to $[-\ell,\ell]$, the series (1.3.1) reduces, respectively, to the *Fourier sine series*

$$f(x) = \sum_{n=1}^{\infty} b_n\sin\frac{n\pi x}{\ell}, \qquad b_n = \frac{2}{\ell}\int_0^{\ell} f(x)\sin\frac{n\pi x}{\ell}\,dx, \tag{1.3.3}$$

or the *Fourier cosine series*

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n\cos\frac{n\pi x}{\ell}, \qquad a_n = \frac{2}{\ell}\int_0^{\ell} f(x)\cos\frac{n\pi x}{\ell}\,dx. \tag{1.3.4}$$

These two expansions arise, respectively, in problems with homogeneous Dirichlet and homogeneous Neumann boundary conditions on the interval $[0,\ell]$, as Chapter 7 demonstrates in detail.

1.4 The Special Functions of Mathematical Physics

The separation of the master equation in cylindrical and spherical coordinates, carried out in detail in Chapter 9, leads to two ordinary differential equations whose solutions are of fundamental importance in mathematical physics: the Bessel equation and the Legendre equation. We collect here, for ease of reference, their basic features. A complete treatment, including the proofs of orthogonality and completeness, is given in Section 9.6 and the worked examples of Chapter 9.

Bessel functions. The *Bessel equation of order n* is the second-order linear ODE

$$x^2 y'' + x y' + (x^2 - n^2)\, y = 0, \tag{1.4.1}$$

which arises whenever the Helmholtz equation $\nabla^2 u + \lambda u = 0$ is separated in cylindrical coordinates. Solving (1.4.1) by the method of Frobenius produces, for each $n \geq 0$, two linearly independent solutions: the *Bessel function of the first kind*,

$$J_n(x) \;=\; \sum_{k=0}^{\infty} \frac{(-1)^k}{k!\,\Gamma(n+k+1)} \left(\frac{x}{2}\right)^{n+2k}, \tag{1.4.2}$$

which is finite at $x = 0$, and the *Bessel function of the second kind* $Y_n(x)$ (also called the Neumann function), which has a logarithmic singularity at $x = 0$. The general solution of (1.4.1) is therefore

$$y(x) = c_1\, J_n(x) + c_2\, Y_n(x).$$

In every boundary-value problem that includes the axis of a cylindrical domain, regularity of the solution there forces $c_2 = 0$ (since Y_n diverges on the axis), leaving J_n as the relevant eigenfunction. The eigenvalues are then determined by the boundary condition imposed at the lateral surface of the cylinder; if, for instance, the function is required to vanish at radial distance a, the admissible eigenvalues are $\lambda_k = \xi_{n,k}/a$, where $\xi_{n,k}$ denotes the k-th positive zero of J_n.

The Bessel functions $J_n(\lambda_k x)$ form an orthogonal family on $[0, a]$ with respect to the weight function x,

$$\int_0^a x\, J_n(\lambda_k x)\, J_n(\lambda_j x)\, dx = \frac{a^2}{2}\left[J_n'(\lambda_k a)\right]^2 \delta_{kj}, \tag{1.4.3}$$

a fact which underlies the Fourier–Bessel expansion used in Chapters 9 and 10.

Legendre polynomials. The *Legendre equation* of order n is

$$(1 - x^2)\, y'' - 2x y' + n(n+1)\, y = 0, \tag{1.4.4}$$

which arises from the separation of the Helmholtz equation in spherical coordinates upon substituting $x = \cos\theta$. The Frobenius method yields a power-series solution that converges on $(-1, 1)$ but in general diverges at $x = \pm 1$. The series *terminates*, producing a polynomial solution finite on the closed interval $[-1, 1]$, precisely when n is a non-negative integer; the resulting polynomial of degree n, normalised so that $P_n(1) = 1$, is

the *Legendre polynomial*

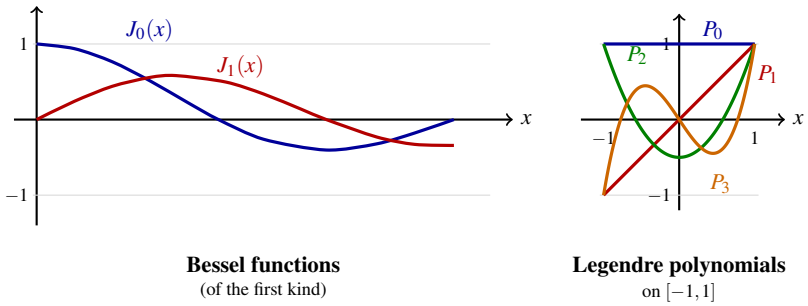
$$P_n(x) \;=\; \frac{1}{2^n\, n!}\,\frac{d^n}{dx^n}\big[(x^2-1)^n\big], \tag{1.4.5}$$

known as Rodrigues's formula. The first few are

$$P_0(x)=1, \quad P_1(x)=x, \quad P_2(x)=\tfrac{1}{2}(3x^2-1), \quad P_3(x)=\tfrac{1}{2}(5x^3-3x).$$

The polynomials P_n form an orthogonal family on $[-1,1]$ with respect to the unit weight,

$$\int_{-1}^{1} P_m(x)\,P_n(x)\,dx = \frac{2}{2n+1}\,\delta_{mn}. \tag{1.4.6}$$



Bessel functions
(of the first kind)

Legendre polynomials
on $[-1,1]$

Fig. 1.1 The standard families of special functions arising in mathematical physics. **Left**: the Bessel functions of the first kind $J_0(x)$ and $J_1(x)$, both regular at the origin and oscillatory beyond their first zero (the Bessel functions of the second kind Y_n, having a logarithmic singularity at $x=0$, are not shown). **Right**: the first four Legendre polynomials $P_n(x)$ on $[-1,1]$, normalised so that $P_n(1)=1$; the polynomial P_n has exactly n zeros in the open interval $(-1,1)$.

For problems that depend on the azimuthal angle φ, separation in spherical coordinates produces a generalisation of (1.4.4), the *associated Legendre equation*

$$(1-x^2)\,y'' - 2x\,y' + \left[n(n+1) - \frac{m^2}{1-x^2}\right] y = 0, \tag{1.4.7}$$

whose regular solutions on $[-1,1]$ are the *associated Legendre polynomials* $P_n^m(x)$ (m,n integers with $0 \le |m| \le n$). Like their unassociated counterparts, they form an orthogonal family on $[-1,1]$:

$$\int_{-1}^{1} P_n^m(x)\,P_k^m(x)\,dx = \frac{2}{2n+1}\,\frac{(n+m)!}{(n-m)!}\,\delta_{nk}. \tag{1.4.8}$$

These functions reappear repeatedly in Chapters 9 and 10 in the solution of boundary-value problems on spherical domains. For their derivation by the Frobenius method, and for the proofs of the orthogonality relations (1.4.3), (1.4.6) and (1.4.8), the reader is referred to Section 9.6.

Notational Conventions

The following conventions are observed throughout the book.

- Independent variables are written in lowercase Latin letters $(x, y, z, t, r, \rho, \theta, \varphi)$; the unknown function is generally denoted by u, with Ψ, Φ or f used where standard or convenient.
- Partial derivatives are denoted interchangeably by $\partial u/\partial x$ and u_x; analogously for higher derivatives, u_{xx}, u_{xy}, etc. The shorthand becomes increasingly useful as the algebraic content of the equations becomes denser.
- The Laplacian is uniformly written ∇^2; in three dimensions it coincides with (1.1.5), and the same symbol is retained, without ambiguity, for its lower- and higher-dimensional analogues.
- The d'Alembertian is written $\Box^2$ and is reserved for the relativistic wave operator introduced in the Introduction.
- Vectors are typeset in boldface (or with an arrow, where tradition prefers it, as in $\mathbf{K}$ and $\mathbf{r}$); their components are written without boldface and indexed by their Cartesian, cylindrical or spherical labels.
- Equations are numbered by section: equation $(N.k.\ell)$ is the ℓ-th numbered equation of Section k of Chapter N. References to equations of the same chapter omit the chapter number and write simply $(k.\ell)$ when no confusion can arise; references across chapters always carry the full label.
- The end of each definition, theorem, exercise or worked example is not marked. The end of each section's narrative is signalled by the appearance of the next section heading. Problems collected at the close of a chapter use Roman numerals, so that, for instance, *Problem V.3* is the third problem at the end of Chapter 7.

With these conventions in place we now turn, in Chapter 2, to the general theory of partial differential equations.

Chapter 2
General Properties of Partial Differential Equations

Abstract This opening chapter introduces the basic vocabulary used throughout the book: partial derivatives, order of an equation, linearity, homogeneity, and the general second-order linear PDE in n independent variables. The general solution of a PDE is shown to depend on arbitrary functions rather than constants. The chapter closes with linear differential operators and the principle of superposition for boundary-value problems—an instrument used over and over in subsequent chapters.

2.1 Fundamental Concepts

In physics, when a function characterising a phenomenon depends on more than one variable, one no longer speaks of an ordinary derivative with respect to a single variable but of partial derivatives with respect to all the variables on which it depends. For instance, if the temperature of a solid body is not uniform—varying both with time and from point to point—then the partial derivatives

$$\frac{\partial u}{\partial t}, \qquad \frac{\partial u}{\partial x}, \qquad \frac{\partial u}{\partial y}, \qquad \frac{\partial u}{\partial z}$$

of the temperature function u will, in general, all be different from zero. For physical phenomena, second-order—and occasionally higher-order—partial derivatives carry a definite physical meaning. In the typical workflow of mathematical physics, one constructs, from a model of the phenomenon under consideration, the equation satisfied by the chosen physical quantity (which depends on certain independent variables and plays a central role in the description); the resulting equation, together with appropriate initial and boundary conditions, constitutes the theory and is to be tested against experiment.

In most cases one obtains, between the function characterising the phenomenon, its partial derivatives, and the independent variables, a functional relation of the form

$$F\left(t, x, \ldots, \tfrac{\partial u}{\partial t}, \tfrac{\partial u}{\partial x}, \ldots, \tfrac{\partial^2 u}{\partial t^2}, \tfrac{\partial^2 u}{\partial t\,\partial x}, \tfrac{\partial^2 u}{\partial x^2}, \ldots\right) = 0.$$

The equation defined by this relation is called a *partial differential equation* (PDE).

Definition 2.1 (Order of a PDE) The order of a partial differential equation is the order of the highest-order partial derivative occurring in it.

For example,

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

is a second-order equation in three independent variables;

$$\left(\frac{\partial u}{\partial x}\right)^2 = \frac{\partial u}{\partial t}$$

is a first-order equation in two independent variables; and

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

is a first-order PDE in three independent variables.

The differential operator that constitutes the essence of a partial differential equation is, as we shall see in Sect. 2.2, called *linear* if it satisfies the linearity criterion. The PDE itself is then called *linear*. In general, a partial differential equation is said to be linear if its solution function u and the partial derivatives of u of every order appear only to the first power and no products of u with its derivatives (or of derivatives among themselves) occur; otherwise it is said to be *nonlinear*. For example,

$$\left(\frac{\partial^2 u}{\partial x^2}\right)^3 + u\frac{\partial u}{\partial y} = u^2$$

is nonlinear. An equation that is linear in its highest-order derivatives only is called *quasilinear*; for example,

$$\frac{\partial u}{\partial x}\frac{\partial^2 u}{\partial x^2} + x\frac{\partial u}{\partial y} = e^y$$

is a quasilinear equation.

In this book, as already mentioned, we shall be concerned mainly with linear and, most often, second-order partial differential equations. The most general second-order linear PDE in n independent variables, with the convention—admissible without loss of generality—that $A_{ij} = A_{ji}$, takes the form

$$\sum_{i,j=1}^{n} A_{ij}\frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^{n} B_i\frac{\partial u}{\partial x_i} + Cu = G, \tag{2.1.1}$$

where A_{ij}, B_i, C and G are functions of the variables x_i. Throughout the book the partial derivatives of u will be denoted briefly by

$$\frac{\partial u}{\partial x} \equiv u_x, \quad \frac{\partial u}{\partial y} \equiv u_y, \quad \frac{\partial^2 u}{\partial x^2} \equiv u_{xx}, \quad \frac{\partial^2 u}{\partial y^2} \equiv u_{yy}, \quad \frac{\partial^2 u}{\partial x \partial y} \equiv u_{xy}, \ldots$$

or, more generally,

$$\frac{\partial u}{\partial x_i} \equiv u_{x_i}, \qquad \frac{\partial^2 u}{\partial x_i^2} \equiv u_{x_i x_i}, \qquad \frac{\partial^2 u}{\partial x_i \partial x_j} \equiv u_{x_i x_j}, \ldots$$

With this notation, equation (2.1.1) reads

$$\sum_{i,j=1}^{n} A_{ij}\, u_{x_i x_j} + \sum_{i=1}^{n} B_i\, u_{x_i} + Cu = G. \tag{2.1.2}$$

Definition 2.2 (Homogeneous and inhomogeneous PDEs) A PDE is said to be *inhomogeneous* if, apart from terms involving the unknown u and partial derivatives of u of any order, a function G as in (2.1.2) is present; otherwise it is said to be *homogeneous*.

It is well known that the general solution of an nth-order ordinary differential equation is a family of functions containing n independent arbitrary constants, to be fixed by suitable conditions. For PDEs, by contrast, the general solution depends on arbitrary *functions* rather than arbitrary constants. As an illustration, consider the second-order equation

$$u_{xy} = 0.$$

Holding x fixed and integrating with respect to y, with $F(x)$ an arbitrary function of x,

$$u_x(x,y) = F(x).$$

Integrating once more with respect to x, with $f(x)$ and $g(y)$ arbitrary functions, the solution is

$$u(x,y) = \int F(x)\,dx + g(y) = f(x) + g(y).$$

Below we present worked exercises on PDEs that may be solved by direct integration in this manner.

Exercise 2.1 Find the general solution of the equation $u_{xy} + u_x = 0$.

Solution. With the substitution $u_x = v$ the equation becomes

$$v_y + v = 0.$$

Rearranged this reads

$$\frac{dv}{v} = -dy,$$

and integration gives, with $\varphi(x)$ denoting an arbitrary function,

$$v(x,y) = u_x = \varphi(x)\, e^{-y}.$$

Integrating this last equation with respect to x, with $g(y)$ a further arbitrary function and $f(x) = \int \varphi(x)\,dx$, the general solution containing two arbitrary functions is

$$u(x,y) = e^{-y} f(x) + g(y).$$

Exercise 2.2 Let a and b be constants. Find the general solution of the equation $u_{xy} = a\,u_y + b$.

Solution. With the substitution $u_y = V(x,y)$ the equation becomes

$$V_x = aV + b,$$

whence

$$\frac{dV}{aV+b} = dx.$$

Integrating, with $\psi(y)$ an arbitrary function of y playing the role of the integration "constant",

$$\frac{1}{a}\ln\left|aV+b\right| = x + \ln\left|\psi(y)\right|,$$

so that

$$V \equiv u_y = \psi(y)\,e^{ax} - \frac{b}{a}.$$

Integrating this expression with respect to y, with $f(x)$ a further arbitrary function and $g(y) = \int \psi(y)\,dy$, the general solution of the equation is

$$u(x,y) = f(x) + e^{ax}g(y) - \frac{b}{a}y.$$

Exercise 2.3 Let a be a constant. Find the general solution of the equation

$$u_{xy} - \frac{1}{ax}u_y = 0.$$

Solution. With the substitution $u_y = V(x,y)$ the equation becomes

$$V_x - \frac{1}{ax}V = 0,$$

which on rearrangement reads

$$\frac{dV}{V} = \frac{dx}{ax}.$$

Integration gives, with $\varphi(y)$ an arbitrary function,

$$V(x,y) \equiv u_y = \varphi(y)\,x^{1/a},$$

and a further integration with respect to y, with $f(x)$ an arbitrary function and $g(y) = \int \varphi(y)\,dy$, gives the general solution

$$u(x,y) = f(x) + g(y)\,x^{1/a}.$$

Exercise 2.4 Find the general solution of $u_{xx} = 0$.

Solution. Integrating once with respect to x, with $f(y)$ an arbitrary function,

$$u_x = f(y).$$

A further integration with respect to x, with $g(y)$ an arbitrary function, gives the general solution

$$u(x,y) = x f(y) + g(y).$$

Exercise 2.5 Find the general solution of $u_{yy} + u_y = 0$.

Solution. With the substitution $u_y \equiv v$ the equation becomes $v_y + v = 0$, which on integration gives, with $f(x)$ an arbitrary function,

$$v(x,y) \equiv u_y = f(x)\, e^{-y}.$$

A further integration with respect to y then yields, with $g(x)$ an arbitrary function, the general solution

$$u(x,y) = g(x) - e^{-y} f(x).$$

Exercise 2.6 Find the general solution of $u_{yy} + u = 0$.

Solution. With $F(x)$ and $G(y)$ arbitrary functions, set

$$u(x,y) = F(x)\, G(y).$$

Substituting this expression into the equation gives

$$F(x)\left(\frac{d^2G}{dy^2} + G\right) = 0,$$

and since $F(x)$ is an arbitrary function this reduces to

$$\frac{d^2G}{dy^2} + G = 0,$$

whose general solution is, with A and B constants,

$$G(y) = A\cos y + B\sin y.$$

The general solution of the given PDE is therefore, on writing $f(x) = A\,F(x)$ and $g(x) = B\,F(x)$ for two arbitrary functions,

$$u(x,y) = f(x)\cos y + g(x)\sin y.$$

2.2 Linear Differential Operators and the Principle of Superposition

An operator represents an operation, governed by definite rules, which when applied to a function produces another function. For example, a differential operator may have any of the following forms:

$$
\begin{aligned}
L[u] &= \frac{\partial u}{\partial x} + \frac{\partial^2 u}{\partial x\,\partial y} + \frac{\partial^3 u}{\partial y^3},\\
M[u] &= \frac{\partial^2 u}{\partial x^2} - x^2\,\frac{\partial^2 u}{\partial y^2},\\
P[u] &= \int_a^b F(x,t)\,u(x,t)\,dt,\\
Q[u] &= u(x,C) + u_t(x,C).
\end{aligned}
$$

The corresponding bare operators read

$$
L \equiv \frac{\partial}{\partial x} + \frac{\partial^2}{\partial x\,\partial y} + \frac{\partial^3}{\partial y^3}, \qquad M \equiv \frac{\partial^2}{\partial x^2} - x^2\,\frac{\partial^2}{\partial y^2}.
$$

Definition 2.3 (Linear operator) An operator is called a *linear operator* if it satisfies the following relations.

1. A constant C may be taken outside the operator,

$$
L[Cu] = C\,L[u], \tag{2.2.1}
$$

   and
2. the result of applying the operator to the sum of two functions equals the sum of the results of applying the operator to each of the functions separately, that is,

$$
L[u_1 + u_2] = L[u_1] + L[u_2]. \tag{2.2.2}
$$

Combining (2.2.1) and (2.2.2) gives

$$
L[C_1 u_1 + C_2 u_2] = C_1\,L[u_1] + C_2\,L[u_2]. \tag{2.2.3}
$$

For a finite number of functions, expression (2.2.3) generalises to

$$
L\left[\sum_{j=1}^{k} C_j\,u_j\right] = \sum_{j=1}^{k} C_j\,L[u_j]. \tag{2.2.4}
$$

On the other hand, if L and M are two linear operators, then their sum, defined by

$$
(L+M)[u] = L[u] + M[u],
$$

is easily seen to be a linear operator as well. Their product, which is the result of applying these operators in succession to the function u, namely

$$
LM[u] = L\big[M[u]\big],
$$

is again a linear operator, as one easily verifies by means of (2.2.4).

The following relations hold for linear differential operators:

1. $L+M = M+L$ (commutativity);

2. $(L+M)+N = L+(M+N)$ (associativity);
3. $(LM)N = L(MN)$ (associativity);
4. $L(M+N) = LM+LN$ (distributivity);

and, in addition, for linear differential operators with *constant* coefficients,

5. $LM = ML$ (commutativity).

For variable-coefficient linear differential operators one usually has $LM \neq ML$. For example, with the operators

(i) $L = \dfrac{\partial^2}{\partial x^2} + x\dfrac{\partial}{\partial y}, \quad M = \dfrac{\partial^2}{\partial y^2} + y\dfrac{\partial}{\partial y};$

(ii) $L = x\dfrac{\partial^2}{\partial x^2} + y, \quad M = x\dfrac{\partial}{\partial x} - \dfrac{\partial^2}{\partial y^2};$

(iii) $L = \dfrac{\partial^2}{\partial y^2} - x\dfrac{\partial}{\partial y}, \quad M = \dfrac{\partial^2}{\partial x^2} + y\dfrac{\partial}{\partial y};$

(iv) $L = y\dfrac{\partial^2}{\partial y^2} - x\dfrac{\partial}{\partial y}, \quad M = x\dfrac{\partial^2}{\partial x^2} + y\dfrac{\partial}{\partial y},$

one verifies that $LM[u] \neq ML[u]$.

Now consider a second-order linear PDE in two independent variables. Such an equation, in its most general form,

$$A(x,y)\,u_{xx} + B(x,y)\,u_{xy} + C(x,y)\,u_{yy} + D(x,y)\,u_x + E(x,y)\,u_y + F(x,y)\,u = G(x,y),$$

may be written, with the operator

$$L \equiv A\frac{\partial^2}{\partial x^2} + B\frac{\partial^2}{\partial x\,\partial y} + C\frac{\partial^2}{\partial y^2} + D\frac{\partial}{\partial x} + E\frac{\partial}{\partial y} + F,$$

in the form

$$L[u] = G,$$

or simply

$$Lu = G.$$

Principle of Superposition (Introduction to Initial-Value Problems). The auxiliary conditions for a boundary-value problem may also be expressed in the operator notation introduced above. For example, the initial-value problem

$$\begin{aligned} u_{tt} - c^2 u_{xx} &= G(x,t), && 0 \le x \le a,\ t > 0, \\ u(x,0) &= g_1(x), && 0 \le x \le a, \\ u_t(x,0) &= g_2(x), && 0 \le x \le a, \\ u(0,t) &= g_3(t), && t > 0, \\ u(a,t) &= g_4(t), && t > 0, \end{aligned}$$

becomes, in operator notation,

$$\begin{aligned} L[u] &= G, \\ M_1[u] &= g_1, \\ M_2[u] &= g_2, \\ M_3[u] &= g_3, \\ M_4[u] &= g_4. \end{aligned}$$

More generally, a linear boundary-value problem whose underlying PDE involves an arbitrary number of independent variables and is of arbitrary order, with $L, M_1, M_2, \ldots, M_n$ all linear operators, may be written

$$\left.\begin{aligned} L[u] &= G, \\ M_1[u] &= g_1, \\ M_2[u] &= g_2, \\ &\vdots \\ M_n[u] &= g_n. \end{aligned}\right\} \tag{2.2.5}$$

By the linearity of the equation and of the auxiliary conditions, problem (2.2.5) can be decomposed into the following sequence of subproblems:

$$\begin{array}{ccccc} L[u_0] = G & L[u_1] = 0 & L[u_2] = 0 & \cdots & L[u_n] = 0 \\ M_1[u_0] = 0 & M_1[u_1] = g_1 & M_1[u_2] = 0 & \cdots & M_1[u_n] = 0 \\ M_2[u_0] = 0 & M_2[u_1] = 0 & M_2[u_2] = g_2 & \cdots & M_2[u_n] = 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ M_n[u_0] = 0 & M_n[u_1] = 0 & M_n[u_2] = 0 & \cdots & M_n[u_n] = g_n. \end{array}$$

In this case it is easily seen that, $u_0, u_1, u_2, \ldots, u_n$ denoting the solutions of the subproblems indicated above, the general solution of problem (2.2.5) is

$$u = \sum_{i=0}^{n} u_i \qquad (i = 0, 1, 2, \ldots, n).$$

Take, as an example, one of the subproblems considered above—the problem satisfied by u_1. Suppose that, with φ_k $(k = 1, 2, \ldots, n)$ denoting (finitely or infinitely many) solutions of the homogeneous system

$$\begin{aligned} L[\varphi_k] &= 0, \\ M_2[\varphi_k] &= 0, \\ M_3[\varphi_k] &= 0, \\ &\vdots \\ M_n[\varphi_k] &= 0, \end{aligned}$$

we have exhibited the constant coefficients $C_1, C_2, \ldots$ satisfying

$$g_1 = M_1[C_1\varphi_1 + C_2\varphi_2 + \cdots] = C_1M_1[\varphi_1] + C_2M_1[\varphi_2] + \cdots = M_1\left[\sum_k C_k\varphi_k\right].$$

The solution of the problem is then

$$u_1 = C_1\varphi_1 + C_2\varphi_2 + \cdots = \sum_k C_k\varphi_k.$$

When this linear combination contains infinitely many terms, the resulting infinite series must be uniformly convergent and term-by-term differentiable sufficiently many times. Furthermore, on setting

$$g_0 \equiv G, \qquad N_0 \equiv L, \qquad N_j \equiv M_j \quad (j = 1,2,\ldots,n),$$

problem (2.2.5) may be written in the form

$$N_k[u] = g_k \qquad (k = 0,1,2,\ldots,n).$$

For (2.2.5) to admit a solution of this superposition form, all of the series $N_k[u_i]$ $(i,k = 0,1,2,\ldots,n)$ that arise in the subproblems must be uniformly convergent.

Problems

2.1 Determine, for each of the following equations, whether it is linear, what its order is, and whether it is homogeneous:

$$\begin{aligned}
&U_{yy} + yU_x = x,\\
&UU_y + \frac{1}{x}U_x = 0,\\
&UU_x - (U_y)^2 = e^x,\\
&U_{xxx} + U_{xyy} + U_{xxy} + U_{yyy} = 1,\\
&U_{xx} + 2U_{xy} + U_{yy} = \tan x,\\
&U_{xx} + U_x + \ln u = 0.
\end{aligned}$$

2.2 Show that the functions $u(x,y) = x^2 - y^2$ and $u(x,y) = e^x \sin y$ are both solutions of the equation

$$u_{xx} + u_{yy} = 0.$$

2.3 Show that the functions $u(x,y) = (x+y)^3$ and $u(x,y) = \sin(x-y)$ are both solutions of the equation

$$u_{xx} - u_{yy} = 0.$$

2.4 Let f denote a differentiable function. Show that $u(x,y) = f(xy)$ satisfies the equation

$$xu_x - yu_y = 0,$$

and verify that the functions $\sin(xy)$, $\cos(xy)$, $\log(xy)$, e^{xy} and $(xy)^3$ are all solutions of the equation above.

2.5 Let f and g denote arbitrary twice-differentiable functions. Show that the function $u(x,y) = f(x)\,g(y)$ satisfies the equation

$$u\,u_{xy} - u_x\,u_y = 0,$$

and determine whether this equation is linear.

2.6 Let L be a linear partial differential operator. If u_1 is a solution of $L[u_1] = f$ and u_2 is a solution of $L[u_2] = 0$, show that $u_1 + u_2$ satisfies the equation $L[u] = f$.

2.7 A simple first-order *nonlinear* equation is given in the form

$$\frac{dy}{dx} + \omega y^n = 0 \qquad (n = 2,3,4,\ldots).$$

With k a constant, show that the function

$$y(x;k) = \left[(n-1)x + k\right]^{-1/(n-1)}$$

is the general solution of the equation when $\omega = \dfrac{1}{n-1}$, but that

$$y(x;k_1) + y(x;k_2)$$

does *not* satisfy the equation. Write down the solution of the equation in the case $n = 1$.

Chapter 3
Mathematical Models and Partial Differential Equations

Abstract We derive, from physical first principles, the governing PDEs for six canonical phenomena of mathematical physics: the vibrating string and membrane, waves in elastic media, heat conduction, the gravitational potential, and sound waves in a gas. The derivations use Newton's laws, Hooke's law, the divergence theorem and small-disturbance linearisation.

We wish to give a few concrete examples of phenomena whose mathematical description leads to partial differential equations, that is, of physical problems that can be modelled only by means of such equations.

3.1 Transverse Vibrations of an Elastic String

Consider, as in Fig. 3.1, an elastic homogeneous string of initial length $\ell = \overline{00'}$ stretched along the x-axis between the points 0 and $0'$ and held under uniform tension. Suppose the string is given a small perturbation in the y-direction (e.g. by being struck), and let $y(x,t)$ denote, t seconds later, the transverse displacement of the string from its equilibrium position. We wish to find the equation of motion governing this displacement. To this end we shall adopt the following experimentally suggested simplifying assumptions:

1. The string is elastic, so that, by definition of elasticity, the tensile force at every point of the string acts along the local tangent to the string.
2. The tensile force T at every point of the string is constant.
3. The weight of the string is negligible compared with the tension.
4. The departure of the perturbed string from its equilibrium configuration is small compared with the length of the string.
5. The slope of the perturbed string at every point is sufficiently small.
6. Only transverse vibrations occur.

From Fig. 3.1 the components of the tensile force T at the two endpoints of the element AB are seen to be

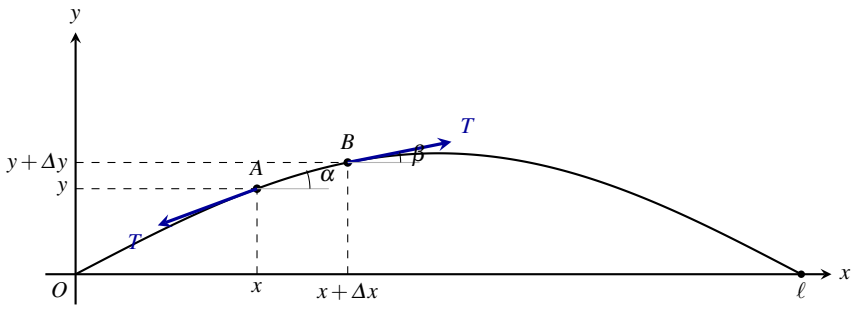


Fig. 3.1 Element AB of the displaced string. The arrows indicate the tensions T exerted on the element by the rest of the string at A and B, directed tangentially outward; α and β are the corresponding angles with the horizontal at x and $x+\Delta x$.

$$\begin{aligned} T_y\,(\text{at } x) &= -T\sin\alpha, \\ T_x\,(\text{at } x) &= -T\cos\alpha, \\ T_{y+\Delta y}\,(\text{at } x+\Delta x) &= T\sin\beta, \\ T_{x+\Delta x}\,(\text{at } x+\Delta x) &= T\cos\beta. \end{aligned}$$

Accordingly the mass of the element AB is $\Delta m \simeq \rho\,\Delta x$. The resultant of the forces acting on AB in the x-direction (longitudinal vibrations being excluded) is

$$T_{x+\Delta x} + T_x = T(\cos\beta - \cos\alpha) \simeq 0 \qquad (\beta \ll 1,\ \alpha \ll 1),$$

while the resultant in the y-direction is, with $\tan\beta = (\partial y/\partial x)_{x+\Delta x}$ and $\tan\alpha = (\partial y/\partial x)_x$,

$$\begin{aligned} T_{y+\Delta y} + T_y &= T(\sin\beta - \sin\alpha) \simeq T(\tan\beta - \tan\alpha) \\ &= T\left[\left(\frac{\partial y}{\partial x}\right)_{x+\Delta x} - \left(\frac{\partial y}{\partial x}\right)_x\right]. \end{aligned} \tag{3.1.1}$$

Newton's second law together with (3.1.1) now gives

$$T\left[\left(\frac{\partial y}{\partial x}\right)_{x+\Delta x} - \left(\frac{\partial y}{\partial x}\right)_x\right] = \Delta m\,\frac{\partial^2 y}{\partial t^2} = \rho\,\Delta x\,\frac{\partial^2 y}{\partial t^2}, \tag{3.1.2}$$

and hence

$$T\left[\frac{(\partial y/\partial x)_{x+\Delta x} - (\partial y/\partial x)_x}{\Delta x}\right] = \rho\,\frac{\partial^2 y}{\partial t^2}.$$

Passing to the limit $\Delta x \to 0$ one obtains

$$T\,\frac{\partial^2 y}{\partial x^2} = \rho\,\frac{\partial^2 y}{\partial t^2},$$

or, on writing $c = \sqrt{T/\rho}$,

$$y_{tt} = c^2\,y_{xx}, \tag{3.1.3}$$

the (one-dimensional) wave equation. If, in addition, an external force per unit length f acts on the string, the equation of motion takes the form

$$y_{tt} = c^2 y_{xx} + f^*, \qquad f^* = \frac{f}{\rho}. \tag{3.1.4}$$

3.2 The Vibrating Membrane

Before deriving the equation of the vibrating membrane we list, as before, the points to be taken into account and the simplifying assumptions:

1. The membrane is elastic, so that the tensile force at every point of the membrane lies in the local tangent plane to the membrane.
2. The tensile force T at every point of the membrane is constant.
3. The weight of the membrane is negligible compared with the tension.
4. The departure of the perturbed membrane from its equilibrium configuration is small compared with the dimensions of the membrane.
5. The slope of the perturbed membrane at every point is sufficiently small.
6. Only transverse vibrations occur.

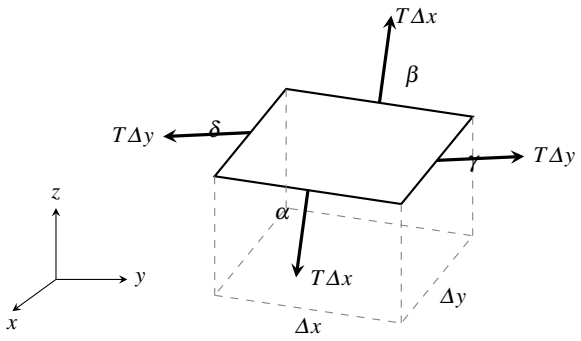


Fig. 3.2 Surface element $\Delta S \simeq \Delta x \Delta y$ of the displaced membrane. Each Δx-edge carries an outward tangential tension of magnitude $T\Delta x$, and each Δy-edge a tension $T\Delta y$. The angles α, β are made by the two Δx-edges with the horizontal (x,y)-plane, and γ, δ by the two Δy-edges.

Take, as in Fig. 3.2, a surface element of area $\Delta S \simeq \Delta x \Delta y$. If T denotes the tensile force per unit length, the forces acting on the edges of the element are $T\Delta x$ and $T\Delta y$. The vertical resultant of the forces acting on the surface element is therefore seen, with the help of Fig. 3.2, to be

$$T\Delta x \sin\beta - T\Delta x \sin\alpha + T\Delta y \sin\delta - T\Delta y \sin\gamma.$$

Since the angles $\alpha, \beta, \gamma, \delta$ are small, their sines may be replaced by their tangents, so that the above expression becomes

$$T\Delta x(\tan\beta-\tan\alpha)+T\Delta y(\tan\delta-\tan\gamma). \tag{3.2.1}$$

By Newton's law the resultant force equals mass times acceleration; with the help of (3.2.1) we obtain

$$T\Delta x(\tan\beta-\tan\alpha)+T\Delta y(\tan\delta-\tan\gamma)=\rho\,\Delta S\frac{\partial^2 z}{\partial t^2}. \tag{3.2.2}$$

On the other hand, with x_1,x_2 values of x between x and $x+\Delta x$ and y_1,y_2 values of y between y and $y+\Delta y$,

$$\tan\alpha=z_y(x_1,y)=\left(\frac{\partial z}{\partial y}\right)_{x_1,y},\quad \tan\beta=z_y(x_2,y+\Delta y)=\left(\frac{\partial z}{\partial y}\right)_{x_2,y+\Delta y},$$
$$\tan\gamma=z_x(x,y_1)=\left(\frac{\partial z}{\partial x}\right)_{x,y_1},\quad \tan\delta=z_x(x+\Delta x,y_2)=\left(\frac{\partial z}{\partial x}\right)_{x+\Delta x,y_2},$$

so that (3.2.2) reads

$$T\Delta x\big[z_y(x_2,y+\Delta y)-z_y(x_1,y)\big]+T\Delta y\big[z_x(x+\Delta x,y_2)-z_x(x,y_1)\big]=\rho\,\Delta x\Delta y\frac{\partial^2 z}{\partial t^2}.$$

Dividing by $\rho\,\Delta x\Delta y$ gives

$$\frac{T}{\rho}\left[\frac{z_y(x_2,y+\Delta y)-z_y(x_1,y)}{\Delta y}+\frac{z_x(x+\Delta x,y_2)-z_x(x,y_1)}{\Delta x}\right]=\frac{\partial^2 z}{\partial t^2}.$$

With $c=\sqrt{T/\rho}$, in the limit $\Delta S\to 0$ ($\Delta x,\Delta y\to 0$),

$$z_{tt}=c^2\big(z_{xx}+z_{yy}\big), \tag{3.2.3}$$

the two-dimensional wave equation. If the membrane is, in addition, under the action of an external force, the equation of motion—written with f denoting the external force per unit area—becomes

$$z_{tt}=c^2\big(z_{xx}+z_{yy}\big)+f^*,\qquad f^*=\frac{f}{\rho}. \tag{3.2.4}$$

3.3 Waves in an Elastic Medium

Consider a perturbation originating at a point of an elastic medium. Neighbouring points then begin to undergo a definite motion, and the medium is placed under stress. We shall assume that the resulting displacements are sufficiently small, and consider the medium as a whole, ignoring its rotational and translational rigid motions. The medium is also taken to be homogeneous and isotropic. Within the elastic medium, take a volume element ΔV as in Fig. 3.3 and denote the stress components acting on the faces of the cubic element (i.e. the forces per unit face area) by

$\tau_{xx}, \tau_{yy}, \tau_{zz}, \tau_{xy}, \tau_{xz}, \tau_{yx}, \tau_{yz}, \tau_{zx}, \tau_{zy}$. The first three are called *normal stresses*, the others *shear stresses*. We assume in addition that

$$\tau_{ij} = \tau_{ji}, \qquad i \neq j, \quad (i, j = x, y, z), \tag{3.3.1}$$

i.e. that the stress tensor τ_{ij} is symmetric.

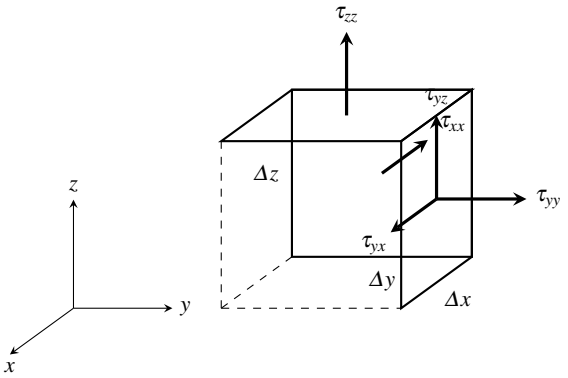


Fig. 3.3 Volume element $\Delta V = \Delta x \Delta y \Delta z$ of the elastic medium. On every face whose outward normal is $+\hat{e}_i$ act three stress components τ_{ij} (one normal, two shear); the figure shows all three on the visible $+\hat{y}$-face, and only the normal stresses τ_{xx}, τ_{zz} on the other two visible faces, to avoid clutter.

Neglecting external (body) forces, the resultant of the forces acting on the volume element in the x-direction is

$$\big[(\tau_{xx})_{x+\Delta x} - (\tau_{xx})_x\big]\Delta y \Delta z + \big[(\tau_{yx})_{y+\Delta y} - (\tau_{yx})_y\big]\Delta x \Delta z + \big[(\tau_{zx})_{z+\Delta z} - (\tau_{zx})_z\big]\Delta x \Delta y. \tag{3.3.2}$$

Now, with $u = u(x, y, z)$ the x-component of the displacement of a point of the elastic medium, Newton's law gives

$$\begin{aligned} &\big[(\tau_{xx})_{x+\Delta x} - (\tau_{xx})_x\big]\Delta y \Delta z + \big[(\tau_{yx})_{y+\Delta y} - (\tau_{yx})_y\big]\Delta x \Delta z \\ &\quad + \big[(\tau_{zx})_{z+\Delta z} - (\tau_{zx})_z\big]\Delta x \Delta y = \rho \, \Delta x \Delta y \Delta z \, \frac{\partial^2 u}{\partial t^2}. \end{aligned} \tag{3.3.3}$$

In the limit $\Delta V \to 0$, taking into account the symmetry $\tau_{ij} = \tau_{ji}$,

$$\frac{\partial \tau_{xx}}{\partial x} + \frac{\partial \tau_{xy}}{\partial y} + \frac{\partial \tau_{xz}}{\partial z} = \rho \, \frac{\partial^2 u}{\partial t^2}. \tag{3.3.4}$$

In the same way, by considering the resultant forces acting on ΔV in the y- and z-directions and the corresponding displacements υ, w, one obtains

$$\frac{\partial \tau_{yx}}{\partial x} + \frac{\partial \tau_{yy}}{\partial y} + \frac{\partial \tau_{yz}}{\partial z} = \rho \frac{\partial^2 \upsilon}{\partial t^2}, \tag{3.3.5}$$

$$\frac{\partial \tau_{zx}}{\partial x} + \frac{\partial \tau_{zy}}{\partial y} + \frac{\partial \tau_{zz}}{\partial z} = \rho \frac{\partial^2 w}{\partial t^2}. \tag{3.3.6}$$

Now define the so-called *linear strains*

$$\left.\begin{aligned} \varepsilon_{xx} &= \frac{\partial u}{\partial x}, \quad \varepsilon_{xy} = \varepsilon_{yx} = \frac{\partial u}{\partial y} + \frac{\partial \upsilon}{\partial x}, \\ \varepsilon_{yy} &= \frac{\partial \upsilon}{\partial y}, \quad \varepsilon_{xz} = \varepsilon_{zx} = \frac{\partial u}{\partial z} + \frac{\partial w}{\partial x}, \\ \varepsilon_{zz} &= \frac{\partial w}{\partial z}, \quad \varepsilon_{yz} = \varepsilon_{zy} = \frac{\partial \upsilon}{\partial z} + \frac{\partial w}{\partial y}. \end{aligned}\right\} \tag{3.3.7}$$

The diagonal quantities $\varepsilon_{xx}, \varepsilon_{yy}, \varepsilon_{zz}$ represent unit elongations and the off-diagonal ones $\varepsilon_{xy}, \varepsilon_{xz}, \varepsilon_{yz}$ represent unit shears.

For an isotropic body the generalised Hooke's law expresses the stresses in terms of the strains by means of two constants λ, μ called the Lamé constants and the dilatation $\theta = \varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz}$:

$$\left.\begin{aligned} \tau_{xx} &= \lambda\theta + 2\mu\varepsilon_{xx}, \quad \tau_{xy} = \mu\varepsilon_{xy}, \\ \tau_{yy} &= \lambda\theta + 2\mu\varepsilon_{yy}, \quad \tau_{xz} = \mu\varepsilon_{xz}, \\ \tau_{zz} &= \lambda\theta + 2\mu\varepsilon_{zz}, \quad \tau_{yz} = \mu\varepsilon_{yz}. \end{aligned}\right\} \tag{3.3.8}$$

To substitute (3.3.8) into (3.3.4) we first rewrite, with the aid of (3.3.7), $\tau_{xx}, \tau_{xy}, \tau_{xz}$ in terms of the ε_{ij}'s,

$$\tau_{xx} = \lambda\theta + 2\mu \frac{\partial u}{\partial x}, \quad \tau_{xy} = \mu\left(\frac{\partial u}{\partial y} + \frac{\partial \upsilon}{\partial x}\right), \quad \tau_{xz} = \mu\left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x}\right). \tag{3.3.9}$$

Differentiating these with respect to x, y, z respectively,

$$\left.\begin{aligned} \frac{\partial \tau_{xx}}{\partial x} &= \lambda \frac{\partial \theta}{\partial x} + 2\mu \frac{\partial^2 u}{\partial x^2}, \\ \frac{\partial \tau_{xy}}{\partial y} &= \mu\left(\frac{\partial^2 \upsilon}{\partial x \partial y} + \frac{\partial^2 u}{\partial y^2}\right), \\ \frac{\partial \tau_{xz}}{\partial z} &= \mu\left(\frac{\partial^2 w}{\partial x \partial z} + \frac{\partial^2 u}{\partial z^2}\right). \end{aligned}\right\} \tag{3.3.10}$$

Substituting (3.3.10) into (3.3.4),

$$\lambda \frac{\partial \theta}{\partial x} + \mu\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}\right) + \mu\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 \upsilon}{\partial x \partial y} + \frac{\partial^2 w}{\partial x \partial z}\right) = \rho \frac{\partial^2 u}{\partial t^2}. \tag{3.3.11}$$

Now set

$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+\frac{\partial^2 u}{\partial z^2}=\nabla^2 u, \tag{3.3.12}$$

and recognise

$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 \upsilon}{\partial x\,\partial y}+\frac{\partial^2 w}{\partial x\,\partial z}=\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial x}+\frac{\partial \upsilon}{\partial y}+\frac{\partial w}{\partial z}\right)=\frac{\partial \theta}{\partial x}. \tag{3.3.13}$$

With (3.3.12), (3.3.13) substituted into (3.3.11) we obtain

$$\lambda\,\frac{\partial \theta}{\partial x}+\mu\,\nabla^2 u+\mu\,\frac{\partial \theta}{\partial x}=(\lambda+\mu)\,\frac{\partial \theta}{\partial x}+\mu\,\nabla^2 u=\rho\,\frac{\partial^2 u}{\partial t^2}. \tag{3.3.14}$$

The same procedure applied to (3.3.5) and (3.3.6) yields the corresponding equations

$$(\lambda+\mu)\,\frac{\partial \theta}{\partial y}+\mu\,\nabla^2 \upsilon=\rho\,\frac{\partial^2 \upsilon}{\partial t^2}, \tag{3.3.15}$$

$$(\lambda+\mu)\,\frac{\partial \theta}{\partial z}+\mu\,\nabla^2 w=\rho\,\frac{\partial^2 w}{\partial t^2}. \tag{3.3.16}$$

Setting $\mathbf{u}=u\,\hat{e}_1+\upsilon\,\hat{e}_2+w\,\hat{e}_3$ and $\theta=\operatorname{div}\mathbf{u}$, equations (3.3.14)–(3.3.16) can be assembled, with the help of vector notation, into the single PDE

$$(\lambda+\mu)\,\operatorname{grad}\operatorname{div}\mathbf{u}+\mu\,\nabla^2\mathbf{u}=\rho\,\frac{\partial^2\mathbf{u}}{\partial t^2}. \tag{3.3.17}$$

Special cases.

(i) If $\operatorname{div}\mathbf{u}=0$, then with $c=\sqrt{\mu/\rho}$ equation (3.3.17) reduces to

$$\mathbf{u}_{tt}=c^2\,\nabla^2\mathbf{u}, \tag{3.3.18}$$

the wave equation governing transverse, equivoluminal waves.

(ii) If $\operatorname{rot}\mathbf{u}=\mathbf{0}$, the vector identity

$$\operatorname{rot}\operatorname{rot}\mathbf{u}=\operatorname{grad}\operatorname{div}\mathbf{u}-\nabla^2\mathbf{u}$$

gives $\operatorname{grad}\operatorname{div}\mathbf{u}=\nabla^2\mathbf{u}$, so that equation (3.3.17) becomes

$$(\lambda+2\mu)\,\nabla^2\mathbf{u}=\rho\,\frac{\partial^2\mathbf{u}}{\partial t^2},$$

or, on writing $c=\sqrt{(\lambda+2\mu)/\rho}$,

$$\mathbf{u}_{tt}=c^2\,\nabla^2\mathbf{u}. \tag{3.3.19}$$

Since here $\operatorname{rot}\mathbf{u}=\mathbf{0}$, the equation describes an irrotational (longitudinal) wave.

3.4 Heat Conduction in Solids

Consider a region B^* bounded by a closed surface S^*, and let $u(x,y,z,t)$ denote the temperature at the point (x,y,z) at time t. If the temperature is non-uniform, a heat current is established in the medium directed from points of higher temperature to points of lower temperature. Since the heat-current density is known to be proportional to the gradient of the temperature, the heat-current density vector $\mathbf{j}$ giving the heat energy crossing in unit time a unit area perpendicular to the direction of flow is, with χ the thermal conductivity, given by Fourier's law

$$\mathbf{j} = -\chi \operatorname{grad} u. \tag{3.4.1}$$

Now consider, inside the region B^*, an arbitrary region B bounded by a closed surface S. The heat $j_n = \mathbf{j}\cdot\mathbf{n}$ leaving B per unit time, where $\mathbf{n}$ is the outward unit normal to S, is

$$\oint_S \mathbf{j}\cdot d\mathbf{S} = \oint_S j_n\, dS. \tag{3.4.2}$$

Using Gauss's theorem (1.1.8) on the right-hand side of (3.4.2),

$$\oint_S j_n\, dS = \int_B \operatorname{div}\mathbf{j}\, dV = \int_B \operatorname{div}\big(-\chi \operatorname{grad} u\big)\, dV = -\chi \int_B \nabla^2 u\, dV. \tag{3.4.3}$$

On the other hand the heat content of B is, with c denoting specific heat and ρ density,

$$\int_B c\rho u\, dV. \tag{3.4.4}$$

The decrease of the heat content of B per unit time, that is the increase rendered with the opposite sign (since $\mathbf{j}$ is directed outwards), is then

$$-\frac{\partial}{\partial t}\int_B c\rho\, u\, dV = -\int_B c\rho\, \frac{\partial u}{\partial t}\, dV, \tag{3.4.5}$$

where the interchange of differentiation and integration is permissible. In the absence of a heat source within the medium, the rate of decrease of the heat content of B must equal the rate at which heat leaves B; combining (3.4.3) and (3.4.5),

$$-\int_B c\rho\, u_t\, dV = -\chi \int_B (\nabla^2 u)\, dV,$$

or

$$\int_B \left[c\rho\, \frac{\partial u}{\partial t} - \chi\, \nabla^2 u \right] dV = 0. \tag{3.4.6}$$

For (3.4.6) to hold for an arbitrary region B the integrand must vanish identically. Setting $\alpha = \chi/(c\rho)$, the equation governing the heat distribution is therefore

$$u_t = \alpha\, \nabla^2 u. \tag{3.4.7}$$

If the medium contains in addition a heat source (or sink), $k^*(x,y,z,t)$ denoting the heat produced (or absorbed) per unit volume per unit time and $k \equiv k^*/(c\rho)$, the equation governing the heat distribution becomes

$$u_t = \alpha \nabla^2 u + k. \tag{3.4.8}$$

As is the case for the wave equation, it should be noted that the function disturbing the homogeneity of the equation represents the presence of an external intervention to the system.

3.5 The Gravitational Potential

According to Newton's law of universal gravitation,

$$|\mathbf{K}| = \frac{Gmm'}{r^2},$$

two material point masses m and m' separated by a distance r attract each other with a force inversely proportional to the square of r. Throughout what follows we shall, with a slight abuse of notation, denote by

$$K \equiv -\frac{Gmm'}{r^2}$$

the (signed) radial component of $\mathbf{K}$ along the position vector $\mathbf{r}$ from m to m', the negative sign expressing the attractive character of the force. We aim to establish the partial differential equation satisfied by the potential function U from which $\mathbf{K}$ derives.

Place one of the material points at the origin of the coordinate system.

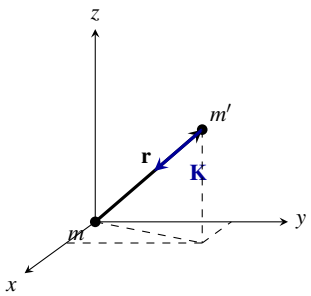


Fig. 3.4 Geometry of the gravitational interaction between m at the origin and m' at the tip of the position vector $\mathbf{r}$ of magnitude r. The attractive force $\mathbf{K}$ acts on m' along the line joining the two masses, anti-parallel to $\mathbf{r}$.

As may be readily seen from Fig. 3.4, with $\cos\alpha_1, \cos\alpha_2, \cos\alpha_3$ the direction cosines of $\mathbf{K}$,

$$K_x = K\cos\alpha_1 = K\frac{x}{r}, \quad K_y = K\cos\alpha_2 = K\frac{y}{r}, \quad K_z = K\cos\alpha_3 = K\frac{z}{r}, \tag{3.5.1}$$

and

$$K = -G\frac{mm'}{x^2+y^2+z^2}, \tag{3.5.2}$$

so that the components K_x, K_y, K_z may be written as

$$\left.\begin{aligned} K_x &= K\frac{x}{r} = -\frac{Gmm'}{x^2+y^2+z^2}\,\frac{x}{\sqrt{x^2+y^2+z^2}} = -\frac{Gmm'x}{(x^2+y^2+z^2)^{3/2}},\\ K_y &= -\frac{Gmm'y}{(x^2+y^2+z^2)^{3/2}},\\ K_z &= -\frac{Gmm'z}{(x^2+y^2+z^2)^{3/2}}. \end{aligned}\right\} \tag{3.5.3}$$

On the other hand, by definition of the potential function,

$$K_x = -\frac{\partial U}{\partial x}, \quad K_y = -\frac{\partial U}{\partial y}, \quad K_z = -\frac{\partial U}{\partial z}. \tag{3.5.4}$$

Combining (3.5.3) and (3.5.4) one easily obtains

$$U = -\frac{Gmm'}{(x^2+y^2+z^2)^{1/2}} = -\frac{Gmm'}{r}. \tag{3.5.5}$$

Direct differentiation now gives

$$\frac{\partial^2 U}{\partial x^2} = -\frac{\partial K_x}{\partial x} = Gmm'\,\frac{(x^2+y^2+z^2)^{3/2} - 3x^2(x^2+y^2+z^2)^{1/2}}{(x^2+y^2+z^2)^3} = Gmm'\,\frac{y^2+z^2-2x^2}{(x^2+y^2+z^2)^{5/2}}, \tag{3.5.6}$$

and similarly

$$\frac{\partial^2 U}{\partial y^2} = Gmm'\,\frac{x^2+z^2-2y^2}{(x^2+y^2+z^2)^{5/2}}, \qquad \frac{\partial^2 U}{\partial z^2} = Gmm'\,\frac{x^2+y^2-2z^2}{(x^2+y^2+z^2)^{5/2}}. \tag{3.5.7}$$

Adding (3.5.6) and (3.5.7) one sees that the potential function from which the universal gravitational force derives satisfies the second-order PDE

$$\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial z^2} = 0, \tag{3.5.8}$$

which, in terms of the Laplace operator ∇^2 introduced in (1.1.5), takes the compact form

$$\nabla^2 U = 0.$$

Equation (3.5.8) is called *Laplace's equation* and is encountered in many branches of physics. The functions which satisfy Laplace's equation form a special family called *harmonic functions*.

3.6 Sound Waves Propagating in a Gas

Hearing consists in the perception, through the ear, of pressure variations occurring in the atmosphere. Such pressure variations originating at a source—e.g. in human speech—constitute what we call sound. As is readily appreciated, sound requires a medium for its propagation. Although it can also propagate in liquids and solids, here we shall content ourselves with deriving the partial differential equation governing only the propagation of sound waves in a gas.

Let p, ρ, υ denote, respectively, the pressure, density, and velocity of the gas. Consider the one-dimensional fluid-dynamical equations of mass conservation (continuity) and momentum conservation (Euler equation):

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x}(\rho \upsilon) = 0, \tag{3.6.1}$$

$$\rho \frac{\partial \upsilon}{\partial t} + \rho \upsilon \frac{\partial \upsilon}{\partial x} = -\frac{\partial p}{\partial x}, \tag{3.6.2}$$

On the other hand, neglecting heat conduction and viscosity in the gas, with K and γ constants, the relation between p and ρ valid in an adiabatic state of the gas is given by

$$p = K\rho^{\gamma}, \tag{3.6.3}$$

which must also be appended to the foregoing equations.

Suppose now that a small perturbation is established in a gas which is initially at rest in equilibrium. Writing the corresponding values of p, ρ, υ, with the dimensionless quantities $\bar{p}, \bar{\rho}, \bar{\upsilon} \ll 1$, in the form

$$\upsilon = \upsilon_0 \bar{\upsilon}, \quad \rho = \rho_0(1+\bar{\rho}), \quad p = p_0(1+\bar{p}), \tag{3.6.4}$$

where υ_0 denotes a characteristic velocity and ρ_0, p_0 the equilibrium density and pressure of the gas, we now insert these expressions into (3.6.1)–(3.6.3), neglecting squares and products of the perturbation quantities. From the continuity equation (3.6.1),

$$\frac{\partial}{\partial t}\left[\rho_0(1+\bar{\rho})\right] + \frac{\partial}{\partial x}\left[\rho_0(1+\bar{\rho})\,\upsilon_0\bar{\upsilon}\right] = \rho_0 \frac{\partial \bar{\rho}}{\partial t} + \rho_0 \upsilon_0 \frac{\partial \bar{\upsilon}}{\partial x} = 0.$$

From the momentum equation (3.6.2),

$$\rho_0(1+\bar{\rho})\,\upsilon_0 \frac{\partial \bar{\upsilon}}{\partial t} + \rho_0(1+\bar{\rho})\,\upsilon_0^2 \bar{\upsilon} \frac{\partial \bar{\upsilon}}{\partial x} = -p_0 \frac{\partial \bar{p}}{\partial x},$$

which, on dropping squares and products of perturbations, reduces to

$$\rho_0 \upsilon_0 \frac{\partial \bar{\upsilon}}{\partial t} = -p_0 \frac{\partial \bar{p}}{\partial x}.$$

Finally, expanding the adiabatic relation (3.6.3) with the help of the binomial series,

$$p_0(1+\bar{p}) = K\big[\rho_0(1+\bar{\rho})\big]^\gamma = K\rho_0^\gamma(1+\bar{\rho})^\gamma \simeq K\rho_0^\gamma(1+\gamma\bar{\rho}),$$

and using $p_0 = K\rho_0^\gamma$ together with the smallness of $\bar{p}$ and $\bar{\rho}$, we obtain $\bar{p} = \gamma\bar{\rho}$. Collecting,

$$\upsilon_0 \frac{\partial \bar{\upsilon}}{\partial t} + \frac{p_0}{\rho_0}\frac{\partial \bar{p}}{\partial x} = 0, \qquad \frac{\partial \bar{\rho}}{\partial t} + \upsilon_0 \frac{\partial \bar{\upsilon}}{\partial x} = 0, \qquad \bar{p} = \gamma\bar{\rho}. \tag{3.6.5}$$

Differentiating the first of (3.6.5) with respect to x, the second with respect to t, comparing the two, and using the third, all the linearised quantities $u \in \{\bar{p}, \bar{\rho}, \bar{\upsilon}\}$ are seen to satisfy the one-dimensional wave equation

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}, \qquad \left(c^2 = \gamma\frac{p_0}{\rho_0}\right). \tag{3.6.6}$$

Chapter 4
Classification of Second-Order Partial Differential Equations

Abstract Second-order linear PDEs are classified as hyperbolic, parabolic, or elliptic by the sign of $B^2 - 4AC$. Each type can be reduced, by a non-singular change of variables, to a canonical form—two for the hyperbolic case and one each for the parabolic and elliptic cases. The classification is extended to n independent variables via the Law of Inertia of quadratic forms.

4.1 The Case of Two Independent Variables

Let $u = u(x,y)$ denote a function of the two independent variables x and y, and write its partial derivatives compactly as

$$u_x = \frac{\partial u}{\partial x}, \quad u_y = \frac{\partial u}{\partial y}, \quad u_{xx} = \frac{\partial^2 u}{\partial x^2}, \quad u_{yy} = \frac{\partial^2 u}{\partial y^2}, \quad u_{xy} = \frac{\partial^2 u}{\partial x\,\partial y}, \ \ldots$$

The most general second-order linear partial differential equation that $u(x,y)$ may satisfy is

$$A(x,y)u_{xx} + B(x,y)u_{xy} + C(x,y)u_{yy} + D(x,y)u_x + E(x,y)u_y + F(x,y)u = G(x,y), \tag{4.1.1}$$

the coefficient functions and the unknown u being assumed at least twice continuously differentiable, with the coefficients not vanishing simultaneously at any point (x,y).

Equation (4.1.1) bears a formal resemblance to the Cartesian equation of a quadratic curve (a conic) in analytic geometry,

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0, \tag{4.1.2}$$

whose classification in terms of the discriminant $\Delta = B^2 - 4AC$ is well known: $\Delta > 0$, $\Delta = 0$, $\Delta < 0$ correspond, respectively, to a hyperbola, a parabola, or an ellipse.

Definition 4.1 (Type of a PDE) The PDE (4.1.1) is said to be *hyperbolic*, *parabolic*, or *elliptic* at the point (x_0, y_0) according as

$$\Delta(x_0,y_0) = B^2(x_0,y_0) - 4A(x_0,y_0)C(x_0,y_0)$$

is positive, zero, or negative respectively. If, throughout a region B of the (x,y)-plane,

$$\Delta(x,y) = B^2(x,y) - 4A(x,y)C(x,y)$$

keeps the same sign (or vanishes identically), the PDE is said to be of that type in B.

The analogy with the conic (4.1.2) is purely formal: the actual solutions of (4.1.1) have, of course, no immediate geometric relationship with hyperbolas, parabolas, or ellipses.

Change of independent variables. Now consider a general change of independent variables

$$\xi = \xi(x,y), \qquad \eta = \eta(x,y), \tag{4.1.3}$$

such that $\forall (x,y) \in B$,

$$J = \begin{vmatrix} \xi_x & \xi_y \\ \eta_x & \eta_y \end{vmatrix} \neq 0. \tag{4.1.4}$$

The non-vanishing of the Jacobian guarantees the existence of the inverse of (4.1.3); our object is to determine the form taken by (4.1.1) under (4.1.3). By the chain rule,

$$\begin{aligned}
u_x &= u_\xi \xi_x + u_\eta \eta_x, \\
u_y &= u_\xi \xi_y + u_\eta \eta_y, \\
u_{xx} &= u_{\xi\xi}\xi_x^2 + 2u_{\xi\eta}\xi_x\eta_x + u_{\eta\eta}\eta_x^2 + u_\xi \xi_{xx} + u_\eta \eta_{xx}, \\
u_{yy} &= u_{\xi\xi}\xi_y^2 + 2u_{\xi\eta}\xi_y\eta_y + u_{\eta\eta}\eta_y^2 + u_\xi \xi_{yy} + u_\eta \eta_{yy}, \\
u_{xy} &= u_{\xi\xi}\xi_x\xi_y + u_{\xi\eta}\left(\xi_x\eta_y + \xi_y\eta_x\right) + u_{\eta\eta}\eta_x\eta_y + u_\xi \xi_{xy} + u_\eta \eta_{xy}.
\end{aligned}$$

Substituting these expressions into (4.1.1) we obtain

$$\left.\begin{aligned}
A^* &= A\xi_x^2 + B\xi_x\xi_y + C\xi_y^2, \\
B^* &= 2A\xi_x\eta_x + B(\xi_x\eta_y + \xi_y\eta_x) + 2C\xi_y\eta_y, \\
C^* &= A\eta_x^2 + B\eta_x\eta_y + C\eta_y^2, \\
D^* &= A\xi_{xx} + B\xi_{xy} + C\xi_{yy} + D\xi_x + E\xi_y, \\
E^* &= A\eta_{xx} + B\eta_{xy} + C\eta_{yy} + D\eta_x + E\eta_y, \\
F^* &= F, \\
G^* &= G,
\end{aligned}\right\} \tag{4.1.5}$$

so that the transformed equation reads

$$A^*u_{\xi\xi} + B^*u_{\xi\eta} + C^*u_{\eta\eta} + D^*u_\xi + E^*u_\eta + F^*u = G^*. \tag{4.1.6}$$

Hence the form of equation (4.1.1) is preserved under the change of variables (4.1.3). A straightforward computation from (4.1.5) shows that the discriminants of (4.1.1) and of its transform (4.1.6) are related by

$$B^{*2} - 4A^*C^* = J^2\big(B^2 - 4AC\big), \tag{4.1.7}$$

so that, since $J \neq 0$, not only the form but also the type of the equation is preserved.

The classification of (4.1.1) depends only on the coefficients $A(x,y)$, $B(x,y)$ and $C(x,y)$. Lumping the lower-order terms of (4.1.1) into a single function

$$H = H(x,y,u,u_x,u_y),$$

we may rewrite (4.1.1) in the abridged form

$$Au_{xx} + Bu_{xy} + Cu_{yy} = H, \tag{4.1.8}$$

and equation (4.1.6), with $H^* = H^*(\xi,\eta,u,u_\xi,u_\eta)$, takes the analogous form

$$A^*u_{\xi\xi} + B^*u_{\xi\eta} + C^*u_{\eta\eta} = H^*. \tag{4.1.9}$$

Canonical forms. Consider (4.1.8) now with $A \neq 0$, $B \neq 0$, $C \neq 0$, and seek new variables ξ, η of the form $\xi = \xi(x,y)$, $\eta = \eta(x,y)$ that, in view of (4.1.5), make the coefficients

$$A^* = A\xi_x^2 + B\xi_x\xi_y + C\xi_y^2 = 0,$$
$$C^* = A\eta_x^2 + B\eta_x\eta_y + C\eta_y^2 = 0$$

of (4.1.9) vanish simultaneously. Both equations have the common form

$$A\zeta_x^2 + B\zeta_x\zeta_y + C\zeta_y^2 = 0.$$

Dividing this expression by ζ_y^2 gives

$$A\left(\frac{\zeta_x}{\zeta_y}\right)^2 + B\left(\frac{\zeta_x}{\zeta_y}\right) + C = 0. \tag{4.1.10}$$

Along a curve $\zeta(x,y) = \text{const.}$ (that is, along $\xi = \text{const.}$ or $\eta = \text{const.}$),

$$d\zeta(x,y) = \zeta_x\,dx + \zeta_y\,dy = 0,$$

so that

$$\frac{dy}{dx} = -\frac{\zeta_x}{\zeta_y}. \tag{4.1.11}$$

Hence, by means of (4.1.11), equation (4.1.10) can be rewritten as the second-order algebraic equation in dy/dx,

$$A\left(\frac{dy}{dx}\right)^2 - B\frac{dy}{dx} + C = 0, \tag{4.1.12}$$

whose roots are

$$\frac{dy}{dx} = \frac{B \pm \sqrt{B^2 - 4AC}}{2A}. \tag{4.1.13}$$

The two ordinary differential equations (4.1.13), called the *characteristic equations* of (4.1.1), are first-order ODEs in the (x,y)-plane, the solutions of which are families of curves called the *characteristic curves* of (4.1.1). Their general solutions involve one integration constant.

Hyperbolic case. If the equation under consideration is of hyperbolic type, then $\Delta = B^2 - 4AC > 0$ and (4.1.13) furnishes two distinct real families of characteristic curves. With $\xi = \xi(x,y)$, $\eta = \eta(x,y)$ chosen as the equations of these two families of characteristics, the coefficients A^* and C^* in the transformed equation (4.1.9) both vanish, and the equation becomes

$$u_{\xi\eta} = \frac{H^*}{B^*} = H_1, \tag{4.1.14}$$

where $B^* \neq 0$ in view of (4.1.7). Equation (4.1.14) is called the *first canonical form of the hyperbolic equation*. Now define new independent variables

$$\alpha = \xi + \eta, \qquad \beta = \xi - \eta. \tag{4.1.15}$$

Then

$$u_\xi = \frac{\partial u}{\partial \alpha}\frac{\partial \alpha}{\partial \xi} + \frac{\partial u}{\partial \beta}\frac{\partial \beta}{\partial \xi} = u_\alpha + u_\beta,$$
$$u_{\xi\eta} = u_{\alpha\alpha} + u_{\alpha\beta} - u_{\beta\alpha} - u_{\beta\beta} = u_{\alpha\alpha} - u_{\beta\beta},$$

so that

$$u_{\alpha\alpha} - u_{\beta\beta} = H_2(\alpha, \beta, u, u_\alpha, u_\beta), \tag{4.1.16}$$

the *second canonical form of the hyperbolic equation*.

Parabolic case. If the equation is of parabolic type, then $\Delta = B^2 - 4AC = 0$ and the characteristic equation (4.1.13) reduces to the single equation

$$\frac{dy}{dx} = \frac{B}{2A}, \tag{4.1.17}$$

which determines a single family of characteristic curves. Choosing ξ to be the equation of this family, one of the two transformation equations is fixed in such a way that $A^* = 0$. The second transformation equation, $\eta(x,y)$, may then be selected arbitrarily, provided $J \neq 0$. With $B^2 = 4AC$, the quadratic form A^* factors as a perfect square,

$$A^* = A\xi_x^2 + B\xi_x\xi_y + C\xi_y^2 = \left(\sqrt{A}\,\xi_x \pm \sqrt{C}\,\xi_y\right)^2 = 0, \tag{4.1.18}$$

and one verifies similarly that

$$B^* = 2A\xi_x\eta_x + B(\xi_x\eta_y + \xi_y\eta_x) + 2C\xi_y\eta_y = 2\left(\sqrt{A}\,\xi_x \pm \sqrt{C}\,\xi_y\right)\left(\sqrt{A}\,\eta_x \pm \sqrt{C}\,\eta_y\right).$$

Since $\eta(x,y)$ is functionally independent of $\xi(x,y)$, (4.1.18) forces

$$B^* = 0.$$

On the other hand, $C^* \neq 0$ for this case, as we shall see in a moment. Consequently the canonical form for parabolic equations is

$$u_{\eta\eta} = H_3. \tag{4.1.19}$$

Of course, if η is taken instead as the solution of (4.1.17) and ξ chosen arbitrarily, one obtains

$$u_{\xi\xi} = H_4$$

as another canonical form.

Elliptic case. If the equation is of elliptic type, then $\Delta = B^2 - 4AC < 0$ and (4.1.13) has no real solutions; instead, the two roots are complex conjugates of one another, so that ξ and η are now complex. Putting

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

the conjugacy of ξ and η guarantees that α and β are real, and one immediately deduces that

$$\xi = \alpha + i\beta, \qquad \eta = \alpha - i\beta. \tag{4.1.20}$$

In this case equation (4.1.8) takes the form

$$\overline{A}(\alpha,\beta)\,u_{\alpha\alpha} + \overline{B}(\alpha,\beta)\,u_{\alpha\beta} + \overline{C}(\alpha,\beta)\,u_{\beta\beta} = \overline{H}(\alpha,\beta,u,u_\alpha,u_\beta), \tag{4.1.21}$$

where $\overline{A}$, $\overline{B}$ and $\overline{C}$ are the analogues of the coefficients in (4.1.5), but now expressed in terms of α and β. Using the transformation (4.1.20), the two equations $A^* = 0$ and $C^* = 0$,

$$\begin{aligned} A^* &= A\xi_x^2 + B\xi_x\xi_y + C\xi_y^2 = 0, \\ C^* &= A\eta_x^2 + B\eta_x\eta_y + C\eta_y^2 = 0, \end{aligned}$$

become, on substituting

$$\begin{aligned} \xi_x &= \xi_\alpha\alpha_x + \xi_\beta\beta_x = \alpha_x + i\beta_x, \\ \xi_y &= \xi_\alpha\alpha_y + \xi_\beta\beta_y = \alpha_y + i\beta_y, \\ \eta_x &= \eta_\alpha\alpha_x + \eta_\beta\beta_x = \alpha_x - i\beta_x, \\ \eta_y &= \eta_\alpha\alpha_y + \eta_\beta\beta_y = \alpha_y - i\beta_y, \end{aligned}$$

the complex pair

$$\begin{aligned}
&\left(A\alpha_x^2 + B\alpha_x\alpha_y + C\alpha_y^2\right) - \left(A\beta_x^2 + B\beta_x\beta_y + C\beta_y^2\right) \\
&\qquad + i\left[2A\alpha_x\beta_x + B(\alpha_x\beta_y + \alpha_y\beta_x) + 2C\alpha_y\beta_y\right] = 0, \\
&\left(A\alpha_x^2 + B\alpha_x\alpha_y + C\alpha_y^2\right) - \left(A\beta_x^2 + B\beta_x\beta_y + C\beta_y^2\right) \\
&\qquad - i\left[2A\alpha_x\beta_x + B(\alpha_x\beta_y + \alpha_y\beta_x) + 2C\alpha_y\beta_y\right] = 0,
\end{aligned}$$

that is to say,

$$(\overline{A} - \overline{C}) + i\overline{B} = 0, \qquad (\overline{A} - \overline{C}) - i\overline{B} = 0.$$

These two equations are simultaneously satisfied if and only if

$$\overline{A} = \overline{C} \quad \text{and} \quad \overline{B} = 0.$$

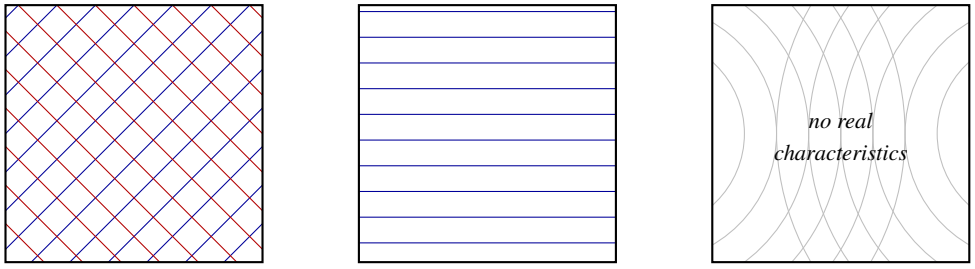
Substituting back into (4.1.21),

$$\overline{A}\,u_{\alpha\alpha} + \overline{A}\,u_{\beta\beta} = \overline{H}(\alpha, \beta, u, u_\alpha, u_\beta),$$

or, dividing through by $\overline{A}$,

$$u_{\alpha\alpha} + u_{\beta\beta} = \frac{\overline{H}}{\overline{A}} = H_5(\alpha, \beta, u, u_\alpha, u_\beta), \tag{4.1.22}$$

which is the *canonical form of the elliptic equation*.



Hyperbolic ($\Delta > 0$)
two real families
e.g. $u_{tt} - u_{xx} = 0$

Parabolic ($\Delta = 0$)
one real family
e.g. $u_t - u_{xx} = 0$

Elliptic ($\Delta < 0$)
no real family
e.g. $u_{xx} + u_{yy} = 0$

Fig. 4.1 The three types of second-order linear PDEs in two independent variables, distinguished by the geometry of their characteristic curves. In the hyperbolic case ($\Delta = B^2 - 4AC > 0$) two distinct real families of characteristics cross transversally, and information propagates along them. In the parabolic case ($\Delta = 0$) the two families coalesce into a single one, along which information propagates instantaneously. In the elliptic case ($\Delta < 0$) no real characteristics exist; no propagation phenomenon is associated, and the natural problem is the boundary-value problem of Chapter 5. The distinction governs which auxiliary conditions yield well-posed problems.

Remark 4.1 We now show that the coefficients of the second-order derivative terms in the canonical forms (4.1.5) cannot vanish simultaneously. Dividing the second of equations (4.1.5), namely

$$B^* = 2A\xi_x\eta_x + B(\xi_x\eta_y + \xi_y\eta_x) + 2C\xi_y\eta_y,$$

by $\xi_y\eta_y$ gives

$$\frac{B^*}{\xi_y\eta_y} = 2A\,\frac{\xi_x}{\xi_y}\,\frac{\eta_x}{\eta_y} + B\left(\frac{\xi_x}{\xi_y} + \frac{\eta_x}{\eta_y}\right) + 2C.$$

For the families $\xi = \text{const.}$ and $\eta = \text{const.}$ of characteristic curves we have, by (4.1.11) and (4.1.13),

$$\frac{\xi_x}{\xi_y} = -\frac{dy}{dx} = -\frac{B+\sqrt{B^2-4AC}}{2A}, \qquad \frac{\eta_x}{\eta_y} = -\frac{dy}{dx} = -\frac{B-\sqrt{B^2-4AC}}{2A},$$

and a brief computation gives

$$\frac{B^*}{\xi_y\eta_y} = -\frac{B^2-4AC}{A}.$$

Hence $\Delta(x,y) > 0$ corresponds to $B^* \neq 0$ (the hyperbolic case) and $\Delta(x,y) < 0$ to $B^* \neq 0$ together with $C^* \neq 0$ (the elliptic case): in either situation the coefficients of the second-order terms cannot vanish simultaneously, so that the type of (4.1.1) is invariant under the change of variables (4.1.3).

To summarise, the second-order linear PDE (4.1.1) is classified according to the coefficients of its second-order derivative terms as follows:

1. if $B^2 - 4AC > 0$, the equation is hyperbolic;
2. if $B^2 - 4AC = 0$, the equation is parabolic;
3. if $B^2 - 4AC < 0$, the equation is elliptic.

The one-dimensional vibrating-string (or wave) equation

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

is hyperbolic, and its corresponding canonical form is

$$\frac{\partial^2 u}{\partial \xi\,\partial \eta} = 0$$

(note that the original equation already exhibits the second canonical form of the hyperbolic equation). The one-dimensional diffusion equation

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

is parabolic and is already in canonical form, while the two-dimensional Laplace equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

is elliptic and is also in canonical form.

Constant-coefficient simplifications. We now show that the canonical forms of constant-coefficient equations can be simplified still further. Consider the canonical forms

$$u_{xy} = a_1 u_x + a_2 u_y + a_3 u + f_1, \qquad \text{(hyperbolic)} \tag{4.1.23}$$

$$u_{xx} - u_{yy} = a_1^* u_x + a_2^* u_y + a_3^* u + f_1^*, \qquad \text{(hyperbolic)} \tag{4.1.24}$$

$$u_{yy} = b_1 u_x + b_2 u_y + b_3 u + f_2, \qquad \text{(parabolic)} \tag{4.1.25}$$

$$u_{xx} + u_{yy} = c_1 u_x + c_2 u_y + c_3 u + f_3, \qquad \text{(elliptic)} \tag{4.1.26}$$

and introduce two as yet undetermined constants α and β via the function transformation

$$v = u\, e^{-(\alpha x + \beta y)}, \tag{4.1.27}$$

the aim being to remove the first-order derivative terms by a judicious choice of α and β. From (4.1.27) one easily obtains, with $u = v e^{\alpha x + \beta y}$,

$$\left.\begin{aligned} u_x &= (v_x + \alpha v)\, e^{\alpha x + \beta y}, \\ u_y &= (v_y + \beta v)\, e^{\alpha x + \beta y}, \\ u_{xx} &= (v_{xx} + 2\alpha v_x + \alpha^2 v)\, e^{\alpha x + \beta y}, \\ u_{xy} &= (v_{xy} + \alpha v_y + \beta v_x + \alpha\beta v)\, e^{\alpha x + \beta y}, \\ u_{yy} &= (v_{yy} + 2\beta v_y + \beta^2 v)\, e^{\alpha x + \beta y}. \end{aligned}\right\} \tag{4.1.28}$$

Carrying these expressions into (4.1.23) yields

$$v_{xy} + (\beta - a_1) v_x + (\alpha - a_2) v_y + (\alpha\beta - \alpha a_1 - \beta a_2 - a_3) v = f_1\, e^{-(\alpha x + \beta y)}. \tag{4.1.29}$$

For the coefficients of the first-order derivative terms in (4.1.29) to vanish, one must have

$$\alpha = a_2, \qquad \beta = a_1.$$

With $\alpha = a_2$, $\beta = a_1$ in (4.1.27), the coefficients of v_x and v_y in (4.1.29) both vanish. Setting $g_1 \equiv f_1\, e^{-(\alpha x + \beta y)}$ and $h_1 \equiv a_1 a_2 + a_3$, one obtains

$$v_{xy} = (a_1 a_2 + a_3) v + g_1,$$

that is to say,

$$v_{xy} = h_1 v + g_1.$$

The expressions (4.1.28) carried into (4.1.25), (4.1.24) and (4.1.26) similarly yield the simplified equations

$$
\begin{aligned}
v_{xy} &= h_1 v + g_1, \\
v_{xx} - v_{yy} &= h_1^* v + g_1^*, \\
v_{xx} + v_{yy} &= h_3 v + g_3, \\
v_{yy} &= k_2 v_x + h_2 v + g_2 \qquad (v_{xx} = \overline{k}_2 v_y + \overline{h}_2 v + \overline{g}_2), \\
v_{xx} + v_{yy} &= h_3 v + g_3,
\end{aligned}
$$

the solutions of which are evidently much easier to obtain than those of the original equations.

Remark 4.2 (Problem) Show that the equations

(i) $u_{xx} - u_{yy} + 3u_x - 2u_y + u = 0$,
(ii) $3u_{xx} + 7u_{xy} + 2u_{yy} + u_y + u = 0$

can be brought, with C a constant and $v = u\,e^{-(\alpha\xi+\beta\eta)}$, to the form

$$v_{\xi\eta} = Cv,$$

and obtain

$$\text{(i)} \quad v_{\xi\eta} = -\tfrac{1}{16} v, \qquad \text{(ii)} \quad v_{\xi\eta} = \tfrac{54}{625} v.$$

4.2 Classification in More Than Two Independent Variables

Let A_{ij}, B_i, C and F be real functions of the variables x_j $(j = 1, 2, \ldots, n)$, and consider the equation

$$\sum_{i,j=1}^{n} A_{ij} \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^{n} B_i \frac{\partial u}{\partial x_i} + Cu + F = 0. \tag{4.2.1}$$

Replace the n variables $x_1, x_2, \ldots, x_n$ by the new variables $\xi_1, \xi_2, \ldots, \xi_n$ via the change

$$\xi_k = \xi_k(x_1, x_2, \ldots, x_n), \qquad (k = 1, 2, \ldots, n), \tag{4.2.2}$$

the functions ξ_k being assumed at least twice differentiable and the Jacobian of the transformation (4.2.2) being non-zero. Then

$$\frac{\partial u}{\partial x_i} = \sum_{k=1}^{n} \frac{\partial u}{\partial \xi_k} \frac{\partial \xi_k}{\partial x_i}, \tag{4.2.3}$$

$$\frac{\partial^2 u}{\partial x_i \partial x_j} = \sum_{k,l=1}^{n} \frac{\partial^2 u}{\partial \xi_k \partial \xi_l} \frac{\partial \xi_k}{\partial x_i} \frac{\partial \xi_l}{\partial x_j} + \sum_{k=1}^{n} \frac{\partial u}{\partial \xi_k} \frac{\partial^2 \xi_k}{\partial x_i \partial x_j}. \tag{4.2.4}$$

Substitution of (4.2.3) and (4.2.4) into (4.2.1) yields, with

$$\overline{A}_{kl} = \sum_{i,j=1}^{n} A_{ij} \frac{\partial \xi_k}{\partial x_i} \frac{\partial \xi_l}{\partial x_j}, \tag{4.2.5}$$

$$\overline{B}_k = \sum_{i,j=1}^{n} A_{ij} \frac{\partial^2 \xi_k}{\partial x_i \partial x_j} + \sum_{i=1}^{n} B_i \frac{\partial \xi_k}{\partial x_i}, \tag{4.2.6}$$

the transformed equation

$$\sum_{k,l=1}^{n} \overline{A}_{kl} \frac{\partial^2 u}{\partial \xi_k \partial \xi_l} + \sum_{k=1}^{n} \overline{B}_k \frac{\partial u}{\partial \xi_k} + Cu + F = 0. \tag{4.2.7}$$

Fix a particular point $M(x_1^0, x_2^0, \ldots, x_n^0)$ and assume that, at M,

$$\frac{\partial \xi_k}{\partial x_i} = \alpha_{ki}. \tag{4.2.8}$$

Then (4.2.5) becomes

$$\overline{A}_{kl} = \sum_{i,j=1}^{n} A_{ij}\, \alpha_{ki} \alpha_{lj}, \tag{4.2.9}$$

which is exactly the transformation law of the coefficients of the quadratic form

$$\sum_{i,j=1}^{n} A_{ij} P_i P_j \tag{4.2.10}$$

under the linear substitution

$$P_i = \sum_{k=1}^{n} \alpha_{ki} q_k. \tag{4.2.11}$$

By substitution, the quadratic form (4.2.10) becomes

$$\sum_{i,j,k,l=1}^{n} A_{ij}\, \alpha_{ki} \alpha_{lj}\, q_k q_l = \sum_{k,l=1}^{n} \overline{A}_{kl}\, q_k q_l. \tag{4.2.12}$$

We see that the coefficients A_{ij} in equation (4.2.1), under a linear substitution (4.2.11) at the point $(x_1^0, x_2^0, \ldots, x_n^0)$, transform exactly as the coefficients of the quadratic form (4.2.10). Of course the A_{ij} in (4.2.10) are now to be understood as the constants obtained by evaluating the A_{ij} in (4.2.1) at the point $(x_1^0, x_2^0, \ldots, x_n^0)$.

Remark 4.3 (Algebraic Law of Inertia) For real A_{ij}, the quadratic form (4.2.10) can always be reduced, by a suitable non-singular transformation (4.2.11), to a non-degenerate sum of $\pm$ squares

$$\sum_{i=1}^{m} \pm q_i^2 \qquad (m \leq n). \tag{4.2.13}$$

The numbers of positive and negative signs in (4.2.13) are determined by the original quadratic form (4.2.10) alone, independently of the transformation (4.2.11). This statement is known as the *Law of Inertia of quadratic forms*.

Suppose now that we have found a non-singular substitution (4.2.11) reducing (4.2.10) to (4.2.13). Then the corresponding transformation (4.2.2) satisfying (4.2.8)

reduces (4.2.1) to the form

$$\overline{A}_{kl}(x_1^0,x_2^0,\ldots,x_n^0) = \pm 1 \qquad (k=l\le m),$$

$$\overline{A}_{kl}(x_1^0,x_2^0,\ldots,x_n^0) = 0 \qquad (k\ne l \text{ and } k=l>m),$$

that is, to

$$\sum_{k,l=1}^{n} \overline{A}_{kl}\,\frac{\partial^2 u}{\partial\xi_k\,\partial\xi_l} + \sum_{k=1}^{n} \overline{B}_k\,\frac{\partial u}{\partial\xi_k} + Cu + F = 0,$$

which is just (4.2.7), the canonical form at the point $(x_1^0,x_2^0,\ldots,x_n^0)$. The classification of (4.2.1) at M is now governed by the signs of the n coefficients $\overline{A}_{kk}(x_1^0,x_2^0,\ldots,x_n^0)$:

- if all n are non-zero and have the same sign, (4.2.1) is *elliptic* at $(x_1^0,x_2^0,\ldots,x_n^0)$;
- if $n-1$ of them have the same sign and the remaining one has the opposite sign, (4.2.1) is *hyperbolic* at $(x_1^0,x_2^0,\ldots,x_n^0)$;
- if at least one of them vanishes, (4.2.1) is *parabolic* at $(x_1^0,x_2^0,\ldots,x_n^0)$.

If (4.2.1) has the corresponding type at every point of a region B, it is said to be elliptic (respectively hyperbolic, parabolic) in B.

4.3 Worked Exercises in Reduction to Canonical Form

Exercise 4.1 Determine the characteristic coordinates of

$$4y^2u_{xx} - e^{2x}u_{yy} - 4y^2u_x + \frac{e^{2x}}{y}u_y = 4^3y^4e^{3x},$$

reduce to canonical form, and find the general solution.

Solution. For

$$4y^2u_{xx} - e^{2x}u_{yy} - 4y^2u_x + \frac{e^{2x}}{y}u_y = 4^3y^4e^{3x}$$

the discriminant is

$$\Delta = 4y^2e^{2x} > 0,$$

so the equation is hyperbolic at every point of the (x,y)-plane. The characteristic coordinates are obtained from

$$\frac{dy}{dx} = \mp\frac{e^x}{2y} \quad \text{whence} \quad \xi = y^2+e^x, \qquad \eta = y^2-e^x.$$

Then

$$\begin{aligned}
u_x &= e^xu_\xi - e^xu_\eta, \quad u_y = 2yu_\xi + 2yu_\eta,\\
u_{xx} &= e^{2x}u_{\xi\xi} - 2e^{2x}u_{\xi\eta} + e^{2x}u_{\eta\eta} + e^xu_\xi - e^xu_\eta,\\
u_{yy} &= 4y^2u_{\xi\xi} + 8y^2u_{\xi\eta} + 4y^2u_{\eta\eta} + 2u_\xi + 2u_\eta,
\end{aligned}$$

and substituting into the equation gives

$$u_{\xi\eta} = -4y^2e^x \quad \text{and} \quad u_{\xi\eta} = \eta^2 - \xi^2.$$

Integrating this equation yields

$$\begin{aligned} u_\xi &= \frac{1}{3}\eta^3 - \xi^2\eta + F(\xi) \quad \big(\textstyle\int F(\xi)\,d\xi = \phi(\xi)\big), \\ u(\xi,\eta) &= \frac{1}{3}\eta^3\xi - \frac{1}{3}\xi^3\eta + \phi(\xi) + \psi(\eta) \\ &= \frac{1}{3}\eta\xi\big(\eta^2 - \xi^2\big) + \phi(\xi) + \psi(\eta), \\ u(x,y) &= -\frac{4}{3}y^2e^x\big(y^4 - e^{2x}\big) + \phi(y^2 + e^x) + \psi(y^2 - e^x). \end{aligned}$$

Exercise 4.2 (Tricomi equation) Determine the type of the equation

$$u_{xx} + xu_{yy} = 0$$

in each region of the (x,y)-plane to which it belongs, and find the canonical forms.

Solution. The discriminant is $\Delta = -4x$.

(i) For $x < 0$, $\Delta > 0$, and the equation is hyperbolic. The characteristic coordinates are obtained from

$$\frac{dy}{dx} = \pm(-x)^{1/2},$$

that is,

$$\xi = y + \frac{2}{3}(-x)^{3/2}, \qquad \eta = y - \frac{2}{3}(-x)^{3/2}.$$

Then

$$\begin{aligned} &u_x = -(-x)^{1/2}u_\xi + (-x)^{1/2}u_\eta, \quad u_y = u_\xi + u_\eta, \\ &u_{xx} = (-x)u_{\xi\xi} - 2(-x)u_{\xi\eta} + (-x)u_{\eta\eta} + \frac{1}{2}(-x)^{-1/2}u_\xi - \frac{1}{2}(-x)^{-1/2}u_\eta, \\ &u_{yy} = u_{\xi\xi} + 2u_{\xi\eta} + u_{\eta\eta}, \end{aligned}$$

and carrying these expressions into the equation yields

$$u_{\xi\eta} = \frac{1}{8}(-x)^{-3/2}u_\xi - \frac{1}{8}(-x)^{-3/2}u_\eta$$

and

$$u_{\xi\eta} = \frac{1}{6(\xi - \eta)}\big(u_\xi - u_\eta\big).$$

(ii) For $x > 0$, $\Delta < 0$, and the equation is elliptic. Its characteristic coordinates are

$$\xi = \frac{3}{2}y - ix^{3/2}, \qquad \eta = \frac{3}{2}y + ix^{3/2},$$

and

$$\alpha = \frac{1}{2}(\xi + \eta) = \frac{3}{2}y, \qquad \beta = \frac{1}{2i}(\xi - \eta) = -x^{3/2},$$

giving the canonical form

$$u_{\alpha\alpha} + u_{\beta\beta} = -\frac{1}{3\beta}u_\beta.$$

Exercise 4.3 Determine the characteristic coordinates of

$$\sin^2 x\, u_{xx} + \sin 2x\, u_{xy} + \cos^2 x\, u_{yy} = x$$

and reduce to canonical form.

Solution. For

$$\sin^2 x\, u_{xx} + \sin 2x\, u_{xy} + \cos^2 x\, u_{yy} = x$$

the discriminant is

$$\Delta = B^2 - 4AC = \sin^2 2x - 4\sin^2 x \cos^2 x = 0,$$

so the equation is parabolic. The characteristic coordinates are

$$\xi = y - \ln(\sin x), \qquad \eta = y \ \text{(arbitrary, with } J \neq 0).$$

Then

$$\begin{aligned}
u_x &= -\cot x\, u_\xi, \qquad u_y = u_\xi + u_\eta, \\
u_{xx} &= \cot^2 x\, u_{\xi\xi} + \frac{1}{\sin^2 x}u_\xi, \\
u_{xy} &= -\cot x\, u_{\xi\xi} - \cot x\, u_{\xi\eta}, \\
u_{yy} &= u_{\xi\xi} + 2u_{\xi\eta} + u_{\eta\eta},
\end{aligned}$$

and on substituting these into the original equation,

$$\cos^2 x\, u_{\eta\eta} + u_\xi = x,$$

that is,

$$u_{\eta\eta} + \frac{u_\xi}{\cos^2\left(\arcsin e^{\eta-\xi}\right)} = \arcsin e^{\eta-\xi}.$$

Exercise 4.4 Determine the characteristic coordinates of

$$e^{2x}u_{xx} + e^x u_{xy} + u_{yy} + e^{2x}u_x - \frac{3}{4}u_y = 0,$$

reduce to canonical form, and find the general solution.

Solution. The discriminant is

$$\Delta = e^{2x} - 4e^{2x} = -3e^{2x} < 0,$$

which is negative for every x, so the equation is elliptic. Its characteristic coordinates are

$$\frac{dy}{dx} = \frac{1}{2}\big(1 \mp \sqrt{3}\,i\big)e^{-x},$$

$$\xi = y + \frac{1}{2}\big(1 + \sqrt{3}\,i\big)e^{-x}, \qquad \eta = y + \frac{1}{2}\big(1 - \sqrt{3}\,i\big)e^{-x},$$

and

$$\alpha = y + \frac{1}{2}e^{-x}, \qquad \beta = \frac{\sqrt{3}}{2}e^{-x}.$$

Then

$$\begin{aligned}
u_x &= -\frac{1}{2}e^{-x}u_\alpha - \frac{\sqrt{3}}{2}e^{-x}u_\beta, \quad u_y = u_\alpha, \\
u_{xx} &= \frac{1}{4}e^{-2x}u_{\alpha\alpha} + \frac{\sqrt{3}}{2}e^{-2x}u_{\alpha\beta} + \frac{3}{4}e^{-2x}u_{\beta\beta} + \frac{1}{2}e^{-x}u_\alpha + \frac{\sqrt{3}}{2}e^{-x}u_\beta, \\
u_{xy} &= -\frac{1}{2}e^{-x}u_{\alpha\alpha} - \frac{\sqrt{3}}{2}e^{-x}u_{\beta\alpha}, \qquad u_{yy} = u_{\alpha\alpha},
\end{aligned}$$

and carrying these expressions into the equation gives

$$u_{\alpha\alpha} + u_{\beta\beta} = u_\alpha.$$

Exercise 4.5 Determine the characteristic coordinates of

$$u_{xy} + 2u_{yy} + 9u_x + u_y = 2$$

and reduce to canonical form.

Solution. The discriminant is $\Delta = 1 > 0$, so the equation is hyperbolic; moreover $A \equiv 0$. Hence the characteristic coordinates may be obtained from

$$\begin{aligned}
\frac{dx}{dy} &= 0, \quad \xi = x, \\
\frac{dx}{dy} &= \frac{B}{C} = \frac{1}{2}, \quad \eta = x - \frac{1}{2}y.
\end{aligned}$$

Then

$$\begin{aligned}
u_x &= u_\xi + u_\eta, \quad u_{xy} = -\frac{1}{2}u_{\xi\eta} - \frac{1}{2}u_{\eta\eta}, \\
u_y &= -\frac{1}{2}u_\eta, \quad u_{yy} = \frac{1}{4}u_{\eta\eta},
\end{aligned}$$

and substituting into the original equation gives

$$u_{\xi\eta} = 18u_\xi + 17u_\eta - 4.$$

Exercise 4.6 Determine the characteristic coordinates of

$$u_{xx}+2u_{xy}+3u_{yy}+4u_x+5u_y+6u=e^x,$$

reduce to canonical form, and apply a substitution that eliminates the first-order derivative terms in the simplified canonical form.

Solution. The discriminant is $\Delta = 4-12=-8<0$, so the equation is elliptic everywhere. Its characteristic coordinates are obtained from

$$\frac{dy}{dx}=\frac{2\mp 2\sqrt{2}\,i}{2},$$

$$\xi = y-\big(1+\sqrt{2}\,i\big)x, \qquad \eta = y+\big(\sqrt{2}\,i-1\big)x,$$

and

$$\alpha = y-x, \qquad \beta = -\sqrt{2}\,x.$$

Then

$$\begin{aligned} u_x &= -u_\alpha-\sqrt{2}\,u_\beta, \quad u_y=u_\alpha, \quad u_{xx}=u_{\alpha\alpha}+2\sqrt{2}\,u_{\alpha\beta}+2u_{\beta\beta},\\ u_{xy} &= -u_{\alpha\alpha}-\sqrt{2}\,u_{\alpha\beta}, \quad u_{yy}=u_{\alpha\alpha}, \end{aligned}$$

and carrying these expressions into the original equation gives

$$u_{\alpha\alpha}+u_{\beta\beta}-\frac{1}{2}u_\alpha-2\sqrt{2}\,u_\beta+3u=\frac{1}{2}e^{-\beta/\sqrt{2}}.$$

To eliminate the first-order derivative terms, apply the function transformation

$$u=v\,e^{(A\alpha+B\beta)},$$

under which

$$\begin{aligned} u_\alpha &= (v_\alpha+Av)\,e^{(A\alpha+B\beta)}, \quad u_\beta=(v_\beta+Bv)\,e^{(A\alpha+B\beta)},\\ u_{\alpha\alpha} &= (v_{\alpha\alpha}+2Av_\alpha+A^2v)\,e^{(A\alpha+B\beta)},\\ u_{\beta\beta} &= (v_{\beta\beta}+2Bv_\beta+B^2v)\,e^{(A\alpha+B\beta)}. \end{aligned}$$

Substituting into the canonical equation yields

$$v_{\alpha\alpha}+v_{\beta\beta}+\big(2A+\tfrac{1}{2}\big)v_\alpha+\big(2B-2\sqrt{2}\big)v_\beta+\left(A^2+B^2+\frac{A}{2}-2\sqrt{2}B+3\right)v=\frac{1}{2}e^{-\beta/\sqrt{2}}\,e^{-(A\alpha+B\beta)}.$$

The coefficients of the first-order derivative terms vanish provided

$$A=-\frac{1}{4}, \quad B=\sqrt{2},$$

and the simplified equation becomes

$$v_{\alpha\alpha}+v_{\beta\beta}+\frac{15}{16}v=\frac{1}{2}e^{\frac{1}{4}\alpha-\frac{3}{\sqrt{2}}\beta}.$$

Exercise 4.7 Determine the characteristic coordinates of

$$u_{xx} + u_{xy} + 4u_x + 5u_y + u = e^x,$$

reduce to canonical form, and apply a substitution that eliminates the first-order derivative terms in the simplified canonical form.

Solution. The discriminant is $\Delta = 1 > 0$, so the equation is hyperbolic everywhere. Its characteristic coordinates are

$$\frac{dy}{dx} = \frac{1 \mp 1}{2}, \qquad \xi = y, \quad \eta = y - x,$$

and reduction to canonical form yields

$$u_{\xi\eta} = 5u_\xi + u_\eta + u - e^{\xi - \eta}.$$

To eliminate the first-order derivative terms, apply

$$u = v\, e^{\alpha\xi + \beta\eta},$$

which gives

$$v_{\xi\eta} + (\alpha - 1)v_\eta + (\beta - 5)v_\xi + (\alpha\beta - 5\alpha - \beta - 1)v = -e^{(1-\alpha)\xi - (1+\beta)\eta}.$$

Setting the coefficients of the first-order derivative terms equal to zero,

$$\alpha = 1, \quad \beta = 5,$$

the simplified canonical form is

$$v_{\xi\eta} - 6v = -e^{-6\eta}.$$

Exercise 4.8 Determine the characteristic coordinates of

$$u_{xx} + 6u_{xy} + 9u_{yy} + u_x + 3u_y = 0,$$

reduce to canonical form, and apply a substitution to eliminate the first-order derivative terms in the simplified canonical form.

Solution. The discriminant is $\Delta = 0$, so the equation is parabolic everywhere. Its characteristic coordinates are

$$\xi = y - 3x,$$

together with $\eta = y$ chosen arbitrarily so that $J \neq 0$. Then the canonical form reads

$$u_{\eta\eta} = -\frac{1}{3}u_\eta.$$

To eliminate the first-order derivative term, apply

$$u = v e^{\beta\eta},$$

giving

$$3v_{\eta\eta} + (6\beta + 1)v_\eta + (3\beta^2 + \beta)v = 0.$$

Setting the coefficient of the first-order derivative term equal to zero,

$$6\beta + 1 = 0, \qquad \beta = -\frac{1}{6},$$

and substituting this value back into the equation yields

$$v_{\eta\eta} = \frac{1}{36}v.$$

Exercise 4.9 Determine the characteristic coordinates of

$$(1-x)u_{xx} - u_{yy} - u_x = 0 \qquad (0 \le x \le 1),$$

reduce to canonical form, and apply the function transformation

$$u = \frac{w}{\sqrt{\xi + \eta}}$$

to eliminate the first-order derivative terms in the simplified canonical form.

Solution. The discriminant is $\Delta = 4(1-x) > 0$, so the equation is hyperbolic. Its characteristic coordinates are

$$\xi = 2(1-x)^{1/2} + y, \qquad \eta = 2(1-x)^{1/2} - y,$$

and the canonical form reads

$$4u_{\xi\eta} + \frac{1}{2(1-x)^{1/2}}\left(u_\xi + u_\eta\right) = 0, \quad u_{\xi\eta} + \frac{1}{2(\xi+\eta)}\left(u_\xi + u_\eta\right) = 0.$$

This is a variable-coefficient equation. To eliminate the first-order derivative terms, apply the function transformation

$$u = \frac{\omega}{\sqrt{\xi + \eta}} = (\xi + \eta)^{-1/2}\omega,$$

which gives

$$\begin{aligned}
u_\xi &= (\xi+\eta)^{-1/2}\omega_\xi - \frac{1}{2}(\xi+\eta)^{-3/2}\omega, \\
u_\eta &= (\xi+\eta)^{-1/2}\omega_\eta - \frac{1}{2}(\xi+\eta)^{-3/2}\omega, \\
u_{\xi\eta} &= (\xi+\eta)^{-1/2}\omega_{\xi\eta} - \frac{1}{2}(\xi+\eta)^{-3/2}\omega_\xi - \frac{1}{2}(\xi+\eta)^{-3/2}\omega_\eta + \frac{3}{4}(\xi+\eta)^{-5/2}\omega,
\end{aligned}$$

and substituting into the canonical equation yields

$$\omega_{\xi\eta} + \frac{1}{4(\xi+\eta)^2}\,\omega = 0.$$

Problems

4.1 After reducing each of the following constant-coefficient equations to canonical form, apply a substitution to eliminate the first-order derivative terms:

(a) $u_{xx} + u_{yy} + 2u_x + 8u_y + u = 0$,
(b) $6u_{xx} - u_{xy} + u = y^2$,
(c) $u_{xy} + u_x + u_y = 3x$,
(d) $u_{yy} - 9u_x + 7u_y = \cos y$.

4.2 Determine the type of each of the following equations and reduce them to canonical form:

(a) $xu_{xx} + u_{yy} = x^2$,
(b) $u_{xx} + y^2 u_{yy} = y$,
(c) $x^2 u_{xx} + 2xy\,u_{xy} + y^2 u_{yy} = e^x$,
(d) $e^x u_{xx} + e^y u_{yy} = u$,
(e) $u_{xx} - yu_{xy} + xu_x + yu_y + u = 0$.

4.3 For each of the following equations, after determining the regions in which it is hyperbolic, parabolic, or elliptic, reduce it to canonical form in each such region:

(a) $xu_{xx} + yu_{yy} = 0$,
(b) $u_{xx} + xy\,u_{yy} = 0$,
(c) $y^2 u_{xx} - x^2 u_{yy} = 0$,
(d) $x^2 u_{xx} - y^2 u_{yy} = 0$,
(e) $x^2 u_{xx} + y^2 u_{yy} = 0$,
(f) $y^2 u_{xx} + x^2 u_{yy} = 0$,
(g) $y^2 u_{xx} + 2xy\,u_{xy} + x^2 u_{yy} = 0$,
(h) $4y^2 u_{xx} - e^{2x} u_{yy} - 4y^2 u_x = 0$,
(i) $x^2 u_{xx} + 2xy\,u_{xy} - 3y^2 u_{yy} - 2xu_x + 4yu_y + 16x^4 u = 0$,
(j) $u_{xx}\sin^2 x - 2y\,u_{xy}\sin x + y^2 u_{yy} = 0$.

Chapter 5
Boundary Conditions and Well-Posedness

Abstract This chapter introduces boundary conditions of Dirichlet, Neumann, and Robin type, and Hadamard's three criteria of well-posedness: existence, uniqueness, and continuous dependence of the solution on the data. The well-posed problems associated with the hyperbolic, parabolic and elliptic types are described and illustrated by canonical examples that anticipate the developments of Chapters 6–10.

5.1 Auxiliary Conditions

A partial differential equation by itself does not determine its solution. The operator equation $L[u] = f$ admits, in general, infinitely many solutions, parametrised by arbitrary functions; to single out one of them one must supplement the equation with *auxiliary conditions*, by which the unknown u is constrained on selected parts of the closure of the domain on which $L[u] = f$ is posed.

The auxiliary conditions fall into two categories, distinguished by the geometric character of the variable along which they are imposed. *Initial conditions* prescribe the value of u, and possibly of its first time derivative, on the initial section $t = t_0$ of the domain; they describe the state of the system at the moment from which its evolution is to be tracked. *Boundary conditions*, in contrast, prescribe the behaviour of u on the spatial boundary of the domain; they encode the system's interaction with its surroundings.

Three kinds of boundary condition recur throughout mathematical physics, and we shall encounter all three in the chapters that follow. Let $\partial\Omega$ denote the boundary of the spatial domain Ω and $\partial/\partial n$ the outward normal derivative on $\partial\Omega$.

1. A *Dirichlet condition* prescribes the value of u on $\partial\Omega$:

$$u\big|_{\partial\Omega} = \varphi.$$

   Physically, it fixes the temperature of a body at its surface, the displacement of a string at its endpoints, or an electrostatic potential on a conductor.

2. A *Neumann condition* prescribes the value of the normal derivative of u:

$$\left.\frac{\partial u}{\partial n}\right|_{\partial\Omega} = \psi.$$

   It fixes the heat flux through the boundary, the slope of a membrane at its rim, or the normal component of an electric field at a surface.
3. A *Robin condition* (also called a *third-kind* or *mixed* condition) prescribes a linear combination of u and its normal derivative:

$$\left.\left(\alpha\, u + \beta\, \frac{\partial u}{\partial n}\right)\right|_{\partial\Omega} = \chi,$$

   with α, β given functions on $\partial\Omega$, not simultaneously zero. It models, for instance, Newton's law of convective heat exchange between a body and an external medium, under which the heat flux through the surface is proportional to the temperature difference.

A boundary condition is called *homogeneous* if its right-hand side vanishes (φ, ψ or χ identically zero); otherwise it is said to be *inhomogeneous*. Linear superposition—the guiding thread of this book, introduced in Section 2.2—reduces every linear inhomogeneous boundary-value problem to a finite collection of homogeneous ones, and it is therefore the homogeneous boundary conditions that one must learn to handle.

Three classes of problem. A problem in which only initial conditions are imposed (the spatial domain being unbounded, or boundary effects being neglected) is called an *initial-value problem* or *Cauchy problem*; Chapter 6 is devoted to it. A problem in which only boundary conditions are imposed (no time variable being present, or time being separated out at the start) is called a *boundary-value problem*; the steady-state heat-conduction problem and the electrostatic potential problem of Chapter 3 are of this type. The hybrid case, in which both initial and boundary conditions are present, is an *initial–boundary-value problem* or *mixed problem*; the bulk of the worked examples of Chapters 7–10 are of this kind.

5.2 Hadamard's Criteria of Well-Posedness

The mere imposition of auxiliary conditions does not guarantee that the resulting problem is mathematically meaningful. A problem may admit no solution, or several solutions, or solutions whose dependence on the data is so unstable that an arbitrarily small perturbation of the data produces an arbitrarily large change in the solution—a circumstance that, in any physical application, deprives the problem of predictive value.

J. Hadamard, at the beginning of the twentieth century, formulated three criteria that together capture what is required of a problem if it is to be mathematically and physically acceptable.

Definition 5.1 (Well-posedness) A problem consisting of a partial differential equation and a set of auxiliary conditions is said to be *well posed*, in the sense of Hadamard, if the following three conditions are satisfied:

1. **Existence.** A solution exists in some prescribed class of functions.
2. **Uniqueness.** The solution is unique in that class.
3. **Continuous dependence.** The solution depends continuously on the data: small changes in the auxiliary conditions, the coefficients of the equation, or the inhomogeneous term produce correspondingly small changes in the solution.

A problem failing any one of these criteria is called *ill posed*.

The first two criteria are the natural mathematical demands; the third is a stability requirement of physical origin. The latter is indispensable in any application: the data of the problem are known only to within some experimental tolerance, and a solution that fails to depend continuously on the data is, for practical purposes, no solution at all.

An archetypal ill-posed problem. A celebrated example of ill-posedness, due to Hadamard, is the Cauchy problem for the two-dimensional Laplace equation on the half-plane $y \geq 0$,

$$u_{xx} + u_{yy} = 0, \qquad u(x,0) = 0, \qquad u_y(x,0) = \frac{1}{n}\sin nx, \tag{5.2.1}$$

with n a positive integer. The solution is

$$u_n(x,y) = \frac{1}{n^2}\sin nx \sinh ny.$$

As $n \to \infty$ the data $u_y(x,0) = n^{-1}\sin nx$ tend uniformly to zero, while at any fixed $y > 0$ the solution $u_n(x,y)$ grows exponentially in n (Fig. 5.1). An arbitrarily small change in the data thus produces an arbitrarily large change in the solution: the third criterion fails, and the Cauchy problem for the Laplace equation is ill posed. The phenomenon is not accidental; it is a structural consequence of the elliptic character of the equation, as the systematic treatment below makes clear.

5.3 Well-Posed Problems Type by Type

The classification theorem of Chapter 4 divides second-order linear PDEs into the hyperbolic, parabolic, and elliptic types. To each type corresponds a characteristic class of well-posed problems, summarised in Table 5.1. We describe each in turn, deferring proofs to the chapters in which the relevant techniques are developed.

Hyperbolic equations. For the wave equation

$$u_{tt} = c^2 \nabla^2 u,$$

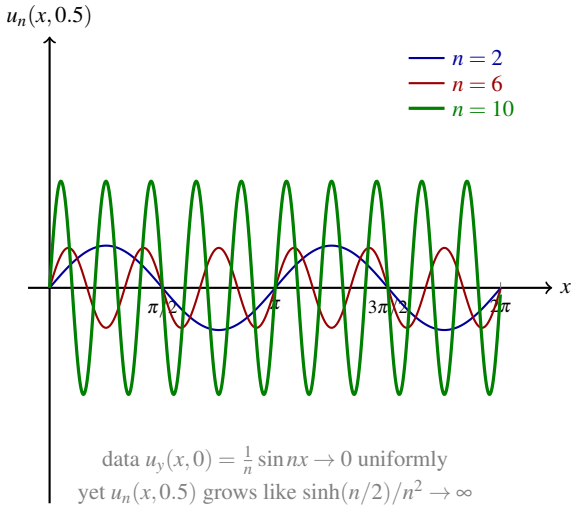


Fig. 5.1 Hadamard's example (5.2.1): the solution profiles $u_n(x,0.5)$ at the fixed altitude $y=0.5$ for $n=2$, 6, 10. The Cauchy data on $y=0$ tend uniformly to zero in n, yet at any fixed positive y the solution amplitude grows exponentially. The third Hadamard criterion (continuous dependence on the data) fails: an arbitrarily small perturbation of the boundary data produces an arbitrarily large change in the solution.

Table 5.1 Well-posed problems associated with the three types of second-order linear PDE. The canonical example in each row is treated in detail in the chapter indicated.

Type	Well-posed problem	Canonical example
Hyperbolic	Cauchy or initial–boundary-value	wave equation
Parabolic	Initial–boundary-value	heat equation
Elliptic	Boundary-value (Dirichlet, Neumann, Robin)	Laplace equation

the prototypical hyperbolic equation, the natural problem is the *Cauchy problem*: in the unbounded case one prescribes, on the initial surface $t=0$, the values of u and u_t,

$$u(\mathbf{r},0)=f(\mathbf{r}),\qquad u_t(\mathbf{r},0)=g(\mathbf{r}),$$

and the solution is determined throughout space-time. For a bounded spatial domain the same initial conditions are supplemented by a boundary condition (Dirichlet, Neumann or Robin) on the lateral boundary, yielding an initial–boundary-value problem. The d'Alembert formula (6.2.7) of Chapter 6 provides the explicit solution in one space dimension; the separation-of-variables method of Chapter 7 treats the bounded case.

Parabolic equations. For the heat equation

$$u_t=\alpha\nabla^2 u,$$

the natural problem is the *initial–boundary-value problem*: on the initial surface $t=0$ one prescribes

$$u(\mathbf{r},0) = f(\mathbf{r}),$$

and on the lateral boundary $\partial\Omega \times (0,T]$ one imposes a single boundary condition (Dirichlet, Neumann or Robin). Note that the parabolic case requires *one* initial condition, not two: the heat equation is first-order in time, and prescribing $u_t(\mathbf{r},0)$ in addition would in general lead to no solution. The backwards Cauchy problem, in which data are imposed on a line $t = t_0 < 0$ and the solution is sought for $t \to 0$, is itself ill posed—heat diffusion smooths the data forwards in time and, by the same token, amplifies them backwards. This irreversibility, characteristic of the parabolic type, is treated in detail in Chapter 7.

Elliptic equations. For the Laplace and Poisson equations

$$\nabla^2 u = 0, \qquad \nabla^2 u = h(\mathbf{r}),$$

no time variable is involved, and the natural problem is the *boundary-value problem*: on the boundary $\partial\Omega$ of a bounded spatial domain Ω one prescribes one of

- $u\big|_{\partial\Omega} = \varphi$ (Dirichlet problem),
- $\dfrac{\partial u}{\partial n}\Big|_{\partial\Omega} = \psi$ (Neumann problem),
- $\big(\alpha u + \beta\, \partial u/\partial n\big)\big|_{\partial\Omega} = \chi$ (Robin problem).

Hadamard's example (5.2.1) shows that imposing initial conditions in place of boundary conditions destroys well-posedness: the elliptic type tolerates only boundary data. The detailed treatment is distributed across Chapters 7, 9 and 10.

The underlying principle. The preceding catalogue is not a list of arbitrary conventions; each entry is dictated by the geometry of the characteristics of the corresponding type, established in Chapter 4. Hyperbolic equations possess two real families of characteristics along which the data propagate; parabolic equations possess a single such family; elliptic equations possess none. The number and nature of the auxiliary conditions that yield a well-posed problem are in every case determined by this characteristic structure, and the proofs of well-posedness given in the subsequent chapters are made possible by the analytical methods—separation of variables, characteristics, Riemann's method, finite integral transforms—that respect it.

With these notions of boundary condition and well-posedness in hand we now proceed, in Chapter 6, to the most fundamental of the problems just enumerated: the Cauchy problem.

Problems

5.1 For each of the following boundary conditions on a bounded spatial domain $\Omega \subset \mathbb{R}^3$ with smooth boundary $\partial\Omega$, identify the type (Dirichlet, Neumann, or Robin) and state whether it is homogeneous or inhomogeneous:

(a) $u(\mathbf{r}) = T_0$ for $\mathbf{r} \in \partial\Omega$, with T_0 a non-zero constant;

(b) $\partial u/\partial n = 0$ on $\partial\Omega$;
(c) $u + 2\,\partial u/\partial n = \sin\theta$ on $\partial\Omega$, where θ is a coordinate on $\partial\Omega$;
(d) $\partial u/\partial n = q(\mathbf{r},t)$ on $\partial\Omega$, with q a prescribed function.

For each, indicate a physical situation (heat conduction, electrostatics, membrane vibration, etc.) in which it arises naturally.

5.2 Consider the one-dimensional heat equation

$$u_t = \alpha\, u_{xx}, \qquad 0 < x < \ell,\ t > 0,$$

together with the auxiliary conditions

$$u(x,0) = f(x), \qquad u(0,t) = u(\ell,t) = 0.$$

Identify the type (initial-value, boundary-value, or initial–boundary-value), classify each auxiliary condition as Dirichlet, Neumann, or Robin, and state which of Hadamard's three criteria (existence, uniqueness, continuous dependence) you would expect this problem to satisfy. (No proof is required; an informal physical argument suffices.)

5.3 Repeat the analysis of Problem 5.2 for the Cauchy problem of the wave equation on the whole real line,

$$u_{tt} = c^2 u_{xx}, \qquad x \in \mathbb{R},\ t > 0,$$
$$u(x,0) = f(x), \qquad u_t(x,0) = g(x),$$

and for the Dirichlet problem of the Laplace equation on the unit disc,

$$\nabla^2 u = 0, \qquad x^2 + y^2 < 1, \qquad u\big|_{x^2+y^2=1} = \varphi(\theta).$$

5.4 Show, by considering the family

$$u_n(x,y) = \frac{1}{n^2}\,\sin nx\,\sinh ny \qquad (n = 1,2,\dots),$$

that the Cauchy problem for the Laplace equation on the strip $\{0 \le y \le 1\}$ with data $u(x,0) = 0$, $u_y(x,0) = n^{-1}\sin nx$ does not depend continuously on the data. (This is the example cited in Section 5.2; the present problem asks the reader to verify the calculation and to identify which of Hadamard's criteria fails and which are still satisfied.)

5.5 A non-uniform string of length ℓ, fixed at both ends, has linear mass density $\rho(x)$ and is subjected to a tension T that may also depend on x. Its small transverse vibrations $u(x,t)$ are governed by the equation

$$\rho(x)\,u_{tt} = \big[T(x)\,u_x\big]_x, \qquad 0 < x < \ell,\ t > 0,$$

together with the conditions $u(0,t) = u(\ell,t) = 0$ and the initial data $u(x,0) = f(x)$, $u_t(x,0) = g(x)$.

(a) Identify the type of the equation (hyperbolic, parabolic, or elliptic), the type of problem (initial-value, boundary-value, or initial–boundary-value), and the classification of each auxiliary condition.

(b) State, without proof, why the problem is well posed.

5.6 Show that the Neumann problem for Poisson's equation,

$$\nabla^2 u = h(\mathbf{r}) \quad \text{in } \Omega, \qquad \left.\frac{\partial u}{\partial n}\right|_{\partial\Omega} = \psi,$$

fails Hadamard's first criterion (existence) unless the data satisfy the *compatibility condition*

$$\int_\Omega h(\mathbf{r})\, dV = \oint_{\partial\Omega} \psi\, dS,$$

and that, even when this condition holds, the second criterion (uniqueness) fails: any solution is determined only up to an additive constant. *Hint.* Apply Gauss's theorem (1.1.8) to $\mathbf{A} = \operatorname{grad} u$.

Chapter 6
The Initial-Value Problem for Partial Differential Equations

Abstract We develop the theory of the Cauchy problem for second-order linear PDEs and quote the Cauchy–Kowalevski theorem. The d'Alembert formula is derived for the homogeneous wave equation and extended to the inhomogeneous case by Green's theorem. The chapter culminates in Riemann's method for hyperbolic equations, illustrated by the telegrapher's equation, where the Riemann function is a Bessel function J_0.

6.1 The Cauchy Problem

It is well known that, given an ordinary differential equation, solving its initial-value problem amounts to finding the solution that satisfies a prescribed set of suitable initial conditions at a given starting point. For instance, to the second-order ODE

$$\frac{d^2u}{dt^2} = f\left(t, u, \frac{du}{dt}\right) \tag{6.1.1}$$

one adjoins, at a starting point $t = t_0$, the values of u and du/dt,

$$u(t_0) = A, \qquad \left(\frac{du}{dt}\right)_{t=t_0} = B. \tag{6.1.2}$$

The PDE (6.1.1) together with (6.1.2) forms an initial-value problem; solving it means determining the solution of (6.1.1) that satisfies the conditions (6.1.2).

A similar problem can be formulated for partial differential equations. We shall introduce it here for a second-order PDE in two independent variables; generalisation to PDEs containing more independent variables is entirely straightforward and presents no difficulty.

Consider, as an example, a second-order PDE arranged so as to be solved for u_{yy},

$$u_{yy} = F(x, y, u, u_x, u_y, u_{xx}, u_{xy}). \tag{6.1.3}$$

For a fixed value $y = y_0$, prescribe u and u_y as

$$u(x,y_0) = f(x), \qquad \left(\frac{\partial u}{\partial y}\right)_{y=y_0} = g(x), \tag{6.1.4}$$

where f and g are given functions of x. The PDE (6.1.3) together with (6.1.4) forms an *initial-value problem*.

In an initial-value problem the prescribed initial values, those that correspond to a definite value $y = y_0$ of one of the independent variables, need not necessarily be given along the line $y = y_0$; they may also be given along a curve C_0 in the (x,y)-plane defined parametrically by

$$x_0 = x_0(\lambda), \qquad y_0 = y_0(\lambda). \tag{6.1.5}$$

In this latter case one speaks, instead of an initial-value problem, of the *Cauchy problem* of the PDE under consideration.

Now consider the PDE

$$A(x,y)u_{xx} + B(x,y)u_{xy} + C(x,y)u_{yy} = F(x,y,u,u_x,u_y), \tag{6.1.6}$$

together with values $f = f(\lambda), g = g(\lambda)$ defined on the curve C_0 given by $x_0 = x_0(\lambda), y_0 = y_0(\lambda)$. Solving the Cauchy problem of (6.1.6) consists in finding, in a neighbourhood of C_0, the solution that satisfies the Cauchy conditions

$$u\big|_{C_0} = f(\lambda), \tag{6.1.7}$$

$$\left(\frac{\partial u}{\partial n}\right)_{C_0} = g(\lambda), \tag{6.1.8}$$

where $\hat{n}$ denotes, for an open curve, the unit normal vector pointing in the direction of increasing arc length and, for a closed curve, the outward unit normal (i.e. on the side opposite to a counter-clockwise traversal of the curve).

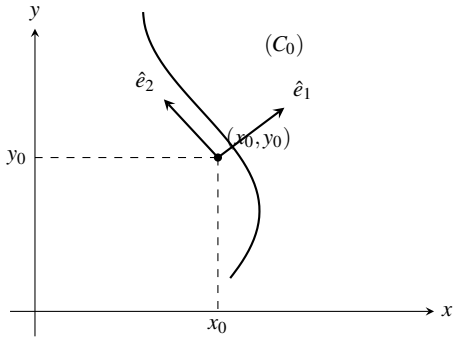


Fig. 6.1 The Cauchy data curve C_0 through (x_0,y_0) with local basis vectors $\hat{e}_1$ (tangent) and $\hat{e}_2$ (normal).

The functions $f = f(\lambda)$ and $g = g(\lambda)$ are called the *Cauchy data* on the curve C_0.

For each point (x_0, y_0) on C_0, the value of u is determined by (6.1.7). Accordingly, C_0 in the parametric form (6.1.5), together with the condition (6.1.7), may be viewed as the orthogonal projection onto the (x,y)-plane of a space curve (C) in the Cartesian space with coordinates (x,y,u). The solution of the Cauchy problem is then an integral surface in (x,y,u)-space passing through this space curve (C). This surface must of course also satisfy the condition (6.1.8), which represents the tangent plane along the curve (C).

If $f = f(\lambda)$ is a differentiable function, then along C_0 we can write

$$\frac{du}{d\lambda} = \frac{\partial u}{\partial x}\frac{dx}{d\lambda} + \frac{\partial u}{\partial y}\frac{dy}{d\lambda} = \frac{df}{d\lambda}, \tag{6.1.9}$$

and

$$\frac{\partial u}{\partial n} = \frac{\partial u}{\partial x}\frac{dx}{dn} + \frac{\partial u}{\partial y}\frac{dy}{dn} = g. \tag{6.1.10}$$

From differential geometry the unit tangent vector $\hat{t}$ is

$$\hat{t} = \frac{d\mathbf{r}}{ds} = \frac{dx}{ds}\hat{e}_x + \frac{dy}{ds}\hat{e}_y = \cos\varphi\,\hat{e}_x + \sin\varphi\,\hat{e}_y, \tag{6.1.11}$$

and the unit normal vector

$$\hat{n} = \frac{d\mathbf{r}}{dn} = \frac{dx}{dn}\hat{e}_x + \frac{dy}{dn}\hat{e}_y = -\sin\varphi\,\hat{e}_x + \cos\varphi\,\hat{e}_y. \tag{6.1.12}$$

Comparing (6.1.11) with (6.1.12), $\hat{n}$ may be written as

$$\hat{n} = -\frac{dy}{ds}\hat{e}_x + \frac{dx}{ds}\hat{e}_y.$$

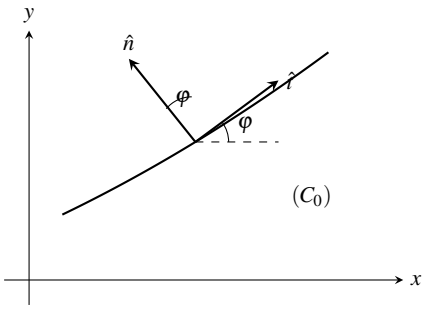


Fig. 6.2 Tangent $\hat{t}$ and outward normal $\hat{n}$ on C_0, with φ measured from the x-axis.

Hence (6.1.10) can be recast as

$$\frac{\partial u}{\partial n} = \hat{n}\cdot(\nabla u) = -\frac{\partial u}{\partial x}\frac{dy}{ds} + \frac{\partial u}{\partial y}\frac{dx}{ds} = g. \tag{6.1.13}$$

Equations (6.1.9) and (6.1.13) can be regarded as a 2×2 algebraic system for the two unknowns $\partial u/\partial x$ and $\partial u/\partial y$ on C_0. The system is uniquely solvable provided its principal determinant does not vanish. This is indeed the case, since

$$\begin{vmatrix} \dfrac{dx}{d\lambda} & \dfrac{dy}{d\lambda} \\ -\dfrac{dy}{ds} & \dfrac{dx}{ds} \end{vmatrix} = \frac{dx^2+dy^2}{d\lambda\, ds} = \frac{ds}{d\lambda} \neq 0,$$

whenever C_0 is a regular curve. Hence u_x and u_y are determined on C_0 by the Cauchy data.

Starting from these expressions of u_x and u_y, we now compute the second-order derivatives. To do so, we differentiate u_x and u_y with respect to the parameter λ along C_0:

$$\frac{d}{d\lambda}\left(\frac{\partial u}{\partial x}\right) = \frac{\partial^2 u}{\partial x^2}\frac{dx}{d\lambda} + \frac{\partial^2 u}{\partial x\partial y}\frac{dy}{d\lambda}, \tag{6.1.14}$$

$$\frac{d}{d\lambda}\left(\frac{\partial u}{\partial y}\right) = \frac{\partial^2 u}{\partial y\partial x}\frac{dx}{d\lambda} + \frac{\partial^2 u}{\partial y^2}\frac{dy}{d\lambda}. \tag{6.1.15}$$

On the other hand,

$$A(x,y)u_{xx} + B(x,y)u_{xy} + C(x,y)u_{yy} = F(x,y,u,u_x,u_y). \tag{6.1.16}$$

If (6.1.16) is now to be evaluated on C_0, the three equations (6.1.14), (6.1.15) and (6.1.16) form a 3×3 algebraic system for the three unknowns u_{xx}, u_{xy} and u_{yy}. The condition that this system be solvable is the non-vanishing of the determinant of its coefficient matrix:

$$\begin{vmatrix} \dfrac{dx}{d\lambda} & \dfrac{dy}{d\lambda} & 0 \\ 0 & \dfrac{dx}{d\lambda} & \dfrac{dy}{d\lambda} \\ A(x,y) & B(x,y) & C(x,y) \end{vmatrix} = C\left(\frac{dx}{d\lambda}\right)^2 - B\frac{dx}{d\lambda}\frac{dy}{d\lambda} + A\left(\frac{dy}{d\lambda}\right)^2 \neq 0.$$

Rearranging, this becomes

$$\left(\frac{dx}{d\lambda}\right)^2\left[A\left(\frac{dy}{dx}\right)^2 - B\frac{dy}{dx} + C\right] \neq 0,$$

that is,

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

But the equation

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

is precisely, as we have seen earlier, the algebraic equation defining the characteristic curves of the PDE (6.1.6). Consequently, the necessary condition for the second-order derivatives of u to be determinable on C_0 is that the curve C_0 along which the Cauchy data are prescribed not coincide with a characteristic curve of (6.1.6).

If $A(x,y), B(x,y), C(x,y), F(x,y,u,u_x,u_y)$ are analytic in their arguments, and if $x_0 = x_0(\lambda), y_0 = y_0(\lambda), f(\lambda), g(\lambda)$ are analytic functions, then the procedure outlined above can be continued to compute higher-order derivatives on C_0, and from these one can construct, in a neighbourhood of C_0, a (convergent series) representation of the solution of the form

$$u(x,y) = \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{1}{k!(n-k)!} \left(\frac{\partial^n u}{\partial x^k \partial y^{n-k}} \right)_{x=x_0, y=y_0} (x-x_0)^k (y-y_0)^{n-k}.$$

We omit the proof here, but the underlying result is the *Cauchy–Kowalevski theorem*, which we state in its general form.

Theorem 6.1 (Cauchy–Kowalevski) *For a second-order PDE*

$$u_{yy} = F(y; x_1, x_2, \ldots, x_n; u, u_y, u_{x_i}, u_{x_i y}, u_{x_i x_j})$$

in $n+1$ independent variables, with prescribed initial values

$$u(y_0, x_1, \ldots, x_n) = f(x_1, \ldots, x_n), \qquad \left(\frac{\partial u}{\partial y} \right)_{y=y_0} = g(x_1, \ldots, x_n),$$

provided F is analytic in a neighbourhood of $(y^0, x_1^0, \ldots, x_n^0, u^0, u_y^0, \ldots)$ and f, g are analytic in a neighbourhood of $(x_1^0, \ldots, x_n^0)$, the Cauchy problem possesses a unique analytic solution in a neighbourhood of $(y^0, x_1^0, \ldots, x_n^0)$.

6.2 Initial-Value Problem for the Homogeneous Wave Equation

We seek the solution of the initial-value problem

$$u_{tt} - c^2 u_{xx} = 0, \qquad u(x,0) = f(x), \qquad u_t(x,0) = g(x),$$

by the *method of characteristics*. We first reduce the equation to canonical form. The characteristic equations are

$$\frac{dx}{dt} = \mp c,$$

and the characteristic coordinates are

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

The canonical form is then

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

whose general solution, with $\Phi(\xi)$ and $\psi(\eta)$ arbitrary functions of ξ and η respectively, is

$$u(\xi,\eta) = \Phi(\xi) + \psi(\eta),$$

that is,

$$u(x,t) = \Phi(x+ct) + \psi(x-ct). \tag{6.2.1}$$

The initial conditions yield

$$u(x,0) = f(x) = \Phi(x) + \psi(x), \tag{6.2.2}$$

$$u_t(x,0) = g(x) = \left(\frac{\partial\Phi}{\partial\xi}\frac{\partial\xi}{\partial t} + \frac{\partial\psi}{\partial\eta}\frac{\partial\eta}{\partial t}\right)_{t=0} = \left(c\frac{d\Phi}{dx} - c\frac{d\psi}{dx}\right). \tag{6.2.3}$$

Integrating (6.2.3) (with x_0 and K constants),

$$\Phi(x) - \psi(x) = \frac{1}{c}\int_{x_0}^{x} g(\tau)\,d\tau + K. \tag{6.2.4}$$

From (6.2.4) and (6.2.2) we obtain

$$\Phi(x) = \tfrac{1}{2}f(x) + \frac{1}{2c}\int_{x_0}^{x} g(\tau)\,d\tau + \frac{K}{2}, \tag{6.2.5}$$

$$\psi(x) = \tfrac{1}{2}f(x) - \frac{1}{2c}\int_{x_0}^{x} g(\tau)\,d\tau - \frac{K}{2}. \tag{6.2.6}$$

Substituting $x \to x+ct$ in (6.2.5) and $x \to x-ct$ in (6.2.6), the solution of our initial-value problem becomes

$$u(x,t) = \tfrac{1}{2}\big[f(x+ct)+f(x-ct)\big] + \frac{1}{2c}\int_{x_0}^{x+ct} g(\tau)\,d\tau - \frac{1}{2c}\int_{x_0}^{x-ct} g(\tau)\,d\tau,$$

or, combining the two integrals over $[x-ct, x+ct]$,

$$\boxed{u(x,t) = \tfrac{1}{2}\big[f(x+ct)+f(x-ct)\big] + \frac{1}{2c}\int_{x-ct}^{x+ct} g(\tau)\,d\tau.} \tag{6.2.7}$$

This is the celebrated *d'Alembert formula*.

Exercises

Exercise 6.1 Solve the initial-value problem

$$u_{tt} = c^2 u_{xx}, \qquad u(x,0) = \sin x, \qquad u_t(x,0) = \cos x.$$

Solution.

$$u(x,t)=\tfrac{1}{2}\big[\sin(x+ct)+\sin(x-ct)\big]+\frac{1}{2c}\int_{x-ct}^{x+ct}\cos\tau\,d\tau=\sin x\cos ct+\frac{1}{c}\cos x\sin ct.$$

Exercise 6.2 Solve the initial-value problem

$$u_{tt}-c^2u_{xx}=0,\qquad u(x,0)=x^3,\qquad u_t(x,0)=x.$$

Solution.

$$u(x,t)=\tfrac{1}{2}\big[(x+ct)^3+(x-ct)^3\big]+\frac{1}{2c}\int_{x-ct}^{x+ct}\tau\,d\tau=x(x^2+3c^2t^2)+xt.$$

6.3 Initial-Value Problem for the Inhomogeneous Wave Equation: Method of Characteristics

Consider the initial-value problem

$$u_{tt}-c^2u_{xx}=H(x,t),\qquad u(x,0)=f(x),\qquad u_t(x,0)=G(x).$$

With $y=ct$ this hyperbolic equation can be rewritten in its second canonical form,

$$u_{xx}-u_{yy}=h(x,y),\qquad u(x,0)=f(x),\qquad u_y(x,0)=g(x),\tag{6.3.1}$$

where

$$h(x,y)=-\frac{H}{c^2},\qquad g(x)=\frac{G}{c}.$$

Choose now an arbitrary point $P_0(x_0,y_0)$ in the (x,y)-plane, and on the initial line $y=0$ a point $Q_0(x_0,0)$. The two families of characteristic curves of (6.3.1) are $x+y=c_1$ and $x-y=c_2$, that is, lines of slopes ∓1. Of these characteristics, the two passing through P_0 meet the x-axis at $P_1(x_0-y_0,0)$ and $P_2(x_0+y_0,0)$, respectively. The triangle $P_0P_1P_2$ thus formed has sides B_0,B_1,B_2, encloses the interior region B, and has boundary L (see Fig. 6.3).

If we now integrate (6.3.1) over the region B,

$$\iint_B(u_{xx}-u_{yy})\,dx\,dy=\iint_B h(x,y)\,dx\,dy,\tag{6.3.2}$$

and apply Green's theorem

$$\iint_B\left(\frac{\partial N}{\partial x}-\frac{\partial M}{\partial y}\right)dx\,dy=\oint_L(M\,dx+N\,dy),$$

the left-hand side becomes

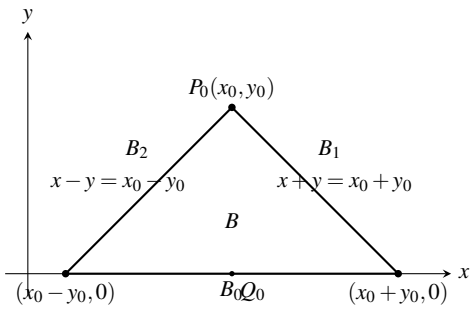


Fig. 6.3 Domain-of-dependence triangle $P_0P_1P_2$ for the one-dimensional wave equation; B_1, B_2 are characteristic sides and B_0 lies on the initial line $y = 0$.

$$\iint_B (u_{xx} - u_{yy})\,dx\,dy = \oint_{(L)} (u_x\,dy + u_y\,dx). \tag{6.3.3}$$

Since the closed contour L is composed of B_0, B_1, B_2, traversed in the order $B_0 \to B_1 \to B_2$, we evaluate each piece in turn.

On B_0: $y = 0$, hence $dy = 0$ and the integrand reduces to $u_y\,dx$,

$$\int_{B_0} \big(u_x\,dy + u_y\,dx\big) = \int_{x_0-y_0}^{x_0+y_0} u_y\,dx = \int_{x_0-y_0}^{x_0+y_0} u_y\,dx. \tag{6.3.4}$$

On B_1: the characteristic $x + y = x_0 + y_0$ gives $dx = -dy$, and

$$u_x\,dy + u_y\,dx = -u_x\,dx - u_y\,dy = -du,$$

so that

$$\int_{B_1} \big(u_x\,dy + u_y\,dx\big) = -\int_{B_1} du = -\int_{(x_0+y_0,0)}^{(x_0,y_0)} du = u(x_0+y_0,0) - u(x_0,y_0). \tag{6.3.5}$$

On B_2: the characteristic $x - y = x_0 - y_0$ gives $dx = dy$, and

$$u_x\,dy + u_y\,dx = u_x\,dx + u_y\,dy = du,$$

so that

$$\int_{B_2} \big(u_x\,dy + u_y\,dx\big) = \int_{B_2} du = \int_{(x_0,y_0)}^{(x_0-y_0,0)} du = u(x_0-y_0,0) - u(x_0,y_0). \tag{6.3.6}$$

Substituting (6.3.4), (6.3.5) and (6.3.6) into (6.3.3) gives

$$\oint_L \big(u_x\,dy + u_y\,dx\big) = -2u(x_0,y_0) + u(x_0-y_0,0) + u(x_0+y_0,0) + \int_{x_0-y_0}^{x_0+y_0} u_y\,dx. \tag{6.3.7}$$

Combining (6.3.2), (6.3.3) and (6.3.7),

$$-2u(x_0,y_0)+u(x_0-y_0,0)+u(x_0+y_0,0)+\int_{x_0-y_0}^{x_0+y_0}u_y\,dx=\iint_B h(x,y)\,dS,$$

hence

$$u(x_0,y_0)=\tfrac{1}{2}\big[u(x_0-y_0,0)+u(x_0+y_0,0)\big]+\tfrac{1}{2}\int_{x_0-y_0}^{x_0+y_0}u_y\,dx-\tfrac{1}{2}\iint_B h(x,y)\,dS. \quad (6.3.8)$$

Since (x_0,y_0) was an arbitrarily chosen point of the (x,y)-plane, we may replace (x_0,y_0) by (x,y) to obtain

$$u(x,y)=\tfrac{1}{2}\big[f(x+y)+f(x-y)\big]+\tfrac{1}{2}\int_{x-y}^{x+y}g(\tau)\,d\tau-\tfrac{1}{2}\iint_B h(x,y)\,dx\,dy. \quad (6.3.9)$$

Returning to the original variables,

$$\boxed{u(x,t)=\tfrac{1}{2}\big[f(x+ct)+f(x-ct)\big]+\frac{1}{2c}\int_{x-ct}^{x+ct}G(\tau)\,d\tau+\frac{1}{2c}\iint_B H(x,t)\,dS.} \quad (6.3.10)$$

Exercises

Exercise 6.3 Solve

$$u_{xx}-u_{yy}=1, \qquad u(x,0)=\sin x, \qquad u_y(x,0)=x.$$

Solution. The integral over B in the last term of (6.3.8) can be written, using the geometry of Fig. 6.4, as

$$\begin{aligned}\iint_B dx\,dy&=\int_0^{y_0}\int_{x_0-y_0+y}^{x_0}dx\,dy+\int_0^{y_0}\int_{x_0}^{x_0+y_0-y}dx\,dy\\&=\int_0^{y_0}\int_{x_0-y_0+y}^{x_0+y_0-y}dx\,dy=y_0^2.\end{aligned}$$

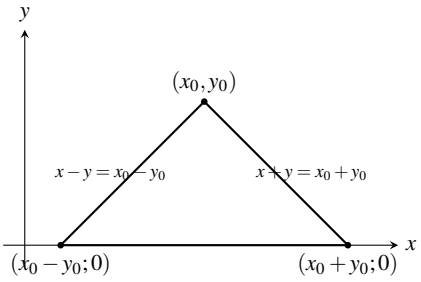


Fig. 6.4 Domain of dependence in Exercise 6.3.

Therefore, with $g(\tau)=\tau$,

$$\begin{aligned} u(x_0,y_0) &= \tfrac{1}{2}\big[\sin(x_0+y_0)+\sin(x_0-y_0)\big]+\tfrac{1}{2}\int_{x_0-y_0}^{x_0+y_0}\tau\,d\tau-\tfrac{1}{2}y_0^2 \\ &= \tfrac{1}{2}\big[\sin(x_0+y_0)+\sin(x_0-y_0)\big]+2x_0y_0-\tfrac{1}{2}y_0^2, \end{aligned}$$

and hence, replacing (x_0,y_0) by (x,y),

$$u(x,y)=\tfrac{1}{2}\big[\sin(x+y)+\sin(x-y)\big]+2xy-\tfrac{1}{2}y^2.$$

Exercise 6.4 Solve the following initial-value problems:

(a) $u_{tt}-c^2u_{xx}=x+ct, \quad u(x,0)=x, \quad u_t(x,0)=\sin x;$
(b) $u_{tt}-c^2u_{xx}=e^x, \quad u(x,0)=5, \quad u_t(x,0)=x^2;$
(c) $u_{tt}-c^2u_{xx}=\sin x, \quad u(x,0)=\cos x, \quad u_t(x,0)=1+x;$
(d) $u_{tt}-c^2u_{xx}=xe^t, \quad u(x,0)=\sin x, \quad u_t(x,0)=0;$
(e) $u_{tt}-c^2u_{xx}=2, \quad u(x,0)=x^2, \quad u_t(x,0)=\cos x.$

6.4 Solution of the Cauchy Problem for Hyperbolic Equations: Riemann's Method

In this section we shall present *Riemann's method*, a procedure that solves the Cauchy problem associated with the most general second-order linear hyperbolic PDE in canonical form,

$$L[u]\equiv u_{xy}+a(x,y)u_x+b(x,y)u_y+c(x,y)u=f(x,y), \tag{6.4.1}$$

where $a(x,y),b(x,y),c(x,y)$ and $f(x,y)$ are differentiable functions defined in some region B^* and

$$L\equiv\frac{\partial^2}{\partial x\,\partial y}+a\frac{\partial}{\partial x}+b\frac{\partial}{\partial y}+c$$

is a linear differential operator.

Let $v=v(x,y)$ be a function having continuous second-order partial derivatives. Then

$$\left.\begin{aligned} vu_{xy}-uv_{xy} &= (vu_x)_y-(uv_y)_x, \\ avu_x &= (avu)_x-u(av)_x, \\ bvu_y &= (bvu)_y-u(bv)_y, \end{aligned}\right\} \tag{6.4.2}$$

so that, upon defining

$$\left.\begin{aligned} U &= avu-uv_y, \\ V &= bvu+vu_x, \end{aligned}\right\} \tag{6.4.3}$$

and the operator

$$M[v]=v_{xy}-(av)_x-(bv)_y+cv, \tag{6.4.4}$$

one verifies the identity

$$vL[u] - uM[v] = U_x + V_y. \tag{6.4.5}$$

The operator M is called the *adjoint operator* of L. If $M = L$, then L is said to be *self-adjoint*.

We now integrate (6.4.5) over a region B contained in B^*. Using Green's theorem along the closed boundary C of B,

$$\iint_{B_*} \big(vL[u] - uM[v]\big)\, dx\, dy = \iint_{B_*} (U_x + V_y)\, dx\, dy = \oint_{(C)} (U\, dy - V\, dx). \tag{6.4.6}$$

Let Λ be a continuous initial-data curve as shown in Fig. 6.5. Since (6.4.1) is in canonical form, x and y are independent variables and serve as the characteristic coordinates of (6.4.1). Assume that Λ has no tangent parallel to the x- or y-axis. We seek the solution at an arbitrarily chosen point $P(\alpha, \beta)$ in the (x, y)-plane.

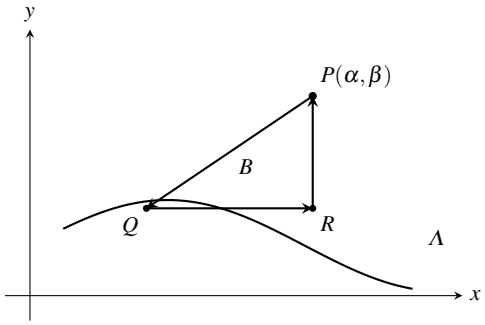


Fig. 6.5 Riemann's geometry: an initial-data curve Λ, the point $P(\alpha, \beta)$, and the characteristic lines $\overline{PQ}$ and $\overline{PR}$ enclosing the region B.

Through P draw $\overline{PQ}$, parallel to the x-axis, meeting Λ at Q, and $\overline{PR}$, parallel to the y-axis, meeting Λ at R. (Hence $x = \alpha =$const and $y = \beta =$const are characteristic lines.) The curve Λ being prescribed as an initial curve, the values of u and u_x (or equivalently u_y) are assumed to be given on Λ.

Letting C denote the closed contour $PQRP$ surrounding the region B, and noting that $dy = 0$ on $\overline{PQ}$ and $dx = 0$ on $\overline{PR}$, (6.4.6) gives

$$\iint_{B_*} \big(vL[u] - uM[v]\big)\, dx\, dy = \int_Q^R (U\, dy - V\, dx) + \int_R^P U\, dy - \int_P^Q V\, dx. \tag{6.4.7}$$

Using (6.4.3),

$$\int_P^Q V\, dx = \int_P^Q (b\, uv)\, dx + \int_P^Q v\, u_x\, dx,$$

and integrating the last term by parts,

$$\int_P^Q v\, u_x\, dx = [uv]_P^Q - \int_P^Q u\, v_x\, dx,$$

hence

$$\int_P^Q V\,dx = [uv]_Q - [uv]_P + \int_P^Q u\,(bv - v_x)\,dx.$$

A symmetric computation along $\overline{PR}$ yields

$$\int_R^P U\,dy = [uv]_P - [uv]_R + \int_R^P u\,(av - v_y)\,dy.$$

Substituting these into (6.4.7) and rearranging,

$$\begin{aligned}[uv]_P =& [uv]_Q + \int_P^Q u\,(bv - v_x)\,dx - \int_R^P u\,(av - v_y)\,dy \\ &- \int_Q^R (U\,dy - V\,dx) + \iint_{B_*} \big(vL[u] - uM[v]\big)\,dx\,dy.\end{aligned} \tag{6.4.8}$$

We now choose $v(x,y;\alpha,\beta)$ to be the solution of $M[v]=0$ satisfying

$$\left.\begin{aligned} &v_x(x,\beta;\alpha,\beta) = b(x,\beta)\,v(x,\beta;\alpha,\beta), && y=\beta,\\ &v_y(\alpha,y;\alpha,\beta) = a(\alpha,y)\,v(\alpha,y;\alpha,\beta), && x=\alpha,\\ &v(\alpha,\beta;\alpha,\beta) = 1. \end{aligned}\right\} \tag{6.4.9}$$

The function $v(x,y;\alpha,\beta)$ so determined is the *Riemann function*.

Since $L[u] = f(x,y)$, equation (6.4.8) simplifies under the conditions (6.4.9) to

$$\begin{aligned}[u]_P =& [uv]_Q - \int_Q^R uv\,(a\,dy - b\,dx) + \int_Q^R \big(uv_y\,dy + vu_x\,dx\big) \\ &+ \iint_{B_*} v\,f(x,y)\,dx\,dy,\end{aligned} \tag{6.4.10}$$

which expresses the value of u at $P(\alpha,\beta)$ as an integral formula involving the values of u and u_x on the arc $\overline{QR}$ of Λ. When the values of u and u_y are given instead on Λ, one uses the identity

$$\int_Q^R (uv)_x\,dx + (uv)_y\,dy = \int_Q^R d(uv) = [uv]_R - [uv]_Q,$$

which, combined with (6.4.10), yields

$$[u]_P = [uv]_R - \int_Q^R \big(uv_x\,dx + vu_y\,dy\big) - \int_Q^R uv\,(a\,dy - b\,dx) + \iint_{B_*} v\,f(x,y)\,dx\,dy. \tag{6.4.11}$$

Adding (6.4.10) and (6.4.11) side by side and using $[u]_P = u(\alpha,\beta)$, one obtains the final symmetric formula:

$$[u]_P = u(\alpha,\beta) = \tfrac{1}{2}\big([uv]_Q + [uv]_R\big) - \int_Q^R uv\,(a\,dy - b\,dx) - \tfrac{1}{2}\int_Q^R u\,(v_x\,dx + v_y\,dy) + \tfrac{1}{2}\int_Q^R v\,(u_x\,dx - u_y\,dy) + \iint_{B_*} v\,f(x,y)\,dx\,dy. \tag{6.4.12}$$

Equation (6.4.12) expresses the solution of the Cauchy problem at $P(\alpha,\beta)$ in terms of the Cauchy data on Λ. As is clear, the solution at (α,β) depends only on the Cauchy data on the arc $\overline{QR}$ of Λ.

Remark 6.1 A point worth emphasising here is the following. For the problem to admit a solution, the characteristics (lines) through P should meet Λ in a single point. Indeed, if as in Fig. 6.6 the curve Λ were intersected by the characteristic line $y = \beta$ at two points Q and S, the solution at $P(\alpha,\beta)$ corresponding to the Cauchy data on the arc $\overline{QR}$ would in general differ from that corresponding to the data on $\overline{RS}$. The same Cauchy problem would then admit two different solutions at the same point, contradicting well-posedness.

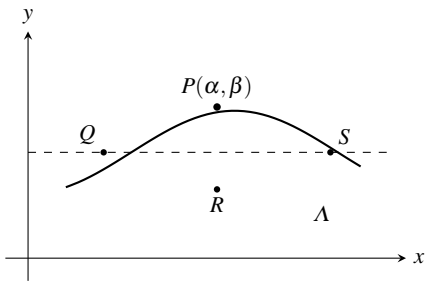


Fig. 6.6 When the characteristic $y = \beta$ through P meets Λ at two points Q, S, the Cauchy problem is overdetermined.

Exercises

Exercise 6.5 (Telegraphers' equation) Consider the telegraphers' equation

$$w_{tt} + a^* w_t + b^* w = c^2 w_{xx}. \tag{6.4.13}$$

By the substitution

$$w = u\,e^{\beta t},$$

with $\beta = -a^*/2$, equation (6.4.13) reduces to

$$u_{tt} - \left(\frac{a^{*2} - 4b^*}{4}\right)u = c^2 u_{xx}. \tag{6.4.14}$$

We shall use Riemann's method to determine the solution of (6.4.14) subject to the initial conditions

$$u(x,0) = f(x), \qquad u_t(x,0) = g(x). \tag{6.4.15}$$

Solution. The substitution

$$\xi = x + ct, \qquad \eta = x - ct, \tag{6.4.16}$$

brings (6.4.14) to canonical form. Setting

$$k \equiv \frac{a^{*2} - 4b^*}{16c^2},$$

the canonical form is

$$L[u] \equiv u_{\xi\eta} + ku = 0. \tag{6.4.17}$$

Under the same canonical change of variables the conditions (6.4.15) become, using $u_t = c(u_\xi - u_\eta)$ and noting that $t = 0$ corresponds to the diagonal $\xi = \eta$ in the (ξ, η)-plane (Fig. 6.7),

$$[u]_{\xi=\eta} = f(\xi), \tag{6.4.18}$$

$$\big[u_\xi - u_\eta\big]_{\xi=\eta} = \frac{1}{c} g(\xi). \tag{6.4.19}$$

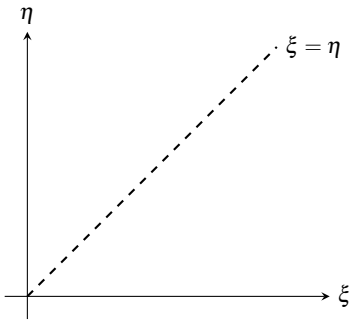


Fig. 6.7 In the characteristic (ξ, η)-plane, the initial line $t = 0$ corresponds to the diagonal $\xi = \eta$.

For (6.4.17) the coefficients of the first-order derivatives are $a = b = 0$, hence the Riemann function is determined by the problem

$$M[v] = v_{\xi\eta} + kv = 0, \tag{6.4.20}$$

$$v_\xi(\xi, \beta; \alpha, \beta) = 0, \tag{6.4.21}$$

$$v_\eta(\alpha, \eta; \alpha, \beta) = 0, \tag{6.4.22}$$

$$v(\alpha, \beta; \alpha, \beta) = 1. \tag{6.4.23}$$

From (6.4.17) and (6.4.20) we have

$$L[v] = M[v] = v_{\xi\eta} + kv,$$

so the operator $L = \partial^2/(\partial\xi\,\partial\eta) + k$ is self-adjoint.

To solve the Riemann problem above, set

$$s = (\xi - \alpha)(\eta - \beta),$$

and assume that the Riemann function has the form

$$v(\xi, \eta; \alpha, \beta) = F(s).$$

Then $v_\xi = (\eta - \beta)F'(s)$ and $v_{\xi\eta} = F'(s) + sF''(s)$, so substituting into (6.4.20) gives

$$sF''(s) + F'(s) + kF(s) = 0.$$

The change of variable $\lambda = 2\sqrt{ks}$ converts this to

$$\lambda^2 F''(\lambda) + \lambda F'(\lambda) + \lambda^2 F(\lambda) = 0,$$

which is the zeroth-order Bessel equation. Its general solution, with $J_0(\lambda)$ and $Y_0(\lambda)$ the two linearly independent series solutions of the equation, is

$$F(\lambda) = AJ_0(\lambda) + BY_0(\lambda).$$

For $\lambda = 0$ the function $Y_0(\lambda)$ diverges to $-\infty$, so boundedness forces $B = 0$. Since $J_0(0) = 1$ and $J_0'(0) = 0$, the normalisation $v(\alpha, \beta; \alpha, \beta) = 1$ requires $A = 1$, and the conditions (6.4.21), (6.4.22) are then automatically satisfied. Hence the Riemann function is

$$F(\lambda) = F\big(2\sqrt{ks}\big) = v(\xi, \eta; \alpha, \beta) = J_0\Big(2\sqrt{k(\xi - \alpha)(\eta - \beta)}\Big). \tag{6.4.24}$$

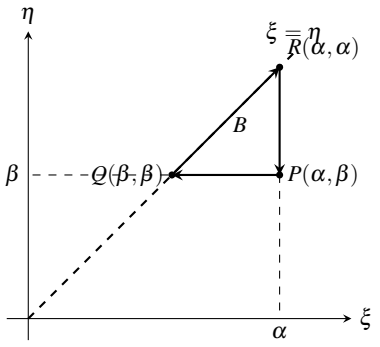


Fig. 6.8 Characteristic triangle PQR in the (ξ, η)-plane for the telegrapher's equation.

The general formal solution (6.4.12) of a hyperbolic equation in canonical form, applied to the present problem, simplifies. With $f = 0$ in (6.4.17) and $a = b = 0$, the formula reduces, along the characteristic arc $\overline{QR}$ on which $\xi = \eta$, to

$$[uv]_P = \tfrac{1}{2}\{[uv]_Q + [uv]_R\} - \tfrac{1}{2}\int_Q^R u\,(v_\xi\,d\xi - v_\eta\,d\eta) + \tfrac{1}{2}\int_Q^R v\,(u_\xi\,d\xi - u_\eta\,d\eta).$$

From (6.4.24),

$$\begin{aligned}
v_\xi &= \frac{dJ_0}{d\lambda}\,\frac{\partial\lambda}{\partial\xi} = \frac{\sqrt{k}\,(\eta-\beta)}{[(\xi-\alpha)(\eta-\beta)]^{1/2}}\,J_0'\big(2\sqrt{k(\xi-\alpha)(\eta-\beta)}\big),\\
[v_\xi]_{\xi=\eta} &= \frac{\sqrt{k}\,(\xi-\beta)}{[(\xi-\alpha)(\xi-\beta)]^{1/2}}\,\big[J_0'(\lambda)\big]_{\xi=\eta},\\
[v_\eta]_{\xi=\eta} &= \frac{\sqrt{k}\,(\xi-\alpha)}{[(\xi-\alpha)(\xi-\beta)]^{1/2}}\,\big[J_0'(\lambda)\big]_{\xi=\eta},\\
[v_\xi - v_\eta]_{\xi=\eta} &= \frac{\sqrt{k}\,(\alpha-\beta)}{[(\xi-\alpha)(\xi-\beta)]^{1/2}}\,\big[J_0'(\lambda)\big]_{\xi=\eta}.
\end{aligned} \tag{6.4.25}$$

With (6.4.18), (6.4.19), (6.4.24) and (6.4.25),

$$\begin{aligned}
u(\alpha,\beta) =& \tfrac{1}{2}\big[f(\alpha)+f(\beta)\big] - \tfrac{1}{2}\int_\beta^\alpha \frac{\sqrt{k}\,(\alpha-\beta)}{[(\tau-\alpha)(\tau-\beta)]^{1/2}}\,J_0'\Big(\big(4k(\tau-\alpha)(\tau-\beta)\big)^{1/2}\Big)f(\tau)\,d\tau\\
&+\frac{1}{2c}\int_\beta^\alpha J_0\Big(\big(4k(\tau-\alpha)(\tau-\beta)\big)^{1/2}\Big)g(\tau)\,d\tau.
\end{aligned} \tag{6.4.26}$$

Replacing (α,β) by (ξ,η) in (6.4.26),

$$\begin{aligned}
u(\xi,\eta) =& \tfrac{1}{2}\big[f(\xi)+f(\eta)\big] - \tfrac{1}{2}\int_\eta^\xi \frac{\sqrt{k}\,(\xi-\eta)}{[(\tau-\xi)(\tau-\eta)]^{1/2}}\,J_0'\Big(\big(4k(\tau-\xi)(\tau-\eta)\big)^{1/2}\Big)f(\tau)\,d\tau\\
&+\frac{1}{2c}\int_\eta^\xi J_0\Big(\big(4k(\tau-\xi)(\tau-\eta)\big)^{1/2}\Big)g(\tau)\,d\tau,
\end{aligned}$$

and the solution in the original variables is

$$u(x,t) = \tfrac{1}{2}\big[f(x+ct)+f(x-ct)\big] + \tfrac{1}{2}\int_{x-ct}^{x+ct} G(x,t,\tau)\,d\tau,$$

where, using $(\tau-\xi)(\tau-\eta) = (\tau-x)^2 - c^2t^2$,

$$\begin{aligned}
G(x,t,\tau) =& -\frac{2\sqrt{k}\,ct\,f(\tau)}{\big[(\tau-x)^2-c^2t^2\big]^{1/2}}\,J_0'\Big(\big(4k\big[(\tau-x)^2-c^2t^2\big]\big)^{1/2}\Big)\\
&+\frac{1}{c}\,g(\tau)\,J_0\Big(\big(4k\big[(\tau-x)^2-c^2t^2\big]\big)^{1/2}\Big).
\end{aligned}$$

Exercise 6.6 Solve, by Riemann's method, the Cauchy problem

$$\left(\sin^2\mu\right)\Phi_{xx} - \left(\cos^2\mu\right)\Phi_{yy} - \left(k^2\sin^2\mu\cos^2\mu\right)\Phi = 0,$$

$$\Phi(0,y) = f_1(y), \quad \Phi_x(0,y) = f_2(y), \quad \Phi(x,0) = g_1(x), \quad \Phi_y(x,0) = g_2(x).$$

Solution. The discriminant satisfies $\Delta > 0$, so the equation is hyperbolic. The substitution

$$\xi = y + (\cot\mu)x, \qquad \eta = y - (\cot\mu)x, \tag{6.4.27}$$

gives

$$\begin{aligned}
\Phi_x &= (\cot\mu)\Phi_\xi - (\cot\mu)\Phi_\eta,\\
\Phi_y &= \Phi_\xi + \Phi_\eta,\\
\Phi_{xx} &= (\cot\mu)^2\Phi_{\xi\xi} - 2(\cot\mu)^2\Phi_{\xi\eta} + (\cot\mu)^2\Phi_{\eta\eta},\\
\Phi_{yy} &= \Phi_{\xi\xi} + 2\Phi_{\xi\eta} + \Phi_{\eta\eta},
\end{aligned}$$

which reduces the equation to the canonical form

$$L[\Phi] = \Phi_{\xi\eta} + k^2\Phi = 0. \tag{6.4.28}$$

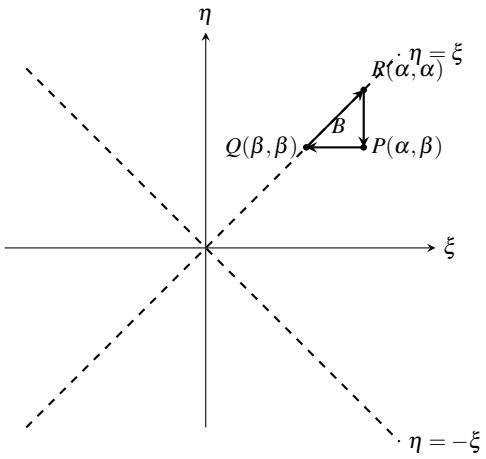


Fig. 6.9 Region B in Exercise 6.6, with characteristic diagonals $\eta = \xi$ (image of $x = 0$) and $\eta = -\xi$ (image of $y = 0$).

The transformed boundary conditions become

$$[\Phi]_{\xi=\eta} = f_1(\xi), \quad [\Phi_\xi - \Phi_\eta]_{\xi=\eta} = \tan\mu\, f_2(\xi), \quad x = 0\ (\xi = \eta), \tag{6.4.29}$$

$$[\Phi]_{\xi=-\eta} = g_1(\xi), \quad [\Phi_\xi + \Phi_\eta]_{\xi=-\eta} = g_2(\xi), \quad y = 0\ (\xi = -\eta). \tag{6.4.30}$$

The Riemann function for (6.4.28) solves

$$M[v] = v_{\xi\eta} + k^2 v = 0, \tag{6.4.31}$$

$$\left.\begin{aligned} v_\xi(\xi,\beta;\alpha,\beta) &= 0,\\ v_\eta(\alpha,\eta;\alpha,\beta) &= 0,\\ v(\alpha,\beta;\alpha,\beta) &= 1. \end{aligned}\right\} \tag{6.4.32}$$

Setting $s=(\xi-\alpha)(\eta-\beta)$ and $v(\xi,\eta;\alpha,\beta)=F(s)$, (6.4.31) becomes

$$sF''+F'+k^2F=0,$$

and with the substitution $\lambda=2k\sqrt{s}$,

$$\lambda^2F''(\lambda)+\lambda F'(\lambda)+\lambda^2F(\lambda)=0,$$

which is again the zeroth-order Bessel equation. Since $Y_0(0)=-\infty$, the bounded solution is

$$v(\xi,\eta;\alpha,\beta)=J_0\Big(2k\big[(\xi-\alpha)(\eta-\beta)\big]^{1/2}\Big). \tag{6.4.33}$$

The conditions $J_0(0)=1$ and $J_0'(0)=0$ ensure that (6.4.32) are met.

Moreover,

$$\begin{aligned} v_\xi &= \frac{dJ_0}{d\lambda}\,\frac{d\lambda}{d\xi} = \frac{k(\eta-\beta)}{[(\xi-\alpha)(\eta-\beta)]^{1/2}}\,J_0'\Big(2k\big[(\xi-\alpha)(\eta-\beta)\big]^{1/2}\Big),\\ v_\eta &= \frac{k(\xi-\alpha)}{[(\xi-\alpha)(\eta-\beta)]^{1/2}}\,J_0'\Big(2k\big[(\xi-\alpha)(\eta-\beta)\big]^{1/2}\Big),\\ [v_\xi+v_\eta]_{\xi=-\eta} &= \frac{k(\alpha-\beta)}{[(\alpha-\xi)(\beta+\eta)]^{1/2}}\,J_0'\Big(2k\big[(\alpha-\xi)(\beta+\eta)\big]^{1/2}\Big), \end{aligned} \tag{6.4.34}$$

$$[v_\xi-v_\eta]_{\xi=\eta} = \frac{k(\alpha-\beta)}{[(\xi-\alpha)(\eta-\beta)]^{1/2}}\,J_0'\Big(2k\big[(\xi-\alpha)(\eta-\beta)\big]^{1/2}\Big). \tag{6.4.35}$$

The general formal solution (6.4.12), in the canonical variables (ξ,η) and with $f=0$, is

$$\begin{aligned} [\Phi v]_P =& \tfrac{1}{2}\big\{[\Phi v]_Q+[\Phi v]_R\big\}-\tfrac{1}{2}\int_Q^R \Phi\,(v_\xi\,d\xi - v_\eta\,d\eta)\\ &+\tfrac{1}{2}\int_Q^R v\,(\Phi_\xi\,d\xi-\Phi_\eta\,d\eta). \end{aligned} \tag{6.4.36}$$

Recalling (6.4.33) we have $v(P)=v(Q)=v(R)=1$, and

$$\Phi(R)=f_1(\alpha), \qquad \Phi(Q)=g_1(\beta).$$

Combining (6.4.31), (6.4.32), (6.4.35) and (6.4.36), integration along the two characteristic arcs of Λ (i.e. along $\xi=\eta$ from R to P, and along $\xi=-\eta$ from P to Q) gives

$$\Phi(\alpha,\beta) = \tfrac{1}{2}\{f_1(\alpha)+g_1(\beta)\} - \tfrac{1}{2}\int_0^\beta g_1(\tau)\,\frac{k(\alpha-\beta)}{[(\alpha-\tau)(\beta+\tau)]^{1/2}}J_0'\Big(2k\big[(\alpha-\tau)(\beta+\tau)\big]^{1/2}\Big)d\tau$$
$$- \tfrac{1}{2}\int_0^\alpha f_1(\tau)\,\frac{k(\alpha-\beta)}{[(\tau-\alpha)(\tau-\beta)]^{1/2}}J_0'\Big(2k\big[(\tau-\alpha)(\tau-\beta)\big]^{1/2}\Big)d\tau$$
$$+ \tfrac{1}{2}\int_0^\beta g_2(\tau)\,J_0\Big(2k\big[(\alpha-\tau)(\beta+\tau)\big]^{1/2}\Big)d\tau$$
$$+ \frac{\tan\mu}{2}\int_0^\alpha f_2(\tau)\,J_0\Big(2k\big[(\tau-\alpha)(\tau-\beta)\big]^{1/2}\Big)d\tau.$$

Since (α,β) is arbitrary, replacing it by (ξ,η) gives the formal solution in canonical variables, and using (6.4.27) to return to original variables, with the identifications

$$(\xi-\tau)(\eta+\tau) = y^2-(\tau-\cot\mu\,x)^2, \qquad (\tau-\xi)(\tau-\eta) = (\tau-y)^2-(\cot\mu\,x)^2,$$

we obtain, with

$$G_1(x,y,\tau) = 2k\cot\mu\,x\left[\frac{J_0'\Big(2k\big[y^2-(\tau-\cot\mu\,x)^2\big]^{1/2}\Big)}{[y^2-(\tau-\cot\mu\,x)^2]^{1/2}}\,g_1(\tau)\right.$$
$$\left. + \frac{J_0'\Big(2k\big[(\tau-y)^2-(\cot\mu\,x)^2\big]^{1/2}\Big)}{[(\tau-y)^2-(\cot\mu\,x)^2]^{1/2}}\,f_1(\tau)\right],$$
$$G_2(x,y,\tau) = J_0\Big(2k\big[y^2-(\tau-\cot\mu\,x)^2\big]^{1/2}\Big)g_2(\tau)$$
$$+\tan\mu\,J_0\Big(2k\big[(\tau-y)^2-(\cot\mu\,x)^2\big]^{1/2}\Big)f_2(\tau),$$

the formal solution

$$\Phi(x,y) = \tfrac{1}{2}\big[f_1(y+\cot\mu\,x)+g_1(y-\cot\mu\,x)\big] - \tfrac{1}{2}\int_{y-\cot\mu\,x}^{y+\cot\mu\,x} G_1\,d\tau + \tfrac{1}{2}\int_{y-\cot\mu\,x}^{y+\cot\mu\,x} G_2\,d\tau.$$

Problems

6.1 Determine, by the method of characteristics, the displacement of a thin, half-infinite string with a free end at $x=0$, given

$$u_{tt}-c^2u_{xx}=0, \qquad 0\le x<\infty,\ t>0,$$

$$u(x,0)=f(x), \quad 0\le x<\infty; \quad u_t(x,0)=g(x), \quad 0\le x<\infty; \quad u_x(0,t)=0, \quad t>0.$$

6.2 By the method of characteristics, determine the displacement of a thin vibrating string in the inhomogeneous boundary-value problem

$$u_{tt}-c^2u_{xx}=0, \qquad x\ge 0,\ t>0,$$

$$u(x,0) = f(x), \quad x \geq 0; \quad u_t(x,0) = g(x), \quad x \geq 0; \quad u(0,t) = p(t), \quad t > 0.$$

6.3 By the method of characteristics, find the displacement of a thin string of finite length ℓ tied at both ends:

$$u_{tt} - c^2 u_{xx} = 0, \qquad 0 \leq x \leq \ell,\ t > 0,$$

$$u(x,0) = f(x), \quad u_t(x,0) = g(x), \quad u(0,t) = 0, \quad u(\ell,t) = 0.$$

6.4 Solve the following initial-value problems by the method of characteristics:

(a) $u_{tt} - 4u_{xx} = 0$, $x \geq 0$, $t > 0$; $u(x,0) = |\sin x|$, $x \geq 0$; $u_t(x,0) = 0$, $x \geq 0$; $u(0,t) = 0$, $t > 0$.

(b) $u_{tt} = u_{xx}$, $0 \leq x < \infty$, $t > 0$; $u(x,0) = \cos\left(\frac{\pi}{2}x\right)$, $0 \leq x < \infty$; $u_t(x,0) = 0$, $0 \leq x < \infty$; $u_x(0,t) = 0$, $t > 0$.

Chapter 7
The Method of Separation of Variables

Abstract This chapter develops the method of separation of variables. After deriving the conditions under which the assumption $u(x,y) = X(x)Y(y)$ reduces a second-order linear PDE to a pair of ODEs, we solve the Dirichlet problem for the vibrating string, obtaining the Fourier-series formula for its displacement. Worked exercises treat the plucked string, heat conduction in a rod, the wave equation with Neumann data, the 3D Laplace equation in a cube, the vibrating beam, the Laplace equation with Robin conditions, and a heat equation with absorption.

7.1 Solvability Conditions

Consider a second-order linear homogeneous PDE in two independent variables, given in the canonical form

$$a^*(\xi,\eta)\,u_{\xi\xi} + b^*(\xi,\eta)\,u_{\xi\eta} + c^*(\xi,\eta)\,u_{\eta\eta} + d^*(\xi,\eta)\,u_\xi + e^*(\xi,\eta)\,u_\eta + f^*(\xi,\eta)\,u = 0. \tag{7.1.1}$$

By a continuous change of variables

$$\xi = \xi(x,y), \qquad \eta = \eta(x,y), \qquad (J \neq 0), \tag{7.1.2}$$

such an equation can always be reduced to the form

$$a(x,y)\,u_{xx} + c(x,y)\,u_{yy} + d(x,y)\,u_x + e(x,y)\,u_y + f(x,y)\,u = 0, \tag{7.1.3}$$

in which, depending on the type of the equation,

1. if hyperbolic, $a = -c$,
2. if parabolic, $a = 0$ or $c = 0$,
3. if elliptic, $a = c$.

To find a solution of (7.1.3) we ask whether the unknown $u(x,y)$ may be represented as a product of a function $X = X(x)$ depending only on x and a function $Y = Y(y)$ depending only on y, both twice differentiable, in the form

$$u(x,y) = X(x)\,Y(y). \tag{7.1.4}$$

If this hypothesis is realisable—if explicit expressions for $X = X(x)$ and $Y = Y(y)$ can be determined—then (7.1.4) provides a function $u = u(x,y)$ that solves the PDE (7.1.3).

Substituting (7.1.4) into (7.1.3) and writing primes for derivatives with respect to the relevant independent variable,

$$a(x,y)X''Y + c(x,y)XY'' + d(x,y)X'Y + e(x,y)XY' + f(x,y)XY = 0. \tag{7.1.5}$$

Dividing through by $a(x,y)$,

$$X''Y \pm XY'' + \frac{d(x,y)}{a(x,y)}X'Y + \frac{c(x,y)}{a(x,y)}XY'' + \frac{f(x,y)}{a(x,y)}XY = 0, \tag{7.1.6}$$

where the sign of the second term is negative for a hyperbolic equation, positive for an elliptic equation, and the term itself is absent for a parabolic equation. Accordingly, if

$$\frac{d(x,y)}{a(x,y)} = a_1(x), \qquad \frac{c(x,y)}{a(x,y)} = b_1(y), \qquad \frac{f(x,y)}{a(x,y)} = a_2(x) + b_2(y), \tag{7.1.7}$$

then equation (7.1.5) reduces to

$$X''Y \pm XY'' + a_1(x)X'Y + b_1(y)XY' + \big(a_2(x) + b_2(y)\big)XY = 0, \tag{7.1.8}$$

which, after division by XY, separates into terms depending only on x and only on y:

$$\frac{X''}{X} + a_1(x)\frac{X'}{X} + a_2(x) = -\left[\pm\frac{Y''}{Y} + b_1(y)\frac{Y'}{Y} + b_2(y)\right]. \tag{7.1.9}$$

Differentiating both sides of this equality with respect to, say, x gives

$$\frac{d}{dx}\left\{\frac{X''}{X} + a_1(x)\frac{X'}{X} + a_2(x)\right\} = 0,$$

so that, with λ an arbitrary constant,

$$\frac{X''}{X} + a_1(x)\frac{X'}{X} + a_2(x) = \lambda, \tag{7.1.10}$$

and consequently, in view of (7.1.9),

$$\mp\frac{Y''}{Y} + b_1(y)\frac{Y'}{Y} + b_2(y) = -\lambda. \tag{7.1.11}$$

Rearranging (7.1.9) and (7.1.10),

$$X'' + a_1(x)X' + \big[a_2(x) - \lambda\big]X = 0, \tag{7.1.12}$$

$$\mp Y'' + b_1(y)Y' + \big[b_2(y) + \lambda\big]Y = 0. \tag{7.1.13}$$

Hence if $X = X(x)$ and $Y = Y(y)$ are solutions of the ordinary differential equations (7.1.12) and (7.1.13) respectively, then $u = u(x,y) = X(x)Y(y)$ is a solution of the PDE (7.1.3). (One should note that for $c = 0$, when (7.1.3) is parabolic, (7.1.13) is a first-order equation.) Thus the solution of the PDE (7.1.3) has been reduced to the solution of as many ODEs as there are independent variables.

The reduction of an arbitrary linear second-order PDE with given coefficients to the canonical form is therefore not strictly necessary for the application of the method, but it does help to lessen the number of separability conditions (7.1.7).

If the given two-variable, second-order homogeneous linear PDE

$$A\,u_{xx} + B\,u_{xy} + C\,u_{yy} + D\,u_x + E\,u_y + F\,u = 0 \tag{7.1.14}$$

has constant coefficients, then it admits a solution by separation of variables without first being reduced to canonical form. Indeed, assuming

$$u(x,y) = X(x)\,Y(y),$$

substitution into (7.1.14) gives

$$AX''Y + BX'Y' + CXY'' + DX'Y + EXY' + FXY = 0,$$

and dividing by AXY one finds

$$\frac{X''}{X} + \frac{B}{A}\frac{X'}{X}\frac{Y'}{Y} + \frac{C}{A}\frac{Y''}{Y} + \frac{D}{A}\frac{X'}{X} + \frac{E}{A}\frac{Y'}{Y} + \frac{F}{A} = 0. \tag{7.1.15}$$

Differentiating (7.1.15) with respect to x,

$$\left(\tfrac{X''}{X}\right)' + \frac{B}{A}\left(\tfrac{X'}{X}\right)'\frac{Y'}{Y} + \frac{D}{A}\left(\tfrac{X'}{X}\right)' = 0,$$

and rearranging,

$$\frac{\left(X''/X\right)'}{(B/A)\left(X'/X\right)'} + \frac{D}{B} = -\frac{Y'}{Y}.$$

Differentiating now with respect to x,

$$\frac{d}{dx}\left\{\frac{\left(X''/X\right)'}{(B/A)\left(X'/X\right)'} + \frac{D}{B}\right\} = 0,$$

so that, λ being an arbitrary constant,

$$\left(\tfrac{X''}{X}\right)' + \left(\tfrac{D}{B} - \lambda\right)\frac{B}{A}\left(\tfrac{X'}{X}\right)' = 0, \tag{7.1.16}$$

together with

$$Y' + \lambda Y = 0. \tag{7.1.17}$$

Integrating (7.1.16) with respect to x introduces a second constant β and gives

$$\frac{X''}{X}+\Big(\frac{D}{B}-\lambda\Big)\frac{B}{A}\frac{X'}{X}=\beta,$$

or equivalently

$$X''+\Big(\frac{D}{B}-\lambda\Big)\frac{B}{A}X'-\beta X=0. \tag{7.1.18}$$

Substituting now $Y'=-\lambda Y$, $Y''=\lambda^2 Y$ from (7.1.17) into (7.1.15) and rearranging,

$$\frac{X''}{X}+\frac{B}{A}\frac{X'}{X}(-\lambda)+\frac{C}{A}\lambda^2+\frac{D}{A}\frac{X'}{X}-\frac{E}{A}\lambda+\frac{F}{A}=0,$$

and so

$$X''+\Big(\frac{D}{B}-\lambda\Big)\frac{B}{A}X'+\Big(\lambda^2-\frac{E}{C}\lambda+\frac{F}{C}\Big)\frac{C}{A}X=0. \tag{7.1.19}$$

Comparison of (7.1.18) and (7.1.19) shows that

$$\beta=-\left(\lambda^2-\frac{E}{C}\lambda+\frac{F}{C}\right)\frac{C}{A}. \tag{7.1.20}$$

Hence in separating the two-variable equation (7.1.14) only one separation constant is introduced. This observation extends to every separable equation: an n-variable PDE separating into n ordinary differential equations introduces $n-1$ separation constants.

Accordingly, if $X=X(x)$ and $Y=Y(y)$ are solutions of the ordinary differential equations (7.1.17) and (7.1.19) respectively, then the solution of the constant-coefficient PDE (7.1.14) of second order is $u=u(x,y)=X(x)Y(y)$.

We now turn briefly to the boundary conditions accompanying a PDE. Although a great variety of boundary conditions occur, the most frequent in mathematical physics fall into three classes; for a problem in n variables they take the forms

(1) Dirichlet (first kind):	$[u]_{x=x_0}=\alpha,$
(2) Neumann (second kind):	$[u_x]_{x=x_0}=\beta,$
(3) Mixed/Robin (third kind):	$[u_x+hu]_{x=x_0}=\gamma,$

where α,β,γ are functions independent of x (depending on the remaining variables) and may even vanish. In the latter case the conditions are called *homogeneous*. If all the boundary data of a problem are of the same type the problem is called the Dirichlet, the Neumann, or the third boundary-value problem; if conditions of different types occur together one speaks of a *Robin problem*.

In framing the boundary conditions in the context of separation of variables it is essential to choose a coordinate system suited to the boundary: rectangular Cartesian coordinates for a rectangular region, spherical coordinates for a spherical region, and so on.

7.2 The Dirichlet Problem for the Vibrating String

Consider a flexible string of length ℓ fastened at both ends, and the initial-boundary-value problem

$$u_{tt} - \alpha^2 u_{xx} = 0, \qquad 0 < x < \ell,\ t > 0, \tag{7.2.1}$$

$$\left.\begin{aligned} u(x,0) &= f(x) \\ u_t(x,0) &= g(x) \end{aligned}\right\} \text{ (initial data)} \qquad 0 \le x \le \ell, \tag{7.2.2}$$

$$\left.\begin{aligned} u(0,t) &= 0 \\ u(\ell,t) &= 0 \end{aligned}\right\} \text{ (first-kind boundary conditions)} \qquad t > 0. \tag{7.2.3}$$

By the method of separation of variables, assume $u(x,t) = X(x)\,T(t)$. Substitution into the equation gives

$$XT'' = \alpha^2 TX'',$$

and dividing by $\alpha^2 XT \neq 0$,

$$\frac{X''}{X} = \frac{1}{\alpha^2}\frac{T''}{T}.$$

Since the left side depends only on x and the right only on t, both sides equal a separation constant λ:

$$\frac{X''}{X} = \frac{1}{\alpha^2}\frac{T''}{T} = \lambda,$$

whence the two ordinary differential equations satisfied by $X(x)$ and $T(t)$ are

$$X'' - \lambda X = 0, \tag{7.2.4}$$

$$T'' - \lambda\alpha^2 T = 0. \tag{7.2.5}$$

The boundary conditions (7.2.3) read

$$u(0,t) = X(0)T(t) = 0, \qquad u(\ell,t) = X(\ell)T(t) = 0,$$

and since $T(t) \not\equiv 0$,

$$X(0) = X(\ell) = 0, \tag{7.2.6}$$

which are homogeneous boundary conditions for the function $X(x)$. Thus, in view of (7.2.4) and (7.2.6), the eigenvalue problem determining $X(x)$ is

$$\left.\begin{aligned} X'' - \lambda X &= 0, \\ X(0) &= 0, \\ X(\ell) &= 0. \end{aligned}\right\} \tag{7.2.7}$$

The values of λ for which this Sturm–Liouville problem admits non-trivial solutions are determined by examining its sign in three cases. In each case the general solution of the ODE $X'' - \lambda X = 0$ is read off from the constant-coefficient results of Section 1.2 (cf. equations (1.2.5) and (1.2.6)).

(i) $\lambda > 0$. With A, B integration constants, the general solution is

$$X(x) = Ae^{-\sqrt{\lambda}x} + Be^{\sqrt{\lambda}x}.$$

Imposing the boundary conditions yields the two homogeneous algebraic equations

$$X(0) = A + B = 0, \qquad X(\ell) = Ae^{-\sqrt{\lambda}\ell} + Be^{\sqrt{\lambda}\ell} = 0.$$

The determinant of the coefficient matrix of this system is

$$\begin{vmatrix} 1 & 1 \\ e^{-\sqrt{\lambda}\ell} & e^{\sqrt{\lambda}\ell} \end{vmatrix} = e^{\sqrt{\lambda}\ell} - e^{-\sqrt{\lambda}\ell} = 2\sinh\sqrt{\lambda}\,\ell \neq 0,$$

so the only solution is the trivial one $A = B = 0$, giving $X(x) \equiv 0$.

(ii) $\lambda = 0$. Now

$$X(x) = A + Bx,$$

and the boundary conditions

$$X(0) = A = 0, \qquad X(\ell) = B\ell = 0$$

again force $X(x) \equiv 0$.

(iii) $\lambda < 0$. Writing $\lambda = -\omega^2$ ($\omega > 0$), the general solution is

$$X(x) = A\cos\omega x + B\sin\omega x.$$

Applying the boundary conditions,

$$X(0) = A = 0, \qquad X(x) = B\sin\omega x,$$

and

$$X(\ell) = B\sin\omega\ell = 0.$$

Non-trivial solutions ($B \neq 0$) require

$$\sin\omega\ell = 0,$$

which yields the values

$$\omega_n = \frac{n\pi}{\ell}, \qquad (n = 1, 2, \ldots),$$

and hence the eigenvalues

$$\lambda_n = -\left(\frac{n\pi}{\ell}\right)^2, \qquad (n = 1, 2, \ldots). \tag{7.2.8}$$

The corresponding eigenfunctions of the Sturm–Liouville problem (7.2.7) are

$$X_n(x) = B_n \sin\frac{n\pi x}{\ell}.$$

For each λ_n, equation (7.2.5) has the solution

$$T_n(t) = C_n \cos\frac{n\pi\alpha}{\ell}t + D_n \sin\frac{n\pi\alpha}{\ell}t, \tag{7.2.9}$$

with C_n, D_n arbitrary constants. The corresponding solutions of the original problem that satisfy the homogeneous boundary conditions are, with $a_n^* \equiv B_n C_n$ and $b_n^* \equiv B_n D_n$,

$$u_n(x,t) = X_n(x)\,T_n(t) = \Big(a_n^* \cos\frac{n\pi}{\ell}\alpha t + b_n^* \sin\frac{n\pi}{\ell}\alpha t\Big)\sin\frac{n\pi}{\ell}x. \tag{7.2.10}$$

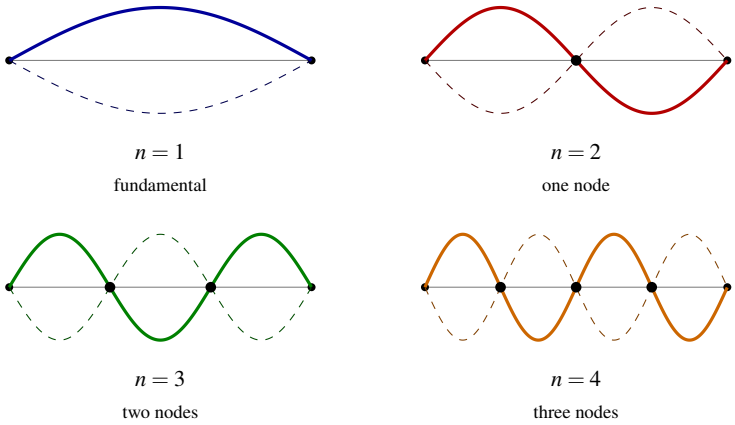


Fig. 7.1 The first four eigenmodes of the vibrating string, $X_n(x) = \sin(n\pi x/\ell)$, drawn at the instant of maximum displacement. The end-points $x = 0$ and $x = \ell$ are clamped (filled dots); the dashed curves show the same modes at the opposite phase of the oscillation. Mode n has $n-1$ interior nodes (small black dots), oscillates with the angular frequency $\omega_n = n\pi\alpha/\ell$, and is the simplest physical realisation of an eigenfunction of the Sturm–Liouville problem (7.2.7). The general solution of the vibrating string is the linear superposition of these modes, weighted by the Fourier coefficients of the initial data.

Since the PDE (7.2.1) is linear, a linear combination of solutions of the form (7.2.10) is also a solution; hence

$$u(x,t) = \sum_{n=1}^{\infty} \Big(a_n \cos\frac{n\pi}{\ell}\alpha t + b_n \sin\frac{n\pi}{\ell}\alpha t\Big)\sin\frac{n\pi}{\ell}x \tag{7.2.11}$$

is a solution satisfying the boundary conditions. The coefficients of this linear combination are determined from the inhomogeneous initial conditions (7.2.2). From them,

$$u(x,0) = f(x) = \sum_{n=1}^{\infty} a_n \sin\frac{n\pi x}{\ell}, \tag{7.2.12}$$

$$u_t(x,0) = g(x) = \sum_{n=1}^{\infty} \Big(\frac{n\pi\alpha}{\ell}\Big) b_n \sin\frac{n\pi x}{\ell}. \tag{7.2.13}$$

The initial data $f(x)$ and $g(x)$ are thus represented by Fourier sine series. The expansion coefficients of (7.2.12) and (7.2.13) are

$$a_n = \frac{2}{\ell}\int_0^\ell f(x)\sin\frac{n\pi x}{\ell}\,dx, \tag{7.2.14}$$

$$b_n = \frac{2}{n\pi\alpha}\int_0^\ell g(x)\sin\frac{n\pi x}{\ell}\,dx. \tag{7.2.15}$$

The solution of the given initial-boundary-value problem may therefore be written

$$\begin{aligned} u(x,t) &= \sum_{n=1}^{\infty}\left(a_n\cos\frac{n\pi\alpha}{\ell}t + b_n\sin\frac{n\pi\alpha}{\ell}t\right)\sin\frac{n\pi}{\ell}x \\ &= \frac{2}{\ell}\sum_{n=1}^{\infty}\left\{\int_0^\ell f(\tau)\sin\frac{n\pi}{\ell}\tau\,d\tau\right\}\cos\frac{n\pi\alpha}{\ell}t\,\sin\frac{n\pi}{\ell}x \\ &\quad + \frac{2}{\pi\alpha}\sum_{n=1}^{\infty}\frac{1}{n}\left\{\int_0^\ell g(\tau)\sin\frac{n\pi}{\ell}\tau\,d\tau\right\}\sin\frac{n\pi\alpha}{\ell}t\,\sin\frac{n\pi}{\ell}x. \end{aligned} \tag{7.2.16}$$

Worked Exercises

Exercise 7.1 A flexible thin string of length ℓ, fastened at both ends $x = 0$ and $x = \ell$, is plucked at $x = a$ to height h as shown in Fig. 7.2. Determine the resulting transverse oscillations.

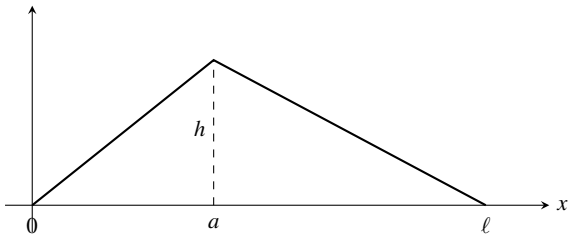


Fig. 7.2 Initial triangular profile of the plucked string in Exercise V.1.

Solution. The problem describing the motion of the string is

$$\left.\begin{aligned} &u_{tt} - \alpha^2 u_{xx} = 0, \\ &u(x,0) = f(x) = \begin{cases} \dfrac{hx}{a}, & 0 \le x \le a, \\ \dfrac{h(\ell - x)}{\ell - a}, & a \le x \le \ell, \end{cases} \\ &u_t(x,0) = 0, \\ &u(0,t) = 0, \\ &u(\ell,t) = 0. \end{aligned}\right\}$$

Following the method of separation of variables, set $u(x,t) = X(x)\,T(t)$, whence

$$\frac{T''}{\alpha^2 T} = \frac{X''}{X} = \lambda = -\omega^2,$$

so that

$$X'' + \omega^2 X = 0, \tag{7.2.17}$$

$$T'' + \omega^2 \alpha^2 T = 0. \tag{7.2.18}$$

The general solution of (7.2.17) is

$$X(x) = A\cos\omega x + B\sin\omega x,$$

and the boundary conditions $X(0) = X(\ell) = 0$ give

$$\omega_n = \frac{n\pi}{\ell}, \qquad (n = 1, 2, \ldots),$$

so that the eigenfunctions are

$$X_n(x) = B_n \sin\frac{n\pi}{\ell} x.$$

The general solution of (7.2.18) is

$$T_n(t) = C_n \cos\frac{n\pi}{\ell}\alpha t + D_n \sin\frac{n\pi}{\ell}\alpha t,$$

and with $a_n^* \equiv B_n C_n$, $b_n^* \equiv B_n D_n$ the solutions satisfying the boundary conditions are

$$u_n(x,t) = \left(a_n^* \cos\frac{n\pi\alpha}{\ell} t + b_n^* \sin\frac{n\pi\alpha}{\ell} t\right)\sin\frac{n\pi}{\ell} x. \tag{7.2.19}$$

Since the PDE is linear, the superposition

$$u(x,t) = \sum_{n=1}^{\infty} \left(a_n \cos\frac{n\pi\alpha}{\ell} t + b_n \sin\frac{n\pi\alpha}{\ell} t\right)\sin\frac{n\pi}{\ell} x$$

is again a solution satisfying the boundary conditions. The expansion coefficients are obtained from the initial conditions: from

$$u(x,0)=\sum_{n=1}^{\infty}a_n\sin\frac{n\pi}{\ell}x=\begin{cases}\dfrac{hx}{a}, & 0\le x\le a,\\ \dfrac{h(\ell-x)}{\ell-a}, & a\le x\le \ell,\end{cases}$$

one finds

$$a_n=\frac{h\ell^3}{\pi^2a(\ell-a)}\frac{1}{n^2}\sin\frac{n\pi a}{\ell},\qquad (n=1,2,\ldots),$$

while $u_t(x,0)=0$ gives

$$b_n\equiv 0,\qquad (n=1,2,\ldots).$$

Hence the displacement of the plucked string is

$$u(x,t)=\frac{h\ell^3}{\pi^2a(\ell-a)}\sum_{n=1}^{\infty}\left(\frac{1}{n^2}\sin\frac{n\pi a}{\ell}\right)\sin\frac{n\pi}{\ell}x\cos\frac{n\pi\alpha}{\ell}t. \tag{7.2.20}$$

Exercise 7.2 Determine the temperature distribution $u(x,t)$ in a homogeneous (uniform) thin rod of length ℓ subject to the initial-boundary conditions

$$u(0,t)=0,\qquad u(\ell,t)=0,\qquad u(x,0)=f(x)\quad (0\le x\le\ell).$$

Solution. The temperature satisfies

$$\left.\begin{aligned}&u_t=ku_{xx},\quad (0\le x\le\ell),\\ &u(x,0)=f(x),\\ &u(0,t)=u(\ell,t)=0.\end{aligned}\right\}$$

By separation of variables, $u(x,t)=X(x)T(t)$. With λ the separation constant, the equation reduces to the two ODEs

$$X''-\lambda X=0,\qquad T'-k\lambda T=0.$$

The boundary conditions are homogeneous Dirichlet conditions; since the temperature distribution must decay in time, $\lambda=-\omega^2<0$ is required. The solution of the temporal equation is then

$$T=Ce^{-k\omega^2t},$$

and the spatial equation under $X(0)=X(\ell)=0$ has the solutions

$$X_n(x)=B_n\sin\frac{n\pi}{\ell}x,\qquad (n=1,2,3,\ldots).$$

For $n=1,2,3,\ldots$,

$$u_n(x,t)=\gamma_n\sin\frac{n\pi}{\ell}xe^{-k\left(\frac{n\pi}{\ell}\right)^2t},\qquad (C_nB_n\equiv\gamma_n).$$

A linear combination of these particular solutions is also a solution of the equation:

$$u(x,t)=\sum_{n=1}^{\infty}\beta_n e^{-k\left(\frac{n\pi}{\ell}\right)^2 t}\sin\frac{n\pi}{\ell}x, \qquad (\gamma_n E_n\equiv\beta_n).$$

The initial condition gives

$$u(x,0)=f(x)=\sum_{n=1}^{\infty}\beta_n\sin\frac{n\pi}{\ell}x.$$

Multiplying both sides by $\sin(m\pi x/\ell)$, integrating over $[0,\ell]$, and using the orthogonality relation of the family $\{\sin(m\pi x/\ell)\}$,

$$\int_0^{\ell}\sin\frac{n\pi x}{\ell}\sin\frac{m\pi x}{\ell}\,dx=\frac{\ell}{2}\,\delta_{mn},$$

one obtains

$$\beta_n=\frac{2}{\ell}\int_0^{\ell}f(x)\sin\frac{n\pi x}{\ell}\,dx.$$

Hence the general solution of the problem is

$$u(x,t)=\frac{2}{\ell}\sum_{n=1}^{\infty}\left[\int_0^{\ell}f(\xi)\sin\frac{n\pi}{\ell}\xi\,d\xi\right]e^{-k\left(\frac{n\pi}{\ell}\right)^2 t}\sin\frac{n\pi}{\ell}x.$$

Exercise 7.3 Solve the following problem by the method of separation of variables:

$$\left.\begin{aligned}&u_{tt}-c^2u_{xx}=0, \quad 0\le x\le\ell,\ t>0,\\ &u(x,0)=\cos x, \quad u_t(x,0)=0,\\ &u_x(0,t)=u_x(\pi,t)=0.\end{aligned}\right\}$$

Solution. By the method of separation of variables, set $u(x,t)=X(x)T(t)$. Letting λ be the separation constant, the equation separates into the two ODEs

$$X''-\lambda X=0, \qquad T''-c^2\lambda T=0. \tag{7.2.21}$$

The boundary conditions $X'(0)=X'(\pi)=0$ for the first equation are satisfied for $\sqrt{\lambda}=im$, i.e. $\lambda=-m^2$, by

$$X_m(x)=A_m\left(e^{imx}+e^{-imx}\right)=\bar{A}_m\cos mx. \tag{7.2.22}$$

The second equation, under the homogeneous initial condition $T_m'(0)=0$, has the solution

$$T_m(t)=C_m\cos mct. \tag{7.2.23}$$

With $a_m^*\equiv\bar{A}_mC_m$,

$$u_m(x,t)=X_m(x)\,T_m(t)=a_m^*\cos mct\,\cos mx. \tag{7.2.24}$$

In contrast to the previous problem, the present one has Neumann conditions, so we must investigate whether $\lambda=0$ produces non-trivial solutions. For $\lambda=0$ the first

equation in (7.2.21) has the solution

$$X_0(x) = A_0 x + B_0,$$

and

$$X_0'(0) = X_0'(\pi) = A_0 = 0$$

gives $X_0(x) = B_0$. The second equation in (7.2.21) has, for $\lambda = 0$, the solution

$$T_0(t) = C_0 t + D_0,$$

and the condition $T_0'(0) = C_0 = 0$ leaves $T_0(t) = D_0$. Hence the $\lambda = 0$ solution of the PDE is

$$u_0(x,t) = B_0 D_0 \equiv \frac{a_0}{2}. \tag{7.2.25}$$

A linear combination of the solutions (7.2.24) and (7.2.25) is also a solution of the problem; so

$$u_m(x,t) = \frac{a_0}{2} + \sum_{m=1}^{\infty} a_m \cos mct \cos mx \tag{7.2.26}$$

satisfies all the conditions except the inhomogeneous initial one. Imposing the initial condition,

$$u(x,0) = \cos x = \frac{a_0}{2} + \sum_{m=1}^{\infty} a_m \cos mx,$$

integrating over $[0,\pi]$ in x yields $a_0 = 0$, and multiplying both sides by $\cos mx$ and integrating gives

$$a_m = \delta_{m1} = \begin{cases} 1, & m = 1, \\ 0, & m \neq 1. \end{cases}$$

Hence the solution of the problem is

$$u(x,t) = \cos ct \cos x = \tfrac{1}{2}\big[\cos(x-ct) + \cos(x+ct)\big].$$

Exercise 7.4 Solve the following problem by the method of separation of variables:

$$\left.\begin{aligned} &\nabla^2 u(x,y,z) = 0, \quad 0 \le x \le \pi,\ 0 \le y \le \pi,\ 0 \le z \le \pi, \\ &u(0,y,z) = u(\pi,y,z) = 0, \\ &u(x,0,z) = u(x,\pi,z) = 0, \\ &u(x,y,\pi) = 0, \\ &u(x,y,0) = f(x,y). \end{aligned}\right\}$$

Solution. For the solution of the boundary-value problem we set, by the method of separation of variables,

$$u(x,y,z) = X(x)Y(y)Z(z).$$

With λ and μ the separation constants, the equation separates into the three ODEs

$$X'' - \mu X = 0, \qquad Y'' - (\lambda - \mu)Y = 0, \qquad Z'' + \lambda Z = 0.$$

The first equation, under the boundary condition $X(0) = X(\pi) = 0$, has the solution

$$X_m(x) = \alpha_m \sin mx, \qquad (m = 1, 2, 3, \ldots).$$

The second equation, under $Y(0) = Y(\pi) = 0$, has the solution

$$Y_n(y) = \beta_n \sin ny, \qquad (n = 1, 2, 3, \ldots),$$

and the third equation, under $Z(\pi) = 0$, has the solution

$$Z_{mn}(z) = \gamma_{mn} \sin\bigl[\sqrt{m^2 + n^2}\,(\pi - z)\bigr].$$

For $m, n = 1, 2, 3, \ldots$,

$$u_{mn}(x, y, z) = X_m(x)\, Y_n(y)\, Z_{mn}(z) = a_{mn} \sin mx \sin ny \sin\bigl[\sqrt{m^2 + n^2}\,(\pi - z)\bigr],$$

and a linear combination of all these particular solutions,

$$u(x, y, z) = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} a_{mn} \sin mx \sin ny \sin\bigl[\sqrt{m^2 + n^2}\,(\pi - z)\bigr],$$

solves the equation under the homogeneous boundary conditions. Imposing the inhomogeneous condition in z,

$$u(x, y, 0) = f(x, y) = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \bigl(a_{mn} \sin \sqrt{m^2 + n^2}\,\pi\bigr) \sin mx \sin ny.$$

Multiplying both sides by $\sin mx$ and $\sin ny$, integrating over $[0, \pi]$ in each variable, and using the orthogonality of the families $\{\sin mx\}$ and $\{\sin ny\}$ on $[0, \pi]$,

$$a_{mn} = \frac{4}{\pi^2 \sinh\bigl(\sqrt{m^2 + n^2}\,\pi\bigr)} \int_0^{\pi} \int_0^{\pi} f(x, y) \sin mx \sin ny \, dx \, dy.$$

Hence the solution of the problem is

$$u(x, y, z) = \frac{4}{\pi^2} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \left[\int_0^{\pi} \int_0^{\pi} f(\xi, \eta) \sin m\xi \sin n\eta \, d\xi \, d\eta \right] \frac{\sin \sqrt{m^2 + n^2}\,(\pi - z)}{\sinh \sqrt{m^2 + n^2}\,\pi} \sin mx \sin ny.$$

Exercise 7.5 Solve the following problem by the method of separation of variables:

$$\left.\begin{aligned} &u_{tt} + a^2 u_{xxxx} = 0, \quad 0 < x < \ell,\ t > 0, \\ &u(x, 0) = f(x), \\ &u_t(x, 0) = g(x), \\ &u(0, t) = u(\ell, t) = 0, \\ &u_{xx}(0, t) = u_{xx}(\ell, t) = 0. \end{aligned}\right\}$$

Solution. By the method of separation of variables, set $u(x,t) = X(x)\,T(t)$. With α the separation constant, the equation separates into the two ODEs

$$X'''' - \alpha X = 0, \qquad T'' + \alpha a^2 T = 0. \tag{7.2.27}$$

With $\alpha^{1/4} \equiv \beta$, the general solution of the spatial equation is

$$X(x) = Ae^{\beta x} + Be^{-\beta x} + Ce^{i\beta x} + De^{-i\beta x}. \tag{7.2.28}$$

The boundary conditions are

$$X(0) = X(\ell) = 0, \qquad X''(0) = X''(\ell) = 0. \tag{7.2.29}$$

The second derivative of (7.2.28) is

$$X''(x) = \beta^2\big(Ae^{\beta x} + Be^{-\beta x}\big) - \beta^2\big(Ce^{i\beta x} + De^{-i\beta x}\big). \tag{7.2.30}$$

At $x = 0$, the conditions (7.2.29) applied to (7.2.28) and (7.2.30) give $B = -A$ and $D = -C$, and at $x = \ell$ they yield the homogeneous algebraic system in A, C

$$\left.\begin{aligned} A\big(e^{\beta\ell} - e^{-\beta\ell}\big) + C\big(e^{i\beta\ell} - e^{-i\beta\ell}\big) = 0, \\ A\big(e^{\beta\ell} - e^{-\beta\ell}\big) - C\big(e^{i\beta\ell} - e^{-i\beta\ell}\big) = 0. \end{aligned}\right\}$$

Adding and subtracting these equations,

$$A\big(e^{\beta\ell} - e^{-\beta\ell}\big) = 0, \qquad C\big(e^{i\beta\ell} - e^{-i\beta\ell}\big) = 0.$$

The first equation forces $A = 0$ (since $e^{\beta\ell} - e^{-\beta\ell} = 2\sinh\beta\ell \neq 0$ for $\beta \neq 0$); the second, since $A \neq 0$ is required for non-trivial solutions, gives

$$\beta_m = i\frac{m\pi}{\ell}, \qquad (m = 1, 2, \ldots).$$

Carrying these values into the second equation, $C \equiv 0$ is obtained, and consequently

$$X_m(x) = \bar{A}_m \sin\frac{m\pi}{\ell}x. \tag{7.2.31}$$

On the other hand, the temporal equation in (7.2.27),

$$T'' + a^2\beta^4 T = 0,$$

has the solution

$$T_m(t) = E_m \cos a\Big(\frac{m\pi}{\ell}\Big)^2 t + F_m \sin a\Big(\frac{m\pi}{\ell}\Big)^2 t. \tag{7.2.32}$$

With (7.2.31) and (7.2.32), the possible solutions of the problem satisfying the boundary conditions, with the constants relabelled, are

$$u_m(x,t) = X_m(x)\,T_m(t) = \sum_{m=1}^{\infty}\left(\bar{E}_m \cos a\Big(\frac{m\pi}{\ell}\Big)^2 t + \bar{F}_m \sin a\Big(\frac{m\pi}{\ell}\Big)^2 t\right)\sin\frac{m\pi}{\ell}x. \tag{7.2.33}$$

Since the PDE is linear, the linear combination of the solutions (7.2.33) is also a solution of the equation; the coefficients of this linear combination are determined by the inhomogeneous initial conditions. Accordingly,

$$u(x,t) = \sum_{m=1}^{\infty}\left(E_m^* \cos a\Big(\frac{m\pi}{\ell}\Big)^2 t + F_m^* \sin a\Big(\frac{m\pi}{\ell}\Big)^2 t\right)\sin\frac{m\pi}{\ell}x. \tag{7.2.34}$$

Imposing the initial conditions on this expression,

$$u(x,0) = f(x) = \sum_{m=1}^{\infty} E_m^* \sin\frac{m\pi}{\ell}x, \tag{7.2.35}$$

$$u_t(x,0) = g(x) = \sum_{m=1}^{\infty} a\Big(\frac{m\pi}{\ell}\Big)^2 F_m^* \sin\frac{m\pi}{\ell}x.$$

The realisability of the initial conditions hinges on the representability of $f(x)$ and $g(x)$ by Fourier sine series; recall from (1.3.3) that, for piecewise smooth functions, this is guaranteed by Dirichlet's theorem. The expansion coefficients of (7.2.34) are

$$E_m^* = \frac{2}{\ell}\int_0^{\ell} f(x)\sin\frac{m\pi}{\ell}x\,dx, \qquad \bar{F}_m = \frac{2}{a\ell}\left(\frac{\ell}{m\pi}\right)^2 \int_0^{\ell} g(x)\sin\frac{m\pi}{\ell}x\,dx,$$

and inserting these into (7.2.34) gives the desired solution of the problem.

Exercise 7.6 Solve the following Laplace problem by the method of separation of variables:

$$\left.\begin{aligned} &\nabla^2 u = 0, \quad 0<x<\pi,\ 0<y<\pi,\ 0<z<\pi,\\ &u(0,y,z) = f(y,z), \quad u(\pi,y,z) = 0,\\ &u_y(x,0,z) = u_y(x,\pi,z) = 0,\\ &u_z(x,y,0) + ku(x,y,0) = 0,\\ &u_z(x,y,\pi) + ku(x,y,\pi) = 0. \end{aligned}\right\}$$

Solution. For the solution of the boundary-value problem set

$$u(x,y,z) = X(x)\,Y(y)\,Z(z).$$

With λ and μ the separation constants, the Laplace equation separates into the three ODEs

$$\left.\begin{aligned} &X'' - \mu X = 0,\\ &Y'' - (\lambda-\mu)Y = 0,\\ &Z'' + \lambda Z = 0, \end{aligned}\right\} \tag{7.2.36}$$

whose solutions are

$$\left.\begin{aligned} X(x) &= A_1 e^{\sqrt{\mu}\,x} + B_1 e^{-\sqrt{\mu}\,x}, \\ Y(y) &= A_2 e^{\sqrt{\lambda-\mu}\,y} + B_2 e^{-\sqrt{\lambda-\mu}\,y}, \\ Z(z) &= A_3 e^{i\sqrt{\lambda}\,z} + B_3 e^{-i\sqrt{\lambda}\,z}. \end{aligned}\right\} \tag{7.2.37}$$

With $X(\pi) = 0$, the first of (7.2.37) gives

$$X(x) = A_1\left[e^{\sqrt{\mu}\,x} - e^{-\sqrt{\mu}\,(x-2\pi)}\right]. \tag{7.2.38}$$

On the other hand, from the second of (7.2.37),

$$Y'(y) = \sqrt{\lambda-\mu}\left[A_2 e^{\sqrt{\lambda-\mu}\,y} - B_2 e^{-\sqrt{\lambda-\mu}\,y}\right],$$

so $Y'(0) = 0$ gives $B_2 = A_2$, and the condition $Y'(\pi) = 0$ then yields

$$Y'(\pi) = \sqrt{\lambda-\mu}\left[A_2 e^{\sqrt{\lambda-\mu}\,\pi} - B_2 e^{-\sqrt{\lambda-\mu}\,\pi}\right] = 0,$$

whence, $A_2 \neq 0$,

$$e^{2\sqrt{\lambda-\mu}\,\pi} = 1 = e^{2im\pi}, \qquad (m = 1,2,\ldots),$$

and $\sqrt{\lambda-\mu} = im$. (For $m = 0$ one has $Y_0(y) \equiv 0$.) Inserting this value of $\sqrt{\lambda-\mu}$ into the second of (7.2.37),

$$Y_m(y) = 2A_2 \cos my. \tag{7.2.39}$$

For the solution $Z(z)$ of the third equation in (7.2.37), the Robin conditions

$$Z'(0) + kZ(0) = 0, \qquad Z'(\pi) + kZ(\pi) = 0,$$

yield the system

$$\left.\begin{aligned} \left(i\sqrt{\lambda}+k\right)A_3 - \left(i\sqrt{\lambda}-k\right)B_3 &= 0, \\ \left(i\sqrt{\lambda}+k\right)e^{i\sqrt{\lambda}\,\pi}A_3 - \left(i\sqrt{\lambda}-k\right)e^{-i\sqrt{\lambda}\,\pi}B_3 &= 0. \end{aligned}\right\}$$

The condition for non-trivial solutions $A_3, B_3 \not\equiv 0$ is the vanishing of the coefficient determinant:

$$\begin{vmatrix} \left(i\sqrt{\lambda}+k\right) & -\left(i\sqrt{\lambda}-k\right) \\ \left(i\sqrt{\lambda}+k\right)e^{i\sqrt{\lambda}\,\pi} & -\left(i\sqrt{\lambda}-k\right)e^{-i\sqrt{\lambda}\,\pi} \end{vmatrix} = 0.$$

Hence

$$e^{2i\sqrt{\lambda}\,\pi} = \frac{\lambda+k^2}{\lambda+k^2} = 1 = e^{2in\pi}, \qquad (n = 1,2,\ldots),$$

giving $\sqrt{\lambda} = n$, $\lambda = n^2$. Here $n \neq 0$, since for $n = 0$ the corresponding Z_0 vanishes under the given conditions. Hence

$$Z_n(z) = \left(A_3 e^{inz} + B_3 e^{-inz}\right) = \bar{A}_3 \cos nz + \bar{B}_3 \sin nz,$$

and the first of the third-kind boundary conditions then gives

$$Z_n(z) = \bar{A}_3\left(\cos nz - \frac{k}{n}\sin nz\right). \tag{7.2.40}$$

By means of (7.2.38), (7.2.39) and (7.2.40), the solution of the Laplace equation that satisfies the homogeneous conditions is

$$u(x,y,z) = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} A_{mn}\left[e^{\sqrt{m^2+n^2}\,x} - e^{-\sqrt{m^2+n^2}\,(x-2\pi)}\right]\cos my\left(\cos nz - \frac{k}{n}\sin nz\right).$$

Imposing the inhomogeneous condition in x on this solution,

$$u(0,y,z) = f(y,z) = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} A_{mn}\left(1 - e^{2\pi\sqrt{m^2+n^2}}\right)\cos my\left(\cos nz - \frac{k}{n}\sin nz\right). \tag{7.2.41}$$

Solutions of the form

$$Z_n(z) \sim \left(\cos nz - \frac{k}{n}\sin nz\right)$$

are solutions of the Sturm–Liouville problem

$$\left.\begin{aligned} &Z'' + n^2 Z = 0,\\ &Z'(0) + kZ(0) = 0,\\ &Z'(\pi) + kZ(\pi) = 0, \end{aligned}\right\}$$

i.e. they satisfy a Sturm–Liouville equation under homogeneous boundary conditions. They thus form an orthogonal family of functions on the interval $[0,\pi]$, satisfying

$$\int_0^{\pi} Z_n Z_m\, dz = \begin{cases} 0, & m \neq n,\\ N, & m = n. \end{cases}$$

The constant N is determined as follows. Noting that $Z_n'' = -n^2 Z_n$, integration by parts gives

$$N = \int_0^{\pi} Z_n^2\, dz = -\frac{1}{n^2}\int_0^{\pi} Z_n Z_n''\, dz = -\frac{1}{n^2}\left[ZZ'\right]_0^{\pi} + \frac{1}{n^2}\int_0^{\pi}\left(Z_n'\right)^2 dz. \tag{7.2.42}$$

On the other hand,

$$Z_n = \cos nz - \frac{k}{n}\sin nz, \qquad \frac{1}{n}Z_n' = -\sin nz - \frac{k}{n}\cos nz,$$

so that

$$Z_n^2 + \frac{1}{n^2}\left(Z_n'\right)^2 = \left(1 + \frac{k^2}{n^2}\right). \tag{7.2.43}$$

Integrating this last expression over $[0,\pi]$,

$$\int_0^\pi Z_n^2\,dz = \left(1+\frac{k^2}{n^2}\right)\pi - \frac{1}{n^2}\int_0^\pi \left(Z_n'\right)^2 dz. \tag{7.2.44}$$

Adding (7.2.42) and (7.2.44) gives

$$2\int_0^\pi Z_n^2\,dz = 2N = \left(1+\frac{k^2}{n^2}\right)\pi - \frac{1}{n^2}\left[Z_nZ_n'\right]_0^\pi. \tag{7.2.45}$$

The boundary conditions imply

$$Z'(0) = -kZ(0), \qquad Z'(\pi) = -kZ(\pi). \tag{7.2.46}$$

Writing (7.2.43) once at $z=0$ and once at $z=\pi$ and using (7.2.46),

$$Z_n^2(0) + \frac{1}{n^2}Z_n'^2(0) = Z_n^2(0)\left(1+\frac{k^2}{n^2}\right),$$

$$Z_n^2(\pi) + \frac{1}{n^2}Z_n'^2(\pi) = Z_n^2(\pi)\left(1+\frac{k^2}{n^2}\right).$$

The right sides of these equalities equal $(1+k^2/n^2)$ for every z by (7.2.43), so

$$Z_n^2(0) = Z_n^2(\pi) = 1. \tag{7.2.47}$$

On the other hand, multiplying the first of (7.2.46) by $Z(0)$ and the second by $Z(\pi)$, and using (7.2.47),

$$Z_n(0)Z_n'(0) = -kZ_n^2(0) = -k,$$
$$Z_n(\pi)Z_n'(\pi) = -kZ_n^2(\pi) = -k,$$

so that, from (7.2.45),

$$N = \int_0^\pi Z_n^2\,dz = \frac{\pi}{2}\left(1+\frac{k^2}{n^2}\right).$$

Hence

$$\int_0^\pi Z_nZ_m\,dz = \frac{\pi}{2}\left(1+\frac{k^2}{n^2}\right)\delta_{mn}.$$

With this orthogonality relation, the coefficients in (7.2.41) are

$$A_{mn} = \frac{4}{\pi^2\left(1+\dfrac{k^2}{n^2}\right)\left(1-e^{2\pi\sqrt{m^2+n^2}}\right)}\int_0^\pi\int_0^\pi f(y,z)\cos my\left(\cos nz - \frac{k}{n}\sin nz\right)dy\,dz.$$

Therefore

$$u(x,y,z)=\frac{4}{\pi^2}\sum_{m=0}^{\infty}\sum_{n=1}^{\infty}\frac{1}{\left(1+\dfrac{k^2}{n^2}\right)}\left[\int_0^{\pi}\int_0^{\pi}f(\xi,\eta)\cos m\xi\left(\cos m\eta-\frac{k}{n}\sin m\eta\right)d\xi\,d\eta\right]$$
$$\times\frac{\left[e^{\sqrt{m^2+n^2}\,x}-e^{-\sqrt{m^2+n^2}\,(x-2\pi)}\right]}{\left[1-e^{2\pi\sqrt{m^2+n^2}}\right]}\cos my\left(\cos nz-\frac{k}{n}\sin nz\right).$$

Exercise 7.7 With h and k constants, solve the initial-value problem

$$\left.\begin{aligned}&u_t=ku_{xx}-hu,\quad -\pi\le x\le\pi,\ t>0,\\&u(x,0)=f(x),\\&u(-\pi,t)=u(\pi,t),\\&u_x(-\pi,t)=u_x(\pi,t).\end{aligned}\right\}$$

First simplify the problem by means of the transformation

$$u(x,t)=e^{-ht}v(x,t)$$

and then apply the method of separation of variables.

Solution. With the transformation $u(x,t)=e^{-ht}v(x,t)$ the problem becomes

$$\left.\begin{aligned}&v_t=kv_{xx},\\&v(x,0)=f(x),\\&v(-\pi,t)=v(\pi,t),\\&v_x(-\pi,t)=v_x(\pi,t).\end{aligned}\right\}$$

By the method of separation of variables, the equation separates into the two ODEs

$$X''-\lambda X=0,\qquad T'-k\lambda T=0,$$

with solutions

$$X(x)=Ae^{\sqrt{\lambda}\,x}+Be^{-\sqrt{\lambda}\,x},\qquad T(t)=Ce^{k\lambda t}. \tag{7.2.48}$$

The periodicity conditions

$$X(\pi)=X(-\pi),\qquad X'(\pi)=X'(-\pi)$$

applied to the first of (7.2.48) give, with $\sqrt{\lambda}=im$ ($m=0,1,2,\ldots$),

$$X_m(x)=\alpha_m\cos mx,$$

and accordingly

$$T_m(t)=C_me^{-m^2kt}.$$

Hence the possible solutions of the problem satisfying the periodicity conditions are

$$v_m(x,t) = \bar{\beta}_m e^{-m^2 kt} \cos mx.$$

A superposition of these solutions is again a solution of the equation; the coefficients β_m being determined by the initial condition, the solution of the problem is

$$v(x,t) = \sum_{m=0}^{\infty} \beta_m e^{-km^2 t} \cos mx.$$

The initial condition gives

$$\beta_m = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos mx \, dx.$$

Recalling the function transformation, $u(x,t)$ becomes

$$u(x,t) = e^{-ht} v(x,t) = \frac{1}{\pi} \sum_{m=0}^{\infty} \left\{ \int_{-\pi}^{\pi} f(x) \cos mx \, dx \right\} e^{-(h+km^2)t} \cos mx.$$

Problems

7.1 Let ℓ and h be constants. By the method of separation of variables solve the initial-boundary-value problem

$$\left.\begin{aligned} &u_{tt} = c^2 u_{xx} + h, \quad 0 \le x \le \ell,\ t > 0, \\ &u(x,0) = 0, \quad u_t(x,0) = 0, \\ &u(0,t) = 0, \quad u(\ell,t) = 0. \end{aligned}\right\}$$

7.2 Solve the following problem by the method of separation of variables:

$$\left.\begin{aligned} &u_{tt} = c^2 u_{xx}, \quad 0 \le x \le \ell,\ t > 0, \\ &u(x,0) = x(\ell - x), \quad u_t(x,0) = 0, \\ &u(0,t) = \ell, \quad u(\ell,t) = 0. \end{aligned}\right\}$$

7.3 With A a constant, solve the following initial-value problem by the method of separation of variables:

$$\left.\begin{aligned} &u_{tt} = c^2 u_{xx} + Ax, \quad 0 \le x \le \ell,\ t > 0, \\ &u(x,0) = 0, \quad u_t(x,0) = 0, \\ &u(0,t) = 0, \quad u(\ell,t) = 0. \end{aligned}\right\}$$

7.4 By the method of separation of variables, solve the initial-boundary-value problem

$$\left.\begin{aligned} &u_{tt}+au_t+bu=c^2u_{xx}, \quad 0\le x\le \ell,\ t>0,\\ &u(x,0)=f(x), \quad u_t(x,0)=0,\\ &u(0,t)=u(\ell,t)=0. \end{aligned}\right\}$$

(a,b,c are constants.)

7.5 With h,k,u_0 constants, solve the following initial-value problem by the method of separation of variables:

$$\left.\begin{aligned} &u_t=ku_{xx}+h, \quad 0\le x\le \ell,\ t>0,\\ &u(x,0)=u_0(1-\cos\pi x), \quad 0\le x\le \ell,\\ &u_t(0,t)=0,\\ &u(\ell,t)=2u_0. \end{aligned}\right\}$$

7.6 Solve the following problem by the method of separation of variables:

$$\left.\begin{aligned} &u_t=4u_{xx}, \quad 0\le x\le \ell,\ t>0,\\ &u(x,0)=x^2(1-x), \quad 0\le x\le \ell,\\ &u_t(0,t)=0,\\ &u(\ell,t)=0. \end{aligned}\right\}$$

7.7 Solve the following initial-value problem by the method of separation of variables:

$$\left.\begin{aligned} &u_t=ku_{xx}, \quad 0\le x\le \pi,\ t>0,\\ &u(x,0)=\sin^2 x,\\ &u(0,t)=0,\\ &u(\pi,t)=0. \end{aligned}\right\}$$

7.8 Solve the problems (V.1)–(V.7) above by the finite Fourier transform method, and compare your answers with the results obtained by the method of separation of variables.

7.9 A truncated cone of base radius a and length ℓ, with both ends fastened against rigid surfaces, is initially displaced according to $f(x)$. Setting $h=\ell a/(a-r)$, determine the longitudinal oscillations of the truncated cone by solving the initial-value problem

$$\left.\begin{aligned} &\left(1-\frac{x}{h}\right)^2\frac{\partial^2 u}{\partial t^2}=c^2\frac{\partial}{\partial x}\left[\left(1-\frac{x}{h}\right)^2\frac{\partial u}{\partial x}\right], \quad 0\le x\le \ell,\ t>0,\\ &u(x,0)=f(x), \quad 0\le x\le \ell,\\ &u_t(x,0)=0, \quad 0\le x\le \ell,\\ &u(0,t)=0, \quad t>0,\\ &u(\ell,t)=0, \quad t>0. \end{aligned}\right\}$$

7.10 A thin rectangular plate has two opposite edges insulated, while one of the other two is held at zero temperature and the other at a temperature distribution given by a

function $f(x)$. Determine the steady (time- independent) temperature distribution by solving the boundary-value problem

$$\left.\begin{aligned} &\nabla^2 u(x,y) = 0, \quad 0 \le x \le a,\ 0 \le y \le b, \\ &u(x,0) = f(x), \\ &u(x,b) = 0, \\ &u_x(0,y) = 0, \\ &u_x(a,y) = 0. \end{aligned}\right\}$$

7.11 Solve the following initial-value problems by the method of separation of variables.

(a) $u_t - c^2\nabla^2 u = 0, \quad 0 \le x \le \pi,\ 0 \le y \le \pi,\ t > 0,$

$$\left.\begin{aligned} &u(x,y,0) = 0, \\ &u(0,y,t) = u(\pi,y,t) = 0, \\ &u(x,0,t) = x(x-\pi)\sin t, \\ &u(x,\pi,t) = 0. \end{aligned}\right\}$$

(b) $\nabla^2 u = 0, \quad 0 \le x \le a,\ 0 \le y \le b,\ 0 \le z \le c,$

$$\left.\begin{aligned} &u(0,y,z) = \sin\frac{\pi}{b}y \sin\frac{\pi}{c}z, \\ &u(a,y,z) = 0, \\ &u(x,0,z) = u(x,b,z) = 0, \\ &u(x,y,0) = u(x,y,c) = 0. \end{aligned}\right\}$$

(c) $u_{tt} = c^2\nabla^2 u, \quad 0 \le x \le a,\ 0 \le y \le b,\ t > 0,$

$$\left.\begin{aligned} &u(x,y,0) = f(x,y), \\ &u_t(x,y,0) = g(x,y), \\ &u_x(0,y,t) = u_x(a,y,t) = 0, \\ &u_y(x,0,t) = u_y(x,b,t) = 0. \end{aligned}\right\}$$

Chapter 8
The Finite Fourier Transform

Abstract The finite Fourier transform is one of the standard integral transforms used to solve non-homogeneous linear partial differential equations on bounded one-dimensional spatial intervals. After a review of the Lagrange variation-of-parameters method (which underlies the Green-function representation of the particular solution of a non-homogeneous linear second-order ordinary differential equation), we define the finite Fourier sine and cosine transforms on $(0,\pi)$, the inversion formulas associated with each, and the operational rules needed to convert second-order spatial derivatives into algebraic expressions involving the boundary data. Seven worked exercises apply the technique to the wave equation, the heat equation, a fissionable fuel-rod problem and the two-dimensional wave equation, and ten end-of-chapter problems are provided.

8.1 Lagrange's Method of Variation of Parameters

As we saw in earlier chapters, the partial differential equations of mathematical physics are predominantly of second order, and applying the method of separation of variables to them leads to ordinary differential equations also of second order. Just as for the variables-separation procedure, when integral transforms are used to solve a non-homogeneous partial differential equation, that equation is reduced to an ordinary differential equation which, although having ordinary derivatives, inherits the non-homogeneity. For this reason it is useful to recall here the method of *variation of parameters* (Lagrange's method), a very powerful technique used to obtain the general solution of a non-homogeneous linear ordinary differential equation. Since the equations that interest us are of second order, we shall confine the discussion to the application of the method to second-order linear non-homogeneous equations. Accordingly, suppose we are given the equation

$$y'' + p(x)\,y' + q(x)\,y = g(x), \tag{8.1.1}$$

and let $y_1(x)$ and $y_2(x)$ be two linearly independent particular solutions of the corresponding homogeneous equation. Denoting the integration parameters by C_1 and C_2,

the homogeneous part has the solution

$$y_H = C_1 y_1(x) + C_2 y_2(x). \tag{8.1.2}$$

The *variation of parameters* method consists in replacing, in the expression for $y_H(x)$, the constants C_1 and C_2 with two x-dependent functions $u_1(x)$ and $u_2(x)$, and then writing

$$\widetilde{y}(x) = u_1(x)\, y_1(x) + u_2(x)\, y_2(x), \tag{8.1.3}$$

the unknown functions $u_1(x)$ and $u_2(x)$ being chosen so that $\widetilde{y}$ satisfies the given non-homogeneous differential equation. In this way one obtains a particular solution of the non-homogeneous equation.

Two conditions are needed to determine $u_1(x)$ and $u_2(x)$ explicitly. The first of them is that the function $\widetilde{y}$ expressed in terms of $u_1(x)$ and $u_2(x)$ satisfy the given differential equation. After this condition has been imposed, the second condition is at our free disposal. Naturally, it is most convenient to choose this second condition so that the calculations are made as simple as possible. Since

$$\widetilde{y}(x) = u_1(x)\, y_1(x) + u_2(x)\, y_2(x), \tag{8.1.4}$$

$$\widetilde{y}'(x) = \left(u_1' y_1 + u_2' y_2\right) + \left(u_1 y_1' + u_2 y_2'\right), \tag{8.1.5}$$

let us select the second condition by requiring that u_1 and u_2 satisfy the relation

$$u_1'(x)\, y_1(x) + u_2'(x)\, y_2(x) = 0. \tag{8.1.6}$$

With this condition in force, one finds

$$\widetilde{y}'' = u_1' y_1' + u_2' y_2' + u_1 y_1'' + u_2 y_2''. \tag{8.1.7}$$

On the other hand, since now

$$\widetilde{y}' = u_1 y_1' + u_2 y_2', \tag{8.1.8}$$

substitution of these last two expressions and of $\widetilde{y} = u_1 y_1 + u_2 y_2$ into the given differential equation gives

$$u_1(x)\left[y_1'' + p(x) y_1' + q(x) y_1\right] + u_2(x)\left[y_2'' + p(x) y_2' + q(x) y_2\right] + u_1' y_1' + u_2' y_2' = g(x). \tag{8.1.9}$$

Since y_1 and y_2 are solutions of the homogeneous equation

$$y'' + p(x) y' + q(x) y = 0, \tag{8.1.10}$$

the bracketed terms in the previous expression are zero. Thus

$$u_1'(x)\, y_1'(x) + u_2'(x)\, y_2'(x) = g(x), \tag{8.1.11}$$

must be satisfied. Hence the determination of u_1 and u_2 rests upon the system formed by equations (8.1.6) and (8.1.11),

$$\begin{cases} u_1' y_1 + u_2' y_2 = 0, \\ u_1' y_1' + u_2' y_2' = g(x). \end{cases} \tag{8.1.12}$$

This is an algebraic system in the unknowns u_1' and u_2'. For this system to admit non-trivial solutions, its coefficient determinant

$$W(y_1, y_2) = y_1 y_2' - y_2 y_1' \neq 0 \tag{8.1.13}$$

must be different from zero. This determinant is called the *Wronskian* determinant. By Cramer's rule it follows that

$$u_1'(x) = -\frac{g(x)\, y_2(x)}{W[y_1(x), y_2(x)]}, \qquad u_2'(x) = \frac{g(x)\, y_1(x)}{W[y_1(x), y_2(x)]}, \tag{8.1.14}$$

that is,

$$u_1(x) = -\int_0^x \frac{g(\tau)\, y_2(\tau)}{W[y_1(\tau), y_2(\tau)]}\, d\tau, \qquad u_2(x) = \int_0^x \frac{g(\tau)\, y_1(\tau)}{W[y_1(\tau), y_2(\tau)]}\, d\tau. \tag{8.1.15}$$

Substituting these in the particular solution of the non-homogeneous equation gives

$$\widetilde{y}(x) = \int_0^x \frac{y_2(x)\, y_1(\tau) - y_1(x)\, y_2(\tau)}{W[y_1(\tau), y_2(\tau)]}\, g(\tau)\, d\tau, \tag{8.1.16}$$

or, setting

$$K(x, \tau) \equiv \frac{y_2(x)\, y_1(\tau) - y_1(x)\, y_2(\tau)}{W[y_1(\tau), y_2(\tau)]}, \tag{8.1.17}$$

we obtain

$$\widetilde{y}(x) = \int_0^x K(x, \tau)\, g(\tau)\, d\tau. \tag{8.1.18}$$

The kernel $K(x, \tau)$ is called the *Green's function* of the problem. Accordingly, the general solution of the equation is, as is known, $y(x) = y_H(x) + \widetilde{y}(x)$.

An example. As an illustration, consider the equation

$$y'' + \omega^2 y = f(x). \tag{8.1.19}$$

Here

$$y_H = A \cos \omega x + B \sin \omega x, \tag{8.1.20}$$

and

$$W = \begin{vmatrix} \cos \omega x & \sin \omega x \\ -\omega \sin \omega x & \omega \cos \omega x \end{vmatrix} = \omega, \tag{8.1.21}$$

so that

$$\begin{aligned}\widetilde{y} &= \frac{1}{\omega}\int_0^x \big[\sin\omega x\cos\omega\tau - \cos\omega x\sin\omega\tau\big] f(\tau)\,d\tau \\ &= \frac{1}{\omega}\int_0^x f(\tau)\,\sin\omega(x-\tau)\,d\tau, \end{aligned} \tag{8.1.22}$$

and consequently the general solution of the given equation is

$$y(x) = y_H(x) + \widetilde{y}(x) = A\cos\omega x + B\sin\omega x + \frac{1}{\omega}\int_0^x f(\tau)\,\sin\omega(x-\tau)\,d\tau. \tag{8.1.23}$$

8.2 The Finite Fourier Transform Method

The finite Fourier transform method is a frequently used technique for finding solutions of non-homogeneous linear partial differential equations. Let $f = f(x)$ be a piecewise-continuous function defined on the interval $[0,\pi]$. Such an interval is taken merely to ensure boundary regularity; changing the length of the interval poses no difficulty whatsoever. The *finite Fourier sine transform* of $f(x)$ is then defined by

$$F_s[f(x)] = F_s(n) = \frac{2}{\pi}\int_0^\pi f(x)\,\sin nx\,dx \qquad (n = 1,2,\dots). \tag{8.2.1}$$

On the other hand, if we consider the Fourier sine expansion of $f(x)$ on $[0,\pi]$ (cf. (1.3.3) with $\ell = \pi$),

$$f(x) = \sum_{n=1}^{\infty} a_n \sin nx, \tag{8.2.2}$$

then since the expansion coefficients are

$$a_n = \frac{2}{\pi}\int_0^\pi f(x)\,\sin nx\,dx \qquad (n = 1,2,\dots), \tag{8.2.3}$$

comparison of (8.2.1) and (8.2.3) shows that (8.2.2) is

$$f(x) = \sum_{n=1}^{\infty} F_s(n)\,\sin nx, \tag{8.2.4}$$

which gives the inverse-transform formula that, once $F_s(n)$ is known, recovers $f(x)$.

The *finite Fourier cosine transform* of $f(x)$ is in turn defined by

$$F_c[f(x)] = F_c(n) = \frac{2}{\pi}\int_0^\pi f(x)\,\cos nx\,dx \qquad (n = 0,1,2,\dots), \tag{8.2.5}$$

with the inverse transform

$$f(x) = \frac{F_c(0)}{2} + \sum_{n=1}^{\infty} F_c(n)\,\cos nx. \tag{8.2.6}$$

Since we are particularly concerned with second-order equations, let us compute the finite Fourier sine transform of $f''(x)$:

$$\begin{aligned} F_s\big[f''(x)\big] &= \frac{2}{\pi}\int_0^\pi f''(x)\,\sin nx\,dx = \frac{2}{\pi}\big[f'(x)\sin nx\big]_0^\pi - \frac{2n}{\pi}\int_0^\pi f'(x)\,\cos nx\,dx \\ &= -\frac{2n}{\pi}\big[f(x)\cos nx\big]_0^\pi - \frac{2n^2}{\pi}\int_0^\pi f(x)\,\sin nx\,dx \\ &= -\frac{2n}{\pi}\big[(-1)^n f(\pi) - f(0)\big] - n^2 F_s(n). \end{aligned} \tag{8.2.7}$$

The finite Fourier cosine transform of $f''(x)$ is computed in the same way, and one finds

$$F_c\big[f''(x)\big] = \frac{2}{\pi}\big[(-1)^n f'(\pi) - f'(0)\big] - n^2 F_c(n). \tag{8.2.8}$$

We now give a few examples to see at close range how the finite Fourier transform is applied to the solution of partial differential equations.

Worked Exercises

Exercise VI.1

Suppose a force acts on an elastic string of length π which is fixed at both ends. Then with the prescribed boundary and initial conditions the motion of the string is determined by

$$\begin{cases} u_{tt} = c^2 u_{xx} + f(x,t), \quad 0 \le x \le \pi, \quad t > 0, \\ u(x,0) = 0, \quad u_t(x,0) = 0, \\ u(0,t) = 0, \quad u(\pi,t) = 0. \end{cases} \tag{8.2.9}$$

Apply the finite Fourier sine transform with respect to the spatial variable x:

$$F_s\big[u_{tt} - c^2 u_{xx} - f(x,t)\big] = 0, \tag{8.2.10}$$

and since the operator representing the finite Fourier sine transform is linear,

$$F_s[u_{tt}] - c^2 F_s[u_{xx}] = F_s[f(x,t)]\,. \tag{8.2.11}$$

Let

$$U_s(n,t) \equiv \frac{2}{\pi}\int_0^\pi u(x,t)\,\sin nx\,dx. \tag{8.2.12}$$

Then

$$F_s[u_{tt}] = \frac{2}{\pi}\int_0^\pi u_{tt}\,\sin nx\,dx = \frac{\partial^2}{\partial t^2}\left[\frac{2}{\pi}\int_0^\pi u(x,t)\,\sin nx\,dx\right] = \frac{d^2 U_s(n,t)}{dt^2}. \tag{8.2.13}$$

On the other hand, integrating by parts and using the boundary conditions,

$$F_s[u_{xx}] = \frac{2n}{\pi}\left[u(0,t) - (-1)^n u(\pi,t)\right] - n^2 U_s(n,t) = -n^2 U_s(n,t). \tag{8.2.14}$$

Also denote the finite Fourier sine transform of $f(x)$ by

$$F_s(n,t) = \frac{2}{\pi}\int_0^\pi f(x,t)\,\sin nx\,dx. \tag{8.2.15}$$

Then by means of (8.2.11), (8.2.13), (8.2.14) and (8.2.15) the transformed equation becomes

$$\frac{d^2U_s(n,t)}{dt^2} + n^2c^2\,U_s(n,t) = F_s(n,t). \tag{8.2.16}$$

We know that the general solution of this non-homogeneous second-order ordinary differential equation is

$$U_s(n,t) = A\cos nct + B\sin nct + \frac{1}{nc}\int_0^t F_s(n,\tau)\,\sin nc(t-\tau)\,d\tau. \tag{8.2.17}$$

Now from the initial conditions

$$F_s[u(x,0)] = \frac{2}{\pi}\int_0^\pi u(x,0)\,\sin nx\,dx = U_s(n,0) = 0, \tag{8.2.18}$$

and

$$F_s[u_t(x,0)] = \left(\frac{dU}{dt}\right)_{t=0} = 0, \tag{8.2.19}$$

so that

$$U(n,0) = A = 0, \qquad U_t(n,0) = B = 0, \tag{8.2.20}$$

that is,

$$U_s(n,t) = \frac{1}{nc}\int_0^t F_s(n,\tau)\,\sin nc(t-\tau)\,d\tau. \tag{8.2.21}$$

Hence the inverse transform of $U_s(n,t)$ is

$$u(x,t) = \sum_{n=1}^{\infty} U_s(n,t)\,\sin nx = \sum_{n=1}^{\infty}\left\{\frac{1}{nc}\int_0^t F(n,\tau)\,\sin nc(t-\tau)\,d\tau\right\}\sin nx. \tag{8.2.22}$$

As a special case let $f(x,t) = a =$ constant. Then

$$F_s[a] = \frac{2}{\pi}\int_0^\pi a\,\sin nx\,dx = \frac{2a}{n\pi}\left[1-(-1)^n\right], \tag{8.2.23}$$

so that $U_s(n,t)$ becomes

$$U_s[n,t] = \frac{1}{nc}\int_0^t \frac{2a}{n\pi}\left[1-(-1)^n\right]\sin nc(t-\tau)\,d\tau$$
$$= \frac{2a}{n^3\pi c^2}\left[1-(-1)^n\right]\left(1-\cos nct\right), \tag{8.2.24}$$

and consequently

$$u(x,t) = \frac{2a}{\pi c^2}\sum_{n=1}^{\infty}\frac{\left[1-(-1)^n\right]}{n^3}\left(1-\cos nct\right)\sin nx. \tag{8.2.25}$$

Exercise VI.2

Determine the temperature distribution in a uniform rod of length π which generates an internal heat source $g(x,t)$ per unit time. The ends of the rod are insulated, and the initial temperature distribution at $t=0$ is assumed to be represented by a function $f(x)$.

"Solution:" The problem is the initial-boundary value problem

$$\begin{cases} u_t = u_{xx} + g(x,t), \quad 0 < x < \pi, \quad t > 0, \\ u(x,0) = f(x), \quad 0 \le x \le \pi, \\ u_x(0,t) = 0, \quad u_x(\pi,t) = 0. \end{cases} \tag{8.2.26}$$

Apply the finite Fourier cosine transform

$$U_c(n,t) = \frac{2}{\pi}\int_0^{\pi} u(x,t)\cos nx\,dx \tag{8.2.27}$$

to the given heat-conduction equation. With

$$G(n,t) = \frac{2}{\pi}\int_0^{\pi} g(x,t)\cos nx\,dx, \tag{8.2.28}$$

one finds

$$\frac{dU_c}{dt} = \frac{2}{\pi}\left[(-1)^n u_x(\pi,t) - u_x(0,t)\right] - n^2 U_c(n,t) + G(n,t)$$
$$= -n^2 U_c(n,t) + G(n,t), \tag{8.2.29}$$

or

$$\frac{dU}{dt} + n^2 U = G(n,t). \tag{8.2.30}$$

Bearing in mind that this equation can be written in the form

$$\frac{d}{dt}\left(e^{n^2 t}U_c\right) = e^{n^2 t}G(n,t), \tag{8.2.31}$$

denoting an integration constant by A, one obtains

$$e^{n^2t}U_c(n,t) = \int_0^t e^{n^2\tau}G(n,\tau)\,d\tau + A, \tag{8.2.32}$$

$$U_c(n,t) = \int_0^t e^{-n^2(t-\tau)}G(n,\tau)\,d\tau + A\,e^{-n^2t}. \tag{8.2.33}$$

For the transform of the initial condition,

$$U_c(n,0) = \frac{2}{\pi}\int_0^\pi u(x,0)\cos nx\,dx = \frac{2}{\pi}\int_0^\pi f(x)\cos nx\,dx. \tag{8.2.34}$$

Hence

$$U_c(n,t) = \int_0^t e^{-n^2(t-\tau)}G(n,\tau)\,d\tau + \frac{2\,e^{-n^2t}}{\pi}\int_0^\pi f(\xi)\cos n\xi\,d\xi. \tag{8.2.35}$$

Hence by the inverse-transform formula one finds for $u(x,t)$ that

$$\begin{aligned} u(x,t) &= \frac{U(0,t)}{2} + \sum_{n=1}^{\infty} U(n,t)\cos nx \\ &= \frac{1}{\pi}\int_0^\pi \left[f(\xi) + \int_0^t g(\xi,\tau)\,d\tau \right] d\xi \\ &\quad + \sum_{n=1}^{\infty} \left\{ \int_0^t e^{-n^2(t-\tau)}G(n,\tau)\,d\tau + \frac{2\,e^{-n^2t}}{\pi}\int_0^\pi f(\xi)\cos n\xi\,d\xi \right\} \cos nx. \end{aligned} \tag{8.2.36}$$

Exercise VI.3

Let a and b be physical constants and consider a rod of length π containing a fissionable fuel which produces neutrons. Assuming the two ends of the rod are ideal reflectors which prevent neutrons from escaping, and that the neutron distribution at the initial instant is known to be $f(x)$, determine the neutron distribution at any subsequent instant.

Solution: According to the data, the problem is

$$\begin{cases} u_t = a\,u_{xx} + b\,u, \quad 0 < x < \pi, \quad t > 0, \\ u(x,0) = f(x), \\ u_x(0,t) = 0, \quad u_x(\pi,t) = 0. \end{cases} \tag{8.2.37}$$

We apply termwise to this equation the transform

$$U_c(n,t) = \frac{2}{\pi}\int_0^\pi u(x,t)\cos nx\,dx. \tag{8.2.38}$$

Since

$$F_c[u_{xx}] = \frac{2}{\pi}\left[(-1)^n u_x(\pi,t) - u_x(0,t)\right] - n^2 U(n,t) = -n^2 U(n,t), \tag{8.2.39}$$

the transformed equation is

$$\frac{dU_c}{dt} + \left(an^2 - b\right)U_c = 0. \tag{8.2.40}$$

Applying the initial condition to this equation's solution

$$U_c(n,t) = A\, e^{-(an^2-b)t} = U_c(n,0)\, e^{-(an^2-b)t} \tag{8.2.41}$$

in the form

$$U_c(n,0) = \frac{2}{\pi}\int_0^\pi u(x,0)\cos nx\, dx = \frac{2}{\pi}\int_0^\pi f(x)\cos nx\, dx, \tag{8.2.42}$$

since

$$A \equiv \frac{2}{\pi}\int_0^\pi f(x)\cos nx\, dx, \tag{8.2.43}$$

the solution is found in the form

$$u(x,t) = \frac{U_c(0,t)}{2} + \sum_{n=1}^{\infty} U_c(n,t)\cos nx. \tag{8.2.44}$$

Here

$$U_c(n,t) = U_c(n,0)\, e^{-(an^2-b)t}, \quad U_c(0,t) = U_c(0,0)\, e^{bt}, \quad \frac{U_c(0,t)}{2} = \frac{U_c(0,0)}{2}\, e^{bt}. \tag{8.2.45}$$

If, as a special case, we take $f(x) = x$ for $0 < x < \pi$, then

$$U_c(0,0) = \frac{2}{\pi}\int_0^\pi x\, dx = \pi, \tag{8.2.46}$$

and

$$U_c(n,0) = \frac{2}{n^2\pi}\left[(-1)^n - 1\right], \qquad (n = 1,2,3,\dots), \tag{8.2.47}$$

and consequently

$$u(x,t) = \frac{\pi\, e^{bt}}{2} + \sum_{n=1}^{\infty}\frac{2}{n^2\pi}\left[(-1)^n - 1\right]e^{-(an^2-b)t}\cos nx. \tag{8.2.48}$$

Exercise VI.4

Solve, by the finite Fourier transform method, the boundary-value problem

$$\begin{cases} u_{tt} = c^2 u_{xx} + f(x,t), \quad 0 \le x \le \pi, \quad t > 0, \\ u(x,0) = 0, \\ u_t(x,0) = 0, \\ u(0,t) = u(\pi,t) = 0. \end{cases} \tag{8.2.49}$$

Solution: Since the problem is a Dirichlet problem, the transform to consider is

$$U_s(n,t) = \frac{2}{\pi}\int_0^\pi u(x,t)\,\sin nx\,dx. \tag{8.2.50}$$

Then

$$F_s[u_{tt}] = \frac{2}{\pi}\frac{\partial^2}{\partial t^2}\int_0^\pi u(x,t)\,\sin nx\,dx = \frac{d^2U_s}{dt^2}, \tag{8.2.51}$$

$$\begin{aligned} F_s[u_{xx}] &= \frac{2}{\pi}\int_0^\pi u_{xx}\sin nx\,dx = \frac{2n}{\pi}\left[U(\pi,t)(-1)^n - U(0,t)\right] - n^2 U_s(n,t) \\ &= -n^2 U_s(n,t), \end{aligned} \tag{8.2.52}$$

$$F_s[f(x,t)] = \frac{2}{\pi}\int_0^\pi f(x,t)\,\sin nx\,dx = F_s(n,t). \tag{8.2.53}$$

Hence the transformed equation

$$\frac{d^2U_s(n,t)}{dt^2} + c^2n^2\,U_s(n,t) = F_s(n,t). \tag{8.2.54}$$

For the solution of this equation

$$U_s(n,t) = A\cos nct + B\sin nct + \frac{1}{cn}\int_0^t F_s(n,\tau)\,\sin cn(t-\tau)\,d\tau, \tag{8.2.55}$$

applying the transformed initial conditions

$$U_s(n,0) = A = 0, \qquad \left(\frac{dU_s}{dt}\right)_{t=0} = cnB = 0, \tag{8.2.56}$$

the transform of the problem's solution is

$$U_s(n,t) = \frac{1}{cn}\int_0^t F_s(n,\tau)\,\sin cn(t-\tau)\,d\tau. \tag{8.2.57}$$

Hence the solution of the problem is

$$u(x,t) = \sum_{n=0}^{\infty}\frac{1}{cn}\left(\int_0^t F_s(n,\tau)\,\sin cn(t-\tau)\,d\tau\right)\sin nx. \tag{8.2.58}$$

Exercise VI.5

Solve by the finite Fourier transform method the boundary-value problem

$$\begin{cases} u_t = u_{xx} + 2x, \quad 0 \le x \le 1, \quad t > 0, \\ u(x,0) = \sin\dfrac{3\pi}{2}x, \\ u(0,t) = u(1,t) = 0. \end{cases} \tag{8.2.59}$$

Solution: Since the problem is a Dirichlet problem, the transform to consider is

$$F_s[u(x,t)] = 2\int_0^1 u(x,t)\sin n\pi x\,dx = U_s(n,t). \tag{8.2.60}$$

Then

$$F_s[u_{xx}] = 2\int_0^1 u_{xx}\sin n\pi x\,dx = -(n\pi)^2 U_s(n,t), \tag{8.2.61}$$

$$F_s[x] = 2\int_0^1 x\sin n\pi x\,dx = -2\frac{(-1)^n}{n\pi}, \tag{8.2.62}$$

so that the transformed equation is

$$\frac{dU_s}{dt} + (n\pi)^2 U_s = \frac{4}{n\pi}(-1)^{n+1}. \tag{8.2.63}$$

For the solution of this equation

$$U_s(n,t) = C_n e^{-n^2\pi^2 t} + \frac{4}{(n\pi)^3}(-1)^{n+1}, \tag{8.2.64}$$

applying the transformed initial condition

$$U_s(n,0) = F_s[u(x,0)] = F_s\left[\sin\frac{3\pi}{2}x\right] = 2\int_0^1 \sin\frac{3\pi}{2}x\sin n\pi x\,dx = 0, \tag{8.2.65}$$

yields

$$C_n = \frac{4}{(n\pi)^3}(-1)^n, \tag{8.2.66}$$

and the transform of the problem's solution is

$$U_s(n,t) = \frac{4}{(n\pi)^3}(-1)^{n+1}\left(1 - e^{-n^2\pi^2 t}\right). \tag{8.2.67}$$

Hence the solution of the problem is

$$u(x,t) = \frac{2}{\pi^3}\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n^3}\left(1 - e^{-n^2\pi^2 t}\right)\sin n\pi x. \tag{8.2.68}$$

Exercise VI.6

Solve by the finite Fourier transform method the initial-value problem

$$\begin{cases} u_{tt} = c^2\nabla^2 u, \quad 0 \le x \le 1, \quad 0 < y < 1, \quad t > 0, \\ u(x,y;0) = 0, \\ u_t(x,y;0) = \sin\pi x \sin 2\pi y, \\ u(0,y;t) = u(1,y,t) = 0, \\ u(x,0;t) = u(x,1;t) = 0. \end{cases} \tag{8.2.69}$$

Solution: Since the problem is a Dirichlet problem and possesses two spatial variables, the transform to consider is

$$U_s(m,n;t) = 4\int_0^1\int_0^1 u(x,y;t)\,\sin m\pi x \sin n\pi y\,dx\,dy. \tag{8.2.70}$$

Then within the framework of the boundary conditions

$$F_s[\nabla^2 u] = -(m^2+n^2)\pi^2 U_s(m,n;t), \tag{8.2.71}$$

so that the transformed equation, with $\alpha_{mn} = \sqrt{m^2+n^2}\,c\pi$, is

$$\frac{d^2U_s}{dt^2} + \alpha_{mn}^2 U_s = 0. \tag{8.2.72}$$

For the solution of this equation

$$U_s(m,n;t) = A_{mn}\cos\alpha_{mn}t + B_{mn}\sin\alpha_{mn}t, \tag{8.2.73}$$

$$U_s(m,n;0) = 0, \tag{8.2.74}$$

$$\begin{aligned} \left[\frac{dU_s}{dt}\right]_{t=0} &= 4\int_0^1\int_0^1 \sin\pi x \sin m\pi x \sin n\pi y \sin 2\pi y\,dy\,dx \\ &= \begin{cases} 1 & (\text{for } m=1,\ n=2), \\ 0 & (\text{for } m\ne 1,\ n\ne 2), \end{cases} \end{aligned} \tag{8.2.75}$$

applying the transformed initial conditions

$$A_{mn} = 0, \qquad U_s = B_{mn}\sin\alpha_{mn}t, \tag{8.2.76}$$

$$\left(\frac{dU_s}{dt}\right)_{t=0} = \alpha_{mn}B_{mn}\cos\alpha_{mn}t\big|_{t=0} = \begin{cases} 1 & m=1,\ n=2, \\ 0 & m\ne 1,\ n\ne 2, \end{cases} \tag{8.2.77}$$

$$B_{12} = \frac{1}{\alpha_{12}} = \frac{1}{\sqrt{5}\,c\pi}. \tag{8.2.78}$$

Thus the transform of the problem's solution is

$$U_s(m,n,t) = \frac{1}{\sqrt{5}\,c\pi} \sin\sqrt{5}\,c\pi t \qquad (m=1,\ n=2), \tag{8.2.79}$$

so that the problem's solution is

$$u(x,y;t) = \frac{1}{\sqrt{5}\,c\pi} \sin\sqrt{5}\,c\pi t \sin\pi x \sin 2\pi y. \tag{8.2.80}$$

Exercise VI.7

A membrane initially at rest occupies the region $0 \le x \le a,\ 0 \le y \le b$ in the (x,y)-plane, fixed along its edges. Determine the oscillations of the membrane given that the initial displacement is $f(x,y)$ and the initial velocity at each point is $g(x,y)$. (Special case: for $t=0$ let $u(x,y;0)=0$ and $\left(\partial u/\partial t\right)_{t=0} = A\,\delta(x-x_0)\,\delta(y-y_0)$; that is, consider that an initially stationary membrane is at the point (x_0,y_0) given a momentum corresponding to a velocity of magnitude A.)

Solution: The initial-value problem to be solved for the two-dimensional wave equation is

$$\begin{cases} \dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial^2 u}{\partial y^2} = \dfrac{1}{c^2}\dfrac{\partial^2 u}{\partial t^2}, \quad 0 \le x \le a, \quad 0 \le y \le b, \quad t > 0, \\ u(x,y;0) = f(x,y), \\ u_t(x,y;0) = g(x,y), \\ u(0,y;t) = u(a,y;t) = 0, \\ u(x,0;t) = u(x,b;t) = 0. \end{cases} \tag{8.2.81}$$

Since this is a Dirichlet problem, we consider the transform

$$U_s(m,n;t) = \frac{4}{ab}\int_0^a \int_0^b u(x,y;t) \sin\frac{m\pi}{a}x \sin\frac{n\pi}{b}y\,dx\,dy. \tag{8.2.82}$$

Then

$$F_s[\nabla^2 u] = F_s\left[\frac{\partial^2 u}{\partial x^2}\right] + F_s\left[\frac{\partial^2 u}{\partial y^2}\right], \tag{8.2.83}$$

and by integration by parts

$$F_s\left[\frac{\partial^2 u}{\partial x^2}\right] = \frac{4}{ab}\int_0^a\int_0^b \frac{\partial^2 u}{\partial x^2}\sin\frac{m\pi}{a}x\sin\frac{n\pi}{a}y\,dx\,dy$$
$$= -\left(\frac{m\pi}{a}\right)^2\frac{4}{ab}\int_0^a\int_0^b u(x,y;t)\sin\frac{m\pi}{a}x\sin\frac{n\pi}{a}y\,dx\,dy$$
$$= -\left(\frac{m\pi}{a}\right)^2 U_s(m,n;t), \tag{8.2.84}$$

and similarly

$$F_s\left[\frac{\partial^2 u}{\partial y^2}\right] = -\left(\frac{n\pi}{b}\right)^2 U_s(m,n;t), \tag{8.2.85}$$

so that

$$F_s[\nabla^2 u] = -\left(\frac{m^2}{a^2}+\frac{n^2}{b^2}\right)\pi^2 U_s(m,n;t), \tag{8.2.86}$$

and

$$F_s\left[\frac{\partial^2 u}{\partial t^2}\right] = \frac{d^2U_s(m,n;t)}{dt^2}. \tag{8.2.87}$$

Thus the transformed equation, with $\omega_{mn} \equiv \left(m^2/a^2+n^2/b^2\right)^{1/2}c\pi$, is

$$\frac{d^2U_s}{dt^2}+\omega_{mn}^2c^2\pi^2U_s = 0. \tag{8.2.88}$$

For the solution of this equation

$$U_s(m,n;t) = A(m,n)\cos\omega_{mn}t + B(m,n)\sin\omega_{mn}t, \tag{8.2.89}$$

applying the transformed initial conditions:

$$U_s(m,n;0) = \frac{4}{ab}\int_0^a\int_0^b u(x,y;0)\sin\frac{m\pi}{a}x\sin\frac{n\pi}{b}y\,dx\,dy$$
$$= \frac{4}{ab}\int_0^a\int_0^b f(x,y)\sin\frac{m\pi}{a}x\sin\frac{n\pi}{b}y\,dx\,dy = F_s(m,n) = A(m,n), \tag{8.2.90}$$

$$\left(\frac{dU_s}{dt}\right)_{t=0} = \left[\frac{4}{ab}\int_0^a\int_0^b \frac{\partial u}{\partial t}\sin\frac{m\pi}{a}x\sin\frac{n\pi}{b}y\,dx\,dy\right]_{t=0}$$
$$= \frac{4}{ab}\int_0^a\int_0^b g(x,y)\sin\frac{m\pi}{a}x\sin\frac{n\pi}{b}y\,dx\,dy = G_s(m,n) = \omega_{mn}B(m,n). \tag{8.2.91}$$

Hence the transform of the problem's solution is

$$U_s(m,n;t) = F_s(m,n)\cos\omega_{mn}t + \frac{1}{\omega_{mn}}G_s(m,n)\sin\omega_{mn}t. \tag{8.2.92}$$

Thus the problem's solution is

$$u(x,y,t) = \frac{4}{ab}\sum_m\sum_n\left[\int_0^a\int_0^b f(\xi,\eta)\sin\frac{m\pi}{a}\xi\,\sin\frac{n\pi}{b}\eta\,d\xi\,d\eta\right]\sin\frac{m\pi}{a}x\sin\frac{n\pi}{b}y\cos\omega_{mn}t$$
$$+\frac{4}{ab}\sum_m\sum_n\frac{1}{\omega_{mn}}\left[\int_0^a\int_0^b g(\xi,\eta)\sin\frac{m\pi}{a}\xi\,\sin\frac{n\pi}{b}\eta\,d\xi\,d\eta\right]\sin\frac{m\pi}{a}x\sin\frac{n\pi}{b}y\sin\omega_{mn}t. \tag{8.2.93}$$

Special case. If $u(x,y,0)=0$ and $\left(\partial u/\partial t\right)_{t=0} = A\,\delta(x-x_0)\,\delta(y-y_0)$ are taken, then

$$A(m,n) = F_s(m,n) = 0, \tag{8.2.94}$$

$$\omega_{mn}B(m,n) = G(m,n) = \frac{4}{ab}\int_0^a\int_0^b A\,\delta(x-x_0)\,\delta(y-y_0)\,\sin\frac{m\pi}{a}x\sin\frac{n\pi}{b}y\,dx\,dy$$
$$= \frac{4A}{ab}\sin\frac{m\pi}{a}x_0\sin\frac{n\pi}{b}y_0, \tag{8.2.95}$$

and

$$U_s(m,n;t) = \frac{4A}{ab}\,\frac{\sin\dfrac{m\pi}{a}x_0\sin\dfrac{n\pi}{b}y_0}{\sqrt{\dfrac{m^2}{a^2}+\dfrac{n^2}{b^2}}\,\pi c}\,\sin\sqrt{\frac{m^2}{a^2}+\frac{n^2}{b^2}}\,c\pi t. \tag{8.2.96}$$

Hence the problem's solution is

$$u(x,y;t) = \frac{4A}{\pi c}\sum_m\sum_n\frac{1}{\sqrt{m^2b^2+n^2a^2}}\,\sin\sqrt{\frac{m^2}{a^2}+\frac{n^2}{b^2}}\,c\pi t\,\sin\frac{m\pi}{a}x_0\sin\frac{m\pi}{a}x\sin\frac{n\pi}{b}y_0\sin\frac{n\pi}{b}y. \tag{8.2.97}$$

Problems

Solve the following problems by the finite Fourier transform method.

(VI.1) $u_t = a\,u_{xx} + b\,u, \quad 0 \le x \le \pi,\ t > 0 \quad (a \text{ and } b \text{ constants})$

$$u(x,0) = f(x),$$
$$u_x(0,t) = u_x(\pi,t) = 0.$$

(VI.2) $u_t = u_{xx} + g(x,t), \quad 0 \le x \le \pi,\ t > 0$

$$u(x,0) = f(x),$$
$$u(0,t) = 0,$$
$$u(\pi,t) = 0.$$

(VI.3) $u_{tt} = u_{xx}, \quad 0 \le x \le \pi,\ t > 0$

$$\begin{aligned} u(x,0) &= 0, \\ u_t(x,0) &= 0, \\ u_x(0,t) &= 0, \\ u_x(\pi,t) &= -b \quad (b \text{ a constant}). \end{aligned}$$

(VI.4) $u_{tt} + c^2 u_{xxxx} = 0, \quad 0 \le x \le \pi,\ t > 0$

$$\begin{aligned} u(x,0) &= 0, \\ u_t(x,0) &= 0, \\ u(0,t) &= u(\pi,t) = 0, \\ u_{xx}(0,t) &= 0, \\ u_{xx}(\pi,t) &= b = \text{constant}. \end{aligned}$$

(VI.5) $\dfrac{1}{c^2}\dfrac{\partial^2 u}{\partial t^2} = \dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial^2 u}{\partial y^2}, \quad 0 \le x \le a,\ 0 \le y \le b,\ t > 0$

$$\begin{aligned} u(x,y,0) &= f(x,y), \\ u_t(x,y,0) &= g(x,y), \\ u(0,y,t) &= u(a,y,t) = 0, \\ u(x,0,t) &= u(x,b,t) = 0. \end{aligned}$$

(VI.6) $\dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial^2 u}{\partial y^2} = \dfrac{1}{c^2}\dfrac{\partial u}{\partial t}, \quad 0 \le x \le a,\ 0 \le y \le b,\ t > 0$

$$\begin{aligned} u(x,y,0) &= f(x,y), \\ u(0,y,t) &= u(a,y,t) = 0, \\ u(x,0,t) &= u(x,b,t) = 0. \end{aligned}$$

(VI.7) $\dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial^2 u}{\partial y^2} = \dfrac{1}{c^2}\dfrac{\partial u}{\partial t}, \quad 0 \le x \le a,\ 0 \le y \le b,\ t > 0$

$$\begin{aligned} u(x,y,0) &= f(x,y), \\ u(0,y,t) &= u(a,y,t) = 0, \\ u(x,b,t) &= 0, \\ u(x,0,t) &= f(x). \end{aligned}$$

(VI.8) $u_t = c^2 \nabla^2 u, \quad 0 \le x \le \pi,\ 0 \le y \le \pi,\ t > 0$

$$\begin{aligned} u(x,y,0) &= 0, \\ u(0,y,t) &= u(\pi,y,t) = 0, \\ u(x,0,t) &= x(x-\pi)\sin t, \\ u(x,\pi,t) &= 0. \end{aligned}$$

(VI.9) Solve the following problems by the finite Fourier transform method.

(a) $u_{tt} = c^2\nabla^2 u + F(x,y,t), \quad 0 \le x \le a,\ 0 \le y \le b,\ t > 0$

$$\begin{aligned} u(x,y,0) &= f(x,y), \\ u_t(x,y,0) &= g(x,y), \\ u_x(0,y,t) &= u_x(a,y,t) = 0, \\ u_y(x,0,t) &= u_y(x,b,t) = 0. \end{aligned}$$

(b) $u_{tt} = c^2\nabla^2 u + xy\sin t, \quad 0 \le x \le \pi,\ 0 \le y \le \pi,\ t > 0$

$$\begin{aligned} u(x,y,0) &= 0, \\ u_t(x,y,0) &= 0, \\ u(0,y,t) &= u(\pi,y,t) = 0, \\ u(x,0,t) &= u(x,\pi,t) = 0. \end{aligned}$$

(c) $u_t = k\nabla^2 u + g(x,y), \quad 0 \le x \le a,\ 0 \le y \le b,\ t > 0$

$$\begin{aligned} u(x,y,0) &= f(x,y), \\ u_x(0,y,t) &= u_x(a,y,t) = 0, \\ u(x,0,t) &= u(x,b,t) = 0. \end{aligned}$$

(VI.10) Solve the following problems both by the finite Fourier transform method and by the method of separation of variables, and compare the results.

(a) $\nabla^2 u = 0, \quad 0 \le x \le 1,\ 0 \le y \le 1,\ 0 \le z \le 1$

$$\begin{aligned} u_x(0,y,z) &= u_x(1,y,z) = 0, \\ u_y(x,0,z) &= u_y(x,1,z) = 0, \\ u_z(x,y,0) &= \cos\pi x \cos\pi y, \\ u_z(x,y,1) &= 0. \end{aligned}$$

(b) $u_{tt} = c^2\nabla^2 u, \quad 0 \le x \le a,\ 0 \le y \le b,\ 0 \le z \le c,\ t > 0$

$$\begin{aligned} u(x,y,z,0) &= f(x,y,z), \\ u_t(x,y,z,0) &= g(x,y,z), \\ u(0,y,z,t) &= u(a,y,z,t) = 0, \\ u(x,0,z,t) &= u(x,b,z,t) = 0, \\ u(x,y,0,t) &= u(x,y,c,t) = 0. \end{aligned}$$

Chapter 9
The Sturm–Liouville Problem

Abstract We develop the Sturm–Liouville theory: separation of the master equation in Cartesian, cylindrical and spherical coordinates; the regular Sturm–Liouville problem with real eigenvalues and eigenfunctions orthogonal with respect to $R(x)$; the variational (Euler–Lagrange) characterisation of the eigenvalues; their asymptotic behaviour through the Liouville transformation; completeness; the orthogonality relations of the Bessel functions J_n and the Legendre polynomials P_n, with application to disk heat conduction, the heat equation in a finite cylinder, the vibrating membrane, the steady annular plate, and the Laplace problem outside and inside a sphere.

9.1 The General Equations of Mathematical Physics and Their Separation

Most of the partial differential equations of theoretical physics may be written either as the equation

$$\nabla^2\psi(\mathbf{r},t)+\lambda'\psi(\mathbf{r},t)=A(\mathbf{r},t)+B(t)\,\frac{\partial\psi(\mathbf{r},t)}{\partial t}+C(t)\,\frac{\partial^2\psi(\mathbf{r},t)}{\partial t^2} \tag{9.1.1}$$

or as one of its specialisations:

1. $A\equiv 0$, $B=$ const., $C=$ const.: *telegraphers' equation* (electrodynamics);
2. $\lambda'\equiv B\equiv C\equiv 0$: *Poisson's equation* (classical theoretical mechanics, electrodynamics);
3. $\lambda'\equiv A\equiv B\equiv 0$, $C=$ const.: *wave equation* (continuum mechanics, electrodynamics);
4. $A\equiv B\equiv C\equiv 0$: *Helmholtz equation* (classical theoretical mechanics, quantum mechanics, neutron diffusion theory);
5. $A\equiv C\equiv 0$, $B=$ const.: *diffusion equation* (theory of heat, kinetic theory of gases);
6. $\lambda'\equiv A\equiv B\equiv C\equiv 0$: *Laplace equation* (classical gravitation theory, electrostatics).

These special cases were discussed in the Introduction. For $A \equiv 0$ the master equation (9.1.1) reduces to the homogeneous equation

$$\nabla^2\psi + \lambda'\psi - B\frac{\partial\psi}{\partial t} - C\frac{\partial^2\psi}{\partial t^2} = 0. \tag{9.1.2}$$

We solve (9.1.2) by separation of variables, looking for a solution of the form

$$\psi(\mathbf{r},t) = U(\mathbf{r})\,T(t), \tag{9.1.3}$$

that is, the product of a function of $\mathbf{r}$ alone and a function of t alone, and asking whether such a separation is possible. Inserting the ansatz (9.1.3) into (9.1.2) and dividing throughout by $U(\mathbf{r})\,T(t)$ gives

$$\frac{1}{U(\mathbf{r})}\nabla^2 U(\mathbf{r}) + \lambda' = \frac{B(t)}{T(t)}\frac{dT(t)}{dt} + \frac{C(t)}{T(t)}\frac{d^2T(t)}{dt^2}, \tag{9.1.4}$$

the left-hand side a function of $\mathbf{r}$ alone and the right-hand side a function of t alone. There are then two possibilities: either the two sides are of the form $f(\mathbf{r}) = g(t)$, or both sides equal a common constant μ. The first option would force a relation $x_i = h_i(t)$ or $t = k(\mathbf{r})$, contradicting the independence of $\mathbf{r}$ and t. Hence both sides must equal a common separation constant. Writing $\lambda' - \mu \equiv \lambda$, we obtain

$$\nabla^2 U(\mathbf{r}) + \lambda\,U(\mathbf{r}) = 0 \tag{9.1.5}$$

and

$$C(t)\frac{d^2T(t)}{dt^2} + B(t)\frac{dT(t)}{dt} - \mu\,T(t) = 0. \tag{9.1.6}$$

The equation in t is an ordinary differential equation whose solution presents no difficulty. The spatial equation (9.1.5)—*Helmholtz's equation*—depends explicitly on the form of the Laplacian ∇^2 in the chosen coordinate system. For instance,

$$\textbf{Cartesian:}\quad \nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}, \tag{9.1.7}$$

$$\textbf{Cylindrical:}\quad \nabla^2 = \frac{\partial^2}{\partial \rho^2} + \frac{1}{\rho}\frac{\partial}{\partial \rho} + \frac{1}{\rho^2}\frac{\partial^2}{\partial \theta^2} + \frac{\partial^2}{\partial z^2}, \tag{9.1.8}$$

$$\textbf{Spherical:}\quad \nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\,\frac{\partial}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}. \tag{9.1.9}$$

The form of the equation therefore depends sensitively on the coordinate system, and on separating the partial differential equation in distinct coordinate systems we obtain different families of ordinary differential equations.

Cartesian coordinates. With $U(x,y,z) = X(x)\,Y(y)\,Z(z)$ substituted into the Helmholtz equation (9.1.5), we get

$$\frac{1}{X}\frac{d^2X}{dx^2}+\frac{1}{Y}\frac{d^2Y}{dy^2}+\frac{1}{Z}\frac{d^2Z}{dz^2}+\lambda=0. \tag{9.1.10}$$

With separation constants μ and γ such that

$$\frac{1}{X}\frac{d^2X}{dx^2}=-\mu^2, \qquad \frac{1}{Y}\frac{d^2Y}{dy^2}=-\gamma^2,$$

we are led to the three ordinary differential equations

$$\frac{d^2X}{dx^2}+\mu^2X=0, \tag{9.1.11}$$

$$\frac{d^2Y}{dy^2}+\gamma^2Y=0, \tag{9.1.12}$$

$$\frac{d^2Z}{dz^2}+\big(\lambda-\mu^2-\gamma^2\big)Z=0. \tag{9.1.13}$$

Cylindrical coordinates. With $U(\rho,\theta,z)=P(\rho)\,\Theta(\theta)\,Z(z)$ substituted into the Helmholtz equation, we obtain

$$\frac{1}{P}\frac{d^2P}{d\rho^2}+\frac{1}{\rho P}\frac{dP}{d\rho}+\frac{1}{\rho^2}\frac{1}{\Theta}\frac{d^2\Theta}{d\theta^2}+\frac{Z''}{Z}+\lambda=0, \tag{9.1.14}$$

which separates with constants μ,γ into

$$\frac{d^2Z}{dz^2}+\big(\lambda-\mu^2\big)Z=0, \tag{9.1.15}$$

$$\frac{d^2\Theta}{d\theta^2}+\gamma^2\Theta=0, \tag{9.1.16}$$

$$\frac{d}{d\rho}\left[\rho\,\frac{dP}{d\rho}\right]+\left(\mu^2\rho-\frac{\gamma^2}{\rho}\right)P=0. \tag{9.1.17}$$

Spherical coordinates. With $U(r,\theta,\varphi)=R(r)\,\Theta(\theta)\,\Phi(\varphi)$ substituted into the Helmholtz equation, we get

$$\frac{1}{R}\frac{d^2R}{dr^2}+\frac{2}{rR}\frac{dR}{dr}+\frac{1}{r^2\sin\theta\,\Theta}\frac{d}{d\theta}\left(\sin\theta\,\frac{d\Theta}{d\theta}\right)+\frac{1}{r^2\sin^2\theta\,\Phi}\frac{d^2\Phi}{d\varphi^2}+\lambda=0, \tag{9.1.18}$$

which separates with constants α,β into

$$\frac{d}{dr}\left[r^2\,\frac{dR(r)}{dr}\right]+\big(\lambda r^2-\alpha\big)R(r)=0, \tag{9.1.19}$$

$$\frac{d}{d\theta}\left[\sin\theta\,\frac{d\Theta}{d\theta}\right]+\left(\alpha\sin\theta-\frac{\beta}{\sin\theta}\right)\Theta=0, \tag{9.1.20}$$

$$\frac{d^2\Phi}{d\varphi^2}+\beta\,\Phi=0. \tag{9.1.21}$$

Equation (9.1.17) is *Bessel's equation*, and the radial equation (9.1.19) is the *Euler–Cauchy equation* (it reduces to a Cauchy equation when $\lambda = 0$, that is, when the Helmholtz equation degenerates into the Laplace equation). The equation (9.1.20) is *Legendre's equation*, whose basic features were summarised in Section 1.4. Cartesian separation gives only harmonic-oscillator equations of the same kind, while cylindrical and spherical separation give, in addition, the Bessel, Cauchy and Legendre equations. We shall see in due course that all three are particular instances of a single more general equation: the Sturm–Liouville equation.

9.2 The Sturm–Liouville Problem and Its Eigenfunctions

Define the linear differential operator

$$L \equiv \frac{d}{dx}\left[P(x)\,\frac{d}{dx}\right] - Q(x). \tag{9.2.1}$$

Definition 9.1 (Sturm–Liouville equation) The eigenvalue equation

$$L\big[y(x)\big] = -\lambda\, R(x)\, y(x), \tag{9.2.2}$$

that is, written explicitly,

$$\frac{d}{dx}\left[P(x)\,\frac{dy(x)}{dx}\right] - \big[Q(x) - \lambda R(x)\big]y(x) = 0, \tag{9.2.3}$$

is called the *Sturm–Liouville equation*. The values of λ for which non-trivial solutions of L exist are the *eigenvalues* of L, and the corresponding solutions of the equation are the *eigenfunctions* of L.

Definition 9.2 (Sturm–Liouville problem) With $x \in [a,b]$ and $P(x), Q(x), R(x)$ real and continuous on $[a,b]$, real (i.e. $P = P^*$, $Q = Q^*$, $R = R^*$), and with real constants k_1, k_2, h_1, h_2 at least one of k_1, k_2 and at least one of h_1, h_2 different from zero, the boundary-value problem consisting of (9.2.3) together with the homogeneous boundary conditions

$$\begin{aligned} &k_1 y(a) + k_2 y'(a) = 0, \\ &h_1 y(b) + h_2 y'(b) = 0 \end{aligned} \qquad \begin{pmatrix} k_1 y^*(a) + k_2 y^{*\prime}(a) = 0 \\ h_1 y^*(b) + h_2 y^{*\prime}(b) = 0 \end{pmatrix} \tag{9.2.4}$$

is called the *Sturm–Liouville problem*. If the functions $P(x), R(x)$ are positive on $[a,b]$, the problem is said to be *regular*.

The values λ_i of the parameter λ for which (9.2.3) admits a non-trivial solution are the eigenvalues, and the solutions $y_i(x)$ corresponding to a given λ_i are the eigenfunctions of the problem. Let $y_j(x)$ denote the eigenfunction corresponding to a second eigenvalue λ_j. Then

$$\frac{d}{dx}\left[P(x)\,\frac{dy_i(x)}{dx}\right]-\big[Q(x)-\lambda_i R(x)\big]y_i(x)=0, \tag{9.2.5}$$

$$\frac{d}{dx}\left[P(x)\,\frac{dy_j(x)}{dx}\right]-\big[Q(x)-\lambda_j R(x)\big]y_j(x)=0. \tag{9.2.6}$$

Taking the complex conjugate of (9.2.6),

$$\frac{d}{dx}\left[P(x)\,\frac{dy_j^*(x)}{dx}\right]-\big[Q(x)-\lambda_j^* R(x)\big]y_j^*(x)=0. \tag{9.2.7}$$

Multiplying (9.2.5) by $y_j^*(x)$ and (9.2.7) by $y_i(x)$, subtracting, and integrating over $[a,b]$,

$$\begin{aligned}\int_a^b\left\{y_j^*\,\frac{d}{dx}\left[P\,\frac{dy_i}{dx}\right]-y_i\,\frac{d}{dx}\left[P\,\frac{dy_j^*}{dx}\right]\right\}dx\\ -(\lambda_j^*-\lambda_i)\int_a^b R(x)\,y_j^*(x)\,y_i(x)\,dx=0.\end{aligned} \tag{9.2.8}$$

The integrand of the first integral can be written as a total derivative:

$$\frac{d}{dx}\left\{P\,\frac{dy_i}{dx}\,y_j^*-P\,\frac{dy_j^*}{dx}\,y_i\right\}, \tag{9.2.9}$$

since by direct computation

$$\begin{aligned}\frac{d}{dx}\left\{\left(P\,\frac{dy_i}{dx}\right)y_j^*-\left(P\,\frac{dy_j^*}{dx}\right)y_i\right\}&=y_j^*\,\frac{d}{dx}\left(P\,\frac{dy_i}{dx}\right)+P\,\frac{dy_i}{dx}\,\frac{dy_j^*}{dx}\\ &\quad-y_i\,\frac{d}{dx}\left(P\,\frac{dy_j^*}{dx}\right)-P\,\frac{dy_j^*}{dx}\,\frac{dy_i}{dx}.\end{aligned} \tag{9.2.10}$$

Therefore (9.2.8) reduces to

$$\int_a^b\frac{d}{dx}\left\{\left(P\,\frac{dy_i}{dx}\right)y_j^*-\left(P\,\frac{dy_j^*}{dx}\right)y_i\right\}dx-(\lambda_j^*-\lambda_i)\int_a^b R(x)\,y_j^*(x)\,y_i(x)\,dx=0, \tag{9.2.11}$$

and hence

$$(\lambda_i-\lambda_j^*)\int_a^b R(x)\,y_j^*(x)\,y_i(x)\,dx=\left[P(x)\left\{y_i(x)\,\frac{dy_j^*(x)}{dx}-y_j^*(x)\,\frac{dy_i(x)}{dx}\right\}\right]_a^b. \tag{9.2.12}$$

The boundary conditions (9.2.4), applied at $x=a$ and at $x=b$, give

$$\begin{aligned}(\lambda_i-\lambda_j^*)\int_a^b R(x)\,y_j^*(x)\,y_i(x)\,dx=\big\{&P(b)\big[y_i(b)\,y_j^{*\prime}(b)-y_j^*(b)\,y_i'(b)\big]\\ &-P(a)\big[y_i(a)\,y_j^{*\prime}(a)-y_j^*(a)\,y_i'(a)\big]\big\}=0.\end{aligned} \tag{9.2.13}$$

For $\lambda_i\neq\lambda_j^*$, equation (9.2.13) expresses that the eigenfunctions of the Sturm–Liouville equation are orthogonal on $[a,b]$ with respect to the weight function $R(x)$:

$$\int_a^b R(x)\, y_j^*(x)\, y_i(x)\, dx = 0. \tag{9.2.14}$$

On the other hand, taking $i = j$ in (9.2.13),

$$(\lambda_i - \lambda_i^*) \int_a^b R(x)\, |y_i(x)|^2\, dx = 0, \tag{9.2.15}$$

and since $R(x)$ is positive the integral is strictly positive, so

$$\lambda_i = \lambda_i^*, \tag{9.2.16}$$

i.e. all eigenvalues of the Sturm–Liouville problem are real. With $R(x)$ continuous and positive on $[a,b]$,

$$\int_a^b R(x)\, |y_i(x)|^2\, dx = N \tag{9.2.17}$$

is always a positive constant. The functions $y_i(x)$ may then be normalised through

$$u_i(x) = \frac{1}{\sqrt{N}}\, y_i(x) \tag{9.2.18}$$

so that

$$\int_a^b R(x)\, |u_i(x)|^2\, dx = 1, \tag{9.2.19}$$

and (9.2.14) together with (9.2.19) reads

$$\int_a^b R(x)\, u_i(x)\, u_j^*(x)\, dx = \delta_{ij}. \tag{9.2.20}$$

The family $\{u_i(x)\}$ satisfying (9.2.20) is called an *orthonormal* family with respect to the weight $R(x)$.

Theorem 9.1 (Sturm–Liouville expansion theorem) *For a regular Sturm–Liouville problem (with $P(x), R(x) > 0$ on $[a,b]$) the eigenvalues form an infinite increasing sequence*

$$\lambda_1 < \lambda_2 < \cdots < \lambda_n < \cdots$$

with

$$\lim_{n\to\infty} \lambda_n = \infty,$$

and the corresponding eigenfunctions $\{y_n(x)\}$ form an orthogonal family on $[a,b]$. Moreover, every piecewise-defined function $f(x)$ on $[a,b]$ admits an absolutely and uniformly convergent expansion

$$f(x) = \sum_{n=1}^{\infty} c_n\, y_n(x), \tag{9.2.21}$$

the orthogonality

$$\int_a^b R(x)\, f(x)\, y_m^*(x)\, dx = \delta_{mn} \int_a^b R(x)\, |y_n(x)|^2\, dx \tag{9.2.22}$$

of $\{y_n(x)\}$ *being borne in mind, the coefficients* c_n *being*

$$c_n = \frac{\int_a^b R(x)\, f(x)\, y_n^*(x)\, dx}{\int_a^b R(x)\, |y_n(x)|^2\, dx} = \frac{1}{N} \int_a^b R(x)\, f(x)\, y_n^*(x)\, dx. \tag{9.2.23}$$

In the light of the foregoing, all the ordinary differential equations that came up in the separation of the Helmholtz equation, together with their above-mentioned homogeneous boundary conditions, will admit solutions which form orthogonal eigenfunction families on their respective intervals. For instance, the equation

$$\frac{d^2 y(x)}{dx^2} + \alpha^2 y(x) = 0 \tag{9.2.24}$$

is a Sturm–Liouville equation with $P(x) = 1$, $Q(x) = 0$, $R(x) = -1$. The equation arising from separation of the Helmholtz equation in cylindrical coordinates,

$$\frac{d}{d\rho}\left[\rho \frac{du}{d\rho}\right] + \left(\mu^2 \rho - \frac{\gamma^2}{\rho}\right) u = 0, \tag{9.2.25}$$

the Bessel equation, is a special case of the Sturm–Liouville equation with $P(\rho) = \rho$, $Q(\rho) = 1/\rho$, $R(\rho) = \rho$. Its solutions, the Bessel functions $J_n(\lambda\rho)$ of the first kind, form an orthogonal family with respect to the weight ρ. Likewise, the equation in θ coming from separation of the Helmholtz equation in spherical coordinates,

$$\frac{d}{d\theta}\left[\sin\theta \frac{d\Theta}{d\theta}\right] + \left(\alpha \sin\theta - \frac{\beta}{\sin\theta}\right) \Theta = 0, \tag{9.2.26}$$

and the Legendre equation are particular cases of the Sturm–Liouville equation with $P(\theta) = \sin\theta$, $Q(\theta) = -\sin\theta$, $R(\theta) = 1/\sin\theta$.

9.3 The Variational Method

For an arbitrary parameter t, let $\varphi_j(t)$ be variables and $\dot{\varphi}_j(t) = \partial\varphi_j/\partial t$ $(j = 1, 2, \ldots, n)$ their time derivatives. Consider a function

$$F = F(\varphi_j, \dot{\varphi}_j, t) \qquad (j = 1, 2, \ldots, n). \tag{9.3.1}$$

For two arbitrary values t_1, t_2 of the parameter t, the problem of determining the $\varphi_j(t)$ that render the integral

$$I = \int_{t_1}^{t_2} F(\varphi_j, \dot{\varphi}_j, t)\, dt \tag{9.3.2}$$

extremal (typically minimal) is called the *variational problem*. We shall show that the $\varphi_j(t)$ rendering (9.3.2) extremal satisfy a system of differential equations: this approach to their determination is the *variational method*.

Suppose now that we wish to find the parametric equations of the curve, in the n-dimensional configuration space whose coordinates are the $\varphi_j(t)$, that joins two arbitrary points corresponding to t_1, t_2 and minimises the path integral (9.3.2). To this end, let ε be a parameter and $\alpha_j(t)$ be functions which take arbitrary values on $[t_1, t_2]$ but satisfy the conditions

$$\alpha_j(t_1) = \alpha_j(t_2) = 0 \qquad (j = 1, 2, \ldots, n), \tag{9.3.3}$$

and consider the one-parameter family of curves

$$\Phi_j(t) = \varphi_j(t) + \varepsilon\, \alpha_j(t) \qquad (j = 1, 2, \ldots, n). \tag{9.3.4}$$

Each member of this family passes through the two points fixed by t_1, t_2, and the member parametrised by a particular value of ε (say $\varepsilon = 0$, that is the curve $\varphi_j = \varphi_j(t)$) renders the integral I extremal. Hence, in terms of the φ_j the integral (9.3.2) reads, after the family is substituted,

$$I(\varepsilon) = \int_{t_1}^{t_2} F(\Phi_j, \dot{\Phi}_j, t)\, dt, \tag{9.3.5}$$

and the requirement that the member of (9.3.4) parametrised by $\varepsilon = 0$ extremise (9.3.5) reads

$$\frac{dI(\varepsilon)}{d\varepsilon} \equiv \delta I = 0. \tag{9.3.6}$$

The quantity δI is the *variation* of the integral I. Performing this derivative,

$$\begin{aligned} \left(\frac{dI(\varepsilon)}{d\varepsilon}\right)_{\varepsilon=0} &= \lim_{\varepsilon\to 0} \int_{t_1}^{t_2} \sum_{j=1}^{n} \left(\frac{\partial F}{\partial \Phi_j} \frac{d\Phi_j}{d\varepsilon} + \frac{\partial F}{\partial \dot{\Phi}_j} \frac{d\dot{\Phi}_j}{d\varepsilon} \right) dt \\ &= \lim_{\varepsilon\to 0} \int_{t_1}^{t_2} \sum_{j=1}^{n} \left(\frac{\partial F}{\partial \Phi_j}\, \alpha_j(t) + \frac{\partial F}{\partial \dot{\Phi}_j}\, \dot{\alpha}_j(t) \right) dt. \end{aligned} \tag{9.3.7}$$

The second term on the right may be integrated by parts:

$$\sum_j \int_{t_1}^{t_2} \frac{\partial F}{\partial \dot{\Phi}_j}\, \dot{\alpha}_j(t)\, dt = \left[\sum_j \frac{\partial F}{\partial \dot{\Phi}_j}\, \alpha_j(t) \right]_{t_1}^{t_2} - \sum_j \int_{t_1}^{t_2} \alpha_j \frac{d}{dt} \left(\frac{\partial F}{\partial \dot{\Phi}_j} \right) dt, \tag{9.3.8}$$

and on account of (9.3.3) the integrated term vanishes; whence

$$\begin{aligned} \delta I &= \lim_{\varepsilon\to 0} \int_{t_1}^{t_2} \sum_{j=1}^{n} \left(\frac{\partial F}{\partial \Phi_j} - \frac{d}{dt} \frac{\partial F}{\partial \dot{\Phi}_j} \right) \alpha_j\, dt = 0 \\ &= \int_{t_1}^{t_2} \sum_{j=1}^{n} \left(\frac{\partial F}{\partial \varphi_j} - \frac{d}{dt} \frac{\partial F}{\partial \dot{\varphi}_j} \right) \alpha_j\, dt = 0. \end{aligned} \tag{9.3.9}$$

Since t_1, t_2 are arbitrary and the integral vanishes for every admissible $\alpha_j(t)$, the integrand must vanish; and since the $\alpha_j(t)$ are themselves arbitrary on the open interval (t_1, t_2), the bracket multiplying each α_j must vanish:

$$\frac{d}{dt}\frac{\partial F}{\partial \dot{\varphi}_j} - \frac{\partial F}{\partial \varphi_j} = 0 \qquad (j = 1, 2, \ldots, n). \tag{9.3.10}$$

That is, the parametric equations $\varphi_j = \varphi_j(t)$ of the curve through any two given points (t_1, t_2) that minimises the path integral I must satisfy (9.3.10). In other words, the solutions of the differential system (9.3.10) are exactly the $\varphi_j = \varphi_j(t)$ that minimise the integral (9.3.2). Equations (9.3.10) are called the *Euler–Lagrange equations*.

9.4 Asymptotic Behaviour of the Sturm–Liouville Eigenvalues

Let $P(x)$ and $R(x)$ be positive on $[0, a]$ and consider the Sturm–Liouville problem

$$\begin{cases} \dfrac{d}{dx}\left[P\dfrac{dy}{dx}\right] - Qy + \lambda R y = 0, \quad x \in [0, a], \\ y(0) = y(a) = 0. \end{cases} \tag{9.4.1}$$

Through the function transformation

$$u(x) = \left[P(x)R(x)\right]^{1/4} y(x) \tag{9.4.2}$$

and the change of variable

$$t = \int_0^x \sqrt{R(s)/P(s)}\, ds, \qquad \frac{dt}{dx} = \sqrt{R/P}, \tag{9.4.3}$$

with

$$\ell^* = \int_0^a \sqrt{R(s)/P(s)}\, ds, \tag{9.4.4}$$

the problem (9.4.1) converts, with $r(t)$ a continuous function on $[0, \ell^*]$, into

$$\begin{cases} \dfrac{d^2u}{dt^2} - r(t)\,u + \lambda u = 0, \quad t \in [0, \ell^*], \\ u(0) = u(\ell^*) = 0 \end{cases} \tag{9.4.5}$$

[see Exercise (VII.1)]. The Euler–Lagrange equation associated with the transformed equation,

$$\frac{d^2u}{dt^2} - r(t)\,u + \lambda u = 0 \implies \frac{d}{dt}\frac{\partial F}{\partial(\partial u/\partial t)} - \frac{\partial F}{\partial u} = 0, \tag{9.4.6}$$

is satisfied by an integrand of the form

$$\frac{d}{dt}\frac{\partial F}{\partial u'} = \frac{d}{dt}\frac{du}{dt}, \qquad \frac{\partial F}{\partial u'} = \frac{du}{dt} = u'. \tag{9.4.7}$$

Hence

$$F(u,u',t) = \tfrac{1}{2}(u')^2 + \phi(u,t), \tag{9.4.8}$$

and

$$-\frac{\partial F}{\partial u} = (-r+\lambda)u, \tag{9.4.9}$$

or equivalently

$$\frac{\partial F}{\partial u} = \frac{\partial \phi}{\partial u} = (r-\lambda)u, \qquad \phi = (r-\lambda)\,\frac{u^2}{2} + c(t). \tag{9.4.10}$$

Therefore

$$F(u,u',t) = \tfrac{1}{2}(u')^2 + \tfrac{1}{2}\big(r(t)-\lambda\big)u^2 + c(t). \tag{9.4.11}$$

Since the integral

$$I = \int_0^{\ell^*} F(u,u',t)\,dt \tag{9.4.12}$$

is minimised by the same F regardless of the choice of $c(t)$, we may take $c(t) = 0$. Discarding also the multiplicative factor $\frac{1}{2}$, the variational integral reads

$$I = \int_0^{\ell^*} \big[(u')^2 + ru^2 - \lambda u^2\big]\,dt. \tag{9.4.13}$$

The function u that minimises this integral is at the same time the solution of (9.4.5)—that is, of

$$u'' - ru + \lambda u = 0. \tag{9.4.14}$$

On the other hand, since the solutions of the Sturm–Liouville equation under the homogeneous Dirichlet conditions are "square-integrable" functions, the integral of the product of the weight function with y^2 is a finite constant, that is,

$$\int_0^a R y^2\,dx = \text{constant}, \tag{9.4.15}$$

and using (9.4.2) and (9.4.3),

$$\int_0^{\ell^*} u^2\,dt = \text{constant}. \tag{9.4.16}$$

Hence the existence of this side-condition allows the integral I to be replaced, up to a constant, by

$$I^* = \int_0^{\ell^*} \big[(u')^2 + ru^2\big]\,dt. \tag{9.4.17}$$

The function u minimising this last integral is the same as the solution of (9.4.14) satisfying the constraint (9.4.16). If now the term ru^2 in the integrand is omitted, the variational integral

$$\int_0^{\ell^*} (u')^2\,dt \tag{9.4.18}$$

differs from the original by at most r_M, the maximum of $|r|$.[1] With (9.4.16) also borne in mind, the equation corresponding to the integral (9.4.18) now will have eigenvalues μ that differ from λ by at most a finite amount,

$$u'' + \mu u = 0, \tag{9.4.19}$$

and the integral (9.4.18) can also be evaluated by integration by parts:

$$\int_0^{\ell^*} (u')^2\,dt = \int_0^{\ell^*} u'\,u'\,dt = \big[u\,u'\big]_0^{\ell^*} - \int_0^{\ell^*} u\,u''\,dt = \mu \int_0^{\ell^*} u^2\,dt = \mu\,. \tag{9.4.20}$$

Likewise the integral (9.4.17), viewed in the same way and bearing (9.4.14) in mind, equals

$$\int_0^{\ell^*} \big[(u')^2 + ru^2\big]\,dt = \lambda\,. \tag{9.4.21}$$

Hence μ and λ differ by at most r_M. The problem

$$\begin{cases} u'' + \mu u = 0, \\ u(0) = u(\ell^*) = 0 \end{cases} \tag{9.4.22}$$

has eigenvalues

$$\mu_n = \frac{n^2\pi^2}{\ell^{*2}} \qquad (n = 0, 1, 2, \ldots), \tag{9.4.23}$$

so, denoting by $\wp(\ell^*)$ a finite quantity remaining bounded as $n \to \infty$, the asymptotic relation

$$\lambda_n = \mu_n + \wp(\ell^*) \tag{9.4.24}$$

holds, and consequently

$$\lim_{n\to\infty} \lambda_n = \lim_{n\to\infty} \big[\mu_n + \wp(\ell^*)\big] = \infty. \tag{9.4.25}$$

Dividing (9.4.24) by λ_n,

$$\frac{n^2}{\lambda_n} = \frac{\ell^{*2}}{\pi^2}\big[1 - \wp(\ell^*)\big], \tag{9.4.26}$$

and

[1] By the mean-value theorem,

$$\int_0^{\ell^*} ru^2\,dt = \ell^*\,r(c)\,u^2(c) \quad c \in [0, \ell^*],$$

$$\int_0^{\ell^*} u^2\,dt = \ell^*\,u^2(c) = 1,$$

so $\int_0^{\ell^*} ru^2\,dt = r(c) < r_M$.

$$\lim_{n\to\infty}\frac{n^2}{\lambda_n}=\frac{\ell^{*2}}{\pi^2}. \tag{9.4.27}$$

Returning to the original notation through (9.4.4),

$$\lim_{n\to\infty}\frac{n^2}{\lambda_n}=\frac{1}{\pi^2}\left(\int_0^a\sqrt{R/P}\,dx\right)^2. \tag{9.4.28}$$

Since the asymptotic behaviour of the eigenvalues of (9.4.14) is independent of the boundary data, exactly the same result (9.4.28) is obtained for arbitrary but homogeneous boundary conditions.

Worked Exercise

(VII.1) Show that the Sturm–Liouville problem

$$\begin{cases}\dfrac{d}{dx}\left[P(x)\dfrac{dy}{dx}\right]-Qy+\lambda Ry=0, \quad x\in[0,a],\\ y(0)=y(a)=0,\end{cases} \tag{9.4.29}$$

in which P and R are positive on $[0,a]$, can be reduced through the function transformation

$$u(x)=\left[P(x)\,R(x)\right]^{1/4}y(x)$$

and the change of variable

$$t=\int_0^x\sqrt{R(x')/P(x')}\,dx', \qquad \ell=\int_0^a\sqrt{R/P}\,dx,$$

with $r(t)$ a continuous function, to the boundary-value problem

$$\begin{cases}\dfrac{d^2u}{dt^2}-r(t)\,u+\lambda u=0, \quad t\in[0,\ell],\\ u(0)=u(\ell)=0.\end{cases}$$

Solution: The function transformation $y=(PR)^{-1/4}u$ gives

$$\frac{dy}{dx}=u\,\frac{d}{dx}(PR)^{-1/4}+(PR)^{-1/4}\,\frac{du}{dx}, \tag{9.4.30}$$

$$P\,\frac{dy}{dx}=Pu\,\frac{d}{dx}(PR)^{-1/4}+P(PR)^{-1/4}\,\frac{du}{dx}, \tag{9.4.31}$$

$$\begin{aligned}\frac{d}{dx}\left[P\,\frac{dy}{dx}\right]&=\frac{d}{dx}\left[P\,\frac{d}{dx}(PR)^{-1/4}\right]u+P\,\frac{d}{dx}(PR)^{-1/4}\,\frac{du}{dx}\\ &\quad+\frac{d}{dx}\left[P(PR)^{-1/4}\right]\frac{du}{dx}+P(PR)^{-1/4}\,\frac{d^2u}{dx^2}.\end{aligned} \tag{9.4.32}$$

On the other hand, with $t = \int_0^x \sqrt{R(x')/P(x')}\,dx'$ and $dt/dx = \sqrt{R/P}$, we have

$$\frac{d}{dx} = \frac{d}{dt}\frac{dt}{dx} = \sqrt{\frac{R}{P}}\frac{d}{dt}. \tag{9.4.33}$$

Equation (9.4.32) therefore becomes

$$\begin{aligned}\frac{d}{dx}\left[P\frac{dy}{dx}\right] &= \sqrt{R/P}\,\frac{d}{dt}\left[P\sqrt{R/P}\,\frac{d}{dt}(PR)^{-1/4}\right]u + P\frac{R}{P}\frac{d}{dt}(PR)^{-1/4}\frac{du}{dt}\\ &\quad + \frac{R}{P}\frac{d}{dt}\left[P(PR)^{-1/4}\right]\frac{du}{dt} + P(PR)^{-1/4}\sqrt{R/P}\,\frac{d}{dt}\left[\sqrt{R/P}\,\frac{du}{dt}\right].\end{aligned} \tag{9.4.34}$$

The last term in (9.4.34), expanded, reads

$$\left[P(PR)^{-1/4}\sqrt{R/P}\,\frac{d}{dt}\sqrt{R/P}\right]\frac{du}{dt} = P(PR)^{-1/4}\frac{R}{P}\frac{d^2u}{dt^2}. \tag{9.4.35}$$

Hence (9.4.34) reorganised gives

$$\begin{aligned}\frac{d}{dx}\left[P\frac{dy}{dx}\right] &= R(PR)^{-1/4}\frac{d^2u}{dt^2}\\ &\quad + \left[R\frac{d}{dt}(PR)^{-1/4} + \frac{R}{P}\frac{d}{dt}\left\{P(PR)^{-1/4}\right\} + P(PR)^{-1/4}\sqrt{R/P}\,\frac{d}{dt}\sqrt{R/P}\right]\frac{du}{dt}\\ &\quad + \sqrt{R/P}\,\frac{d}{dt}\left[\sqrt{PR}\,\frac{d}{dt}(PR)^{-1/4}\right]u.\end{aligned} \tag{9.4.36}$$

The coefficient of du/dt on the right of (9.4.36) is seen to vanish on direct computation. The remaining expression for $\frac{d}{dx}[P\,dy/dx]$, substituted into the Sturm–Liouville equation, gives

$$R(PR)^{-1/4}\frac{d^2u}{dt^2} + \left\{\sqrt{R/P}\,\frac{d}{dt}\left[(PR)^{1/2}\frac{d}{dt}(PR)^{-1/4}\right] - (PR)^{-1/4}Q\right\}u + \lambda(PR)^{-1/4}Ru = 0. \tag{9.4.37}$$

Dividing this last equation by $(PR)^{-1/4}R$ and writing

$$r(t) \equiv \frac{Q}{R} - (PR)^{-1/4}\frac{d}{dt}\left[(PR)^{1/2}\frac{d}{dt}(PR)^{-1/4}\right], \tag{9.4.38}$$

the Sturm–Liouville equation reduces to

$$\frac{d^2u}{dt^2} - r(t)\,u + \lambda u = 0. \tag{9.4.39}$$

On the other hand, since $P(x)$ and $R(x)$ are positive on $[0,a]$, the homogeneous boundary conditions $u(x) = [P(x)R(x)]^{1/4}y(x)$ together with $y(0) = y(a) = 0$ become

$$u(0) = u(\ell) = 0, \tag{9.4.40}$$

where $x=0$ corresponds to $t=0$ and $x=a$ to $t=\ell$. Combined, these give the converted initial-value problem.

9.5 Completeness of the Eigenfunctions of the Sturm–Liouville Problem

For an arbitrary continuous function $y=y(x)$, with y, y' and x varying as

$$I(y)=\int_a^b F(y,y',x)\,dx \qquad (a\le x\le b), \tag{9.5.1}$$

the function $y(x)$ that minimises this integral is at the same time the solution of the corresponding Euler–Lagrange equation. The completeness of the eigenfunction system $\{u_n\}_{n=1}^{\infty}$ generated by a regular Sturm–Liouville problem is equivalent to the statement that any function $f\in L^2_R([a,b])$ admits a mean-square convergent expansion

$$f(x)=\sum_{n=1}^{\infty} c_n\,u_n(x), \qquad c_n=\int_a^b R(x)\,f(x)\,u_n^*(x)\,dx, \tag{9.5.2}$$

or equivalently to Parseval's identity

$$\int_a^b R(x)\,|f(x)|^2\,dx=\sum_{n=1}^{\infty}|c_n|^2. \tag{9.5.3}$$

The proof rests on the spectral theorem for self-adjoint compact operators applied to the Green-function operator of the Sturm–Liouville problem.

Sketch of proof. Order the eigenfunctions $\{u_n\}$ by the min–max characterisation of §9.3. For $f\in L^2_R([a,b])$ set $S_N=\sum_{n=1}^{N}c_nu_n$ with $c_n=\langle f,u_n\rangle_R$. Bessel's inequality gives $\sum|c_n|^2\le\|f\|_R^2$, so $\{S_N\}$ is Cauchy in L^2_R and converges to some $S\in L^2_R$. The variational characterisation of λ_n implies $\inf_{g\perp\operatorname{span}\{u_1,\dots,u_N\}}\|f-g\|_R^2\to 0$ as $N\to\infty$, hence $S=f$ and Parseval's identity holds.

9.6 Orthogonality Relations for the Bessel and Legendre Functions

I. Bessel functions. For $n\ge 0$, the equation

$$x^2y''(x)+xy'(x)+\bigl(x^2-n^2\bigr)y(x)=0 \tag{9.6.1}$$

is called the *Bessel equation of order n*. Solving this variable-coefficient equation by the method of Frobenius [see Problem (VII.1)], its two linearly independent solutions $J_n(x)$ and $Y_n(x)$ are respectively the Bessel functions of the first and the second kind of order n, and the general solution is

$$y(x) = c_1 J_n(x) + c_2 Y_n(x). \tag{9.6.2}$$

For $x \to 0$, the $J_n(x)$ remain bounded whereas $\lim_{x\to 0} Y_n(x) \to -\infty$. This circumstance is constantly to be borne in mind in physical applications.

The change of variable $x \to \lambda x$ converts the equation satisfied by $y = y(\lambda x)$ into

$$x^2 y''(x) + x y'(x) + \left(\lambda^2 x^2 - n^2\right) y(x) = 0, \tag{9.6.3}$$

whose solution is therefore

$$y(\lambda x) = c_1 J_n(\lambda x) + c_2 Y_n(\lambda x). \tag{9.6.4}$$

Now let $y_j(x)$ and $y_k(x)$ be eigenfunctions of (9.6.3) corresponding to the eigenvalues λ_j and λ_k. The corresponding equations are

$$x^2 y_j''(x) + x y_j'(x) + \left(\lambda_j^2 x^2 - n^2\right) y_j(x) = 0, \tag{9.6.5}$$

$$x^2 y_k''(x) + x y_k'(x) + \left(\lambda_k^2 x^2 - n^2\right) y_k(x) = 0. \tag{9.6.6}$$

Multiplying the first by y_k and the second by y_j and subtracting,

$$x \frac{d}{dx}\left[y_k y_j' - y_j y_k'\right] + \left(y_k y_j' - y_j y_k'\right) = \left(\lambda_k^2 - \lambda_j^2\right) x y_j y_k, \tag{9.6.7}$$

and rearranging,

$$\frac{d}{dx}\left[x\left(y_k y_j' - y_j y_k'\right)\right] = \left(\lambda_k^2 - \lambda_j^2\right) x y_j y_k. \tag{9.6.8}$$

Now a being an arbitrary positive constant, integrating the last equation over $[0, a]$ gives

$$\left(\lambda_k^2 - \lambda_j^2\right) \int_0^a x y_j y_k \, dx = \left[x\left(y_k y_j' - y_j y_k'\right)\right]_0^a. \tag{9.6.9}$$

Now suppose that the solutions of the Bessel equation are $y_j = J_n(\lambda_j x)$ and $y_k = J_n(\lambda_k x)$, the Bessel functions of the first kind of order n. Then (9.6.9) becomes

$$\int_0^a x J_n(\lambda_j x) J_n(\lambda_k x) \, dx = \frac{a\left[\lambda_j J_n(\lambda_k a) J_n'(\lambda_j a) - \lambda_k J_n(\lambda_j a) J_n'(\lambda_k a)\right]}{\lambda_k^2 - \lambda_j^2}. \tag{9.6.10}$$

Suppose now that, with at least one of h_1, h_2 different from zero, the J_n satisfy the boundary condition

$$\left[h_1 y + h_2 y'\right]_{x=a} = 0. \tag{9.6.11}$$

The eigenfunctions $J_n(\lambda_j x)$ and $J_n(\lambda_k x)$ both satisfy the same homogeneous condition, so

$$h_1 J_n(\lambda_j a) + h_2 \lambda_j J_n'(\lambda_j a) = 0, \tag{9.6.12}$$

$$h_1 J_n(\lambda_k a) + h_2 \lambda_k J_n'(\lambda_k a) = 0. \tag{9.6.13}$$

Multiplying (9.6.12) by $J_n(\lambda_k a)$ and (9.6.13) by $J_n(\lambda_j a)$ and subtracting, for $h_2 \neq 0$,

$$\lambda_j J_n(\lambda_k a) J_n'(\lambda_j a) - \lambda_k J_n(\lambda_j a) J_n'(\lambda_k a) = 0. \tag{9.6.14}$$

Substituting (9.6.14) into (9.6.10) yields the orthogonality relation

$$\int_0^a x J_n(\lambda_j x) J_n(\lambda_k x)\, dx = 0, \tag{9.6.15}$$

i.e. that the first-kind Bessel functions satisfying (9.6.12)–(9.6.13) form an orthogonal family on $[0,a]$ with respect to the weight x. (Since the Bessel equation is itself a Sturm–Liouville equation, this result was indeed to be expected.)

For $j = k$, that is $\lambda_k \to \lambda_j$, equation (9.6.10) gives the indeterminate form

$$\int_0^a x\big[J_n(\lambda_j x)\big]^2 dx = \lim_{\lambda_k \to \lambda_j} \left[\frac{a\lambda_j J_n(\lambda_k a) J_n'(\lambda_j a) - a\lambda_k J_n(\lambda_j a) J_n'(\lambda_k a)}{\lambda_k^2 - \lambda_j^2} \right] = \tfrac{0}{0}, \tag{9.6.16}$$

and applying l'Hôpital's rule,

$$\begin{aligned} \int_0^a x\big[J_n(\lambda_j x)\big]^2 dx = \lim_{\lambda_k \to \lambda_j} \frac{1}{2\lambda_k} \big[& a^2 \lambda_j J_n'(\lambda_k a) J_n'(\lambda_j a) - a J_n(\lambda_j a) J_n'(\lambda_k a) \\ & - a^2 \lambda_k J_n(\lambda_j a) J_n''(\lambda_k a) \big]. \end{aligned} \tag{9.6.17}$$

Substituting $x \to \lambda_k a$ into the original Bessel equation

$$x^2 J_n''(x) + x J_n'(x) + (x^2 - n^2) J_n(x) = 0 \tag{9.6.18}$$

gives

$$\lambda_k^2 a^2 J_n''(\lambda_k a) + \lambda_k a J_n'(\lambda_k a) + \big(\lambda_k^2 a^2 - n^2\big) J_n(\lambda_k a) = 0, \tag{9.6.19}$$

and from this, isolating $a\lambda_k J_n''(\lambda_k a)$,

$$a\lambda_k J_n''(\lambda_k a) = -J_n'(\lambda_k a) - \lambda_k a \left(1 - \frac{n^2}{\lambda_k^2 a^2}\right) J_n(\lambda_k a). \tag{9.6.20}$$

Substituting this into (9.6.17) we obtain

$$\int_0^a x\big[J_n(\lambda_j x)\big]^2 dx = \frac{a^2}{2} \left\{ \big[J_n'(\lambda_j a)\big]^2 + \left(1 - \frac{n^2}{\lambda_j^2 a^2}\right) \big[J_n(\lambda_j a)\big]^2 \right\}. \tag{9.6.21}$$

This together with (9.6.15) gives

$$\int_0^a x J_n(\lambda_j x) J_n(\lambda_k x)\, dx = \frac{a^2}{2} \left\{ \big[J_n'(\lambda_j a)\big]^2 + \left(1 - \frac{n^2}{\lambda_j^2 a^2}\right) \big[J_n(\lambda_j a)\big]^2 \right\} \delta_{jk}, \tag{9.6.22}$$

the orthogonality relation of the Bessel functions of the first kind that satisfy the third-kind homogeneous boundary conditions. Whenever the J_n satisfy first-kind homogeneous boundary conditions ($h_2 = 0$, $h_1 \neq 0$), the second term in the parenthesis on the right vanishes; whenever they satisfy second-kind homogeneous boundary conditions

($h_1 = 0$, $h_2 \neq 0$), the first term vanishes. For example, suppose that λ_j are the positive roots of $J_n(\lambda_j a) = 0$. Then the $J_n(\lambda_j x)$ are the eigenfunctions which are an orthogonal family on $[0,a]$ satisfying first-kind homogeneous conditions there, and the expansion of an arbitrary function $f(x)$ on $[0,a]$ in terms of these is

$$f(x) = \sum_{k=1}^{\infty} c_k J_n(\lambda_k x) \qquad (k = 1,2,3,\ldots). \tag{9.6.23}$$

Multiplying both sides of (9.6.23) by $xJ_n(\lambda_j x)$ and integrating over $[0,a]$, with (9.6.22) in mind,

$$\int_0^a xJ_n(\lambda_k x)\, f(x)\, dx = \sum_{k=1}^{\infty} c_k \int_0^a xJ_n(\lambda_j x) J_n(\lambda_k x)\, dx = c_j \frac{a^2}{2} \left[J_n'(\lambda_j a)\right]^2, \tag{9.6.24}$$

or, for any k, the expansion coefficients of (9.6.23) are

$$c_k = \frac{2}{a^2 \left[J_n'(\lambda_k a)\right]^2} \int_0^a xJ_n(\lambda_k x)\, f(x)\, dx. \tag{9.6.25}$$

Moreover, since the Bessel functions satisfy the recursion relation

$$xJ_n'(x) = nJ_n(x) - xJ_{n+1}(x) \tag{9.6.26}$$

[see Exercise (VII.2)], substituting $x \to \lambda_k a$ gives

$$\lambda_k a J_n'(\lambda_k a) = nJ_n(\lambda_k a) - \lambda_k a J_{n+1}(\lambda_k a), \tag{9.6.27}$$

and since λ_k are positive roots of $J_n(\lambda_k a) = 0$, the relation between the n-th and $(n+1)$-th Bessel functions becomes

$$J_n'(\lambda_k a) = -J_{n+1}(\lambda_k a). \tag{9.6.28}$$

Hence (9.6.25) can also be written

$$c_k = \frac{2}{a^2 \left[J_{n+1}(\lambda_k a)\right]^2} \int_0^a xJ_n(\lambda_k x)\, f(x)\, dx. \tag{9.6.29}$$

II. Legendre functions. For $x \in [-1,+1]$, the equation

$$\left(1 - x^2\right) y''(x) - 2xy'(x) + n(n+1) y(x) = 0 \tag{9.6.30}$$

is called *Legendre's equation*. Solving this variable-coefficient equation by the Frobenius method [see Problem (VII.2)],

$$y(x) = C_0\left[1 - \frac{n(n+1)}{2!}x^2 - \frac{n(n-2)(n+1)(n+3)}{4!}x^4 - \cdots\right] + C_1\left[x - \frac{(n-1)(n+1)}{3!}x^3 - \frac{(n-1)(n-3)(n+2)(n+4)}{5!}x^5 - \cdots\right]. \tag{9.6.31}$$

For $|x| < 1$ both series converge whatever the value of n (integer or not). For $x = \pm 1$ both series tend to infinity (i.e. both diverge). For n a non-negative integer:

(i) if n is even, the first series terminates;
(ii) if n is odd, the second series terminates;

that is, in either case it terminates and a polynomial solution exists. The other solution is always an infinite series, and it diverges at $x = \pm 1$. When n is a non-negative integer, denoting the polynomial solutions of the Legendre equation by $P_n(x)$ (Legendre polynomials of the first kind) and the corresponding infinite-series divergent solutions by $Q_n(x)$ (Legendre functions of the second kind), the general solution reads

$$y(x) = AP_n(x) + BQ_n(x). \tag{9.6.32}$$

From the series (9.6.31) the first few Legendre polynomials read

$$P_0 = C_0, \quad P_1 = C_1 x, \quad P_2 = C_0\left(1 - 3x^2\right), \quad P_3 = C_1\left(x - \frac{5}{3}x^3\right), \ \ldots \tag{9.6.33}$$

The constants C_0 and C_1 being arbitrary, they are conventionally fixed in such a way that, in every case, $P_n(1) = 1$. Then the first few Legendre polynomials read

$$\begin{aligned} P_0(x) &= 1, \\ P_1(x) &= x, \\ P_2(x) &= \tfrac{1}{2}\left(3x^2 - 1\right), \\ P_3(x) &= \tfrac{1}{2}\left(5x^3 - 3x^2\right), \\ &\vdots \end{aligned} \tag{9.6.34}$$

Now consider (9.6.30) for the eigenvalues m and n:

$$\left(1 - x^2\right)P_n'' - 2xP_n' + n(n+1)P_n = 0, \tag{9.6.35}$$

$$\left(1 - x^2\right)P_m'' - 2xP_m' + m(m+1)P_m = 0. \tag{9.6.36}$$

Multiplying the first by P_m and the second by P_n, subtracting, and integrating over $[-1, +1]$,

$$\left[m(m+1) - n(n+1)\right]\int_{-1}^{+1} P_n P_m \, dx = \left[\left(1 - x^2\right)\left(P_m P_n' - P_n P_m'\right)\right]_{-1}^{+1}. \tag{9.6.37}$$

For $m \neq n$ this gives

$$\int_{-1}^{+1} P_n P_m \, dx = 0, \tag{9.6.38}$$

that is, the Legendre polynomials form an orthogonal family on $[-1,+1]$. As they are the solutions of the Legendre equation, which is itself a Sturm–Liouville equation, this result was indeed to be expected. Likewise the Legendre polynomials admit the generating-function representation

$$\frac{1}{\sqrt{1-2tx+t^2}} = \sum_{n=0}^{\infty} P_n(x)\,t^n \tag{9.6.39}$$

[see Problem (VII.3)]. Squaring this,

$$\frac{1}{1-2tx+t^2} = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} P_n(x)\,P_m(x)\,t^{n+m}, \tag{9.6.40}$$

and integrating over $[-1,+1]$,

$$\int_{-1}^{+1} \frac{dx}{1-2tx+t^2} = \frac{1}{t}\ln\left(\frac{1+t}{1-t}\right) = \sum_{n=0}^{\infty} \frac{2t^{2n}}{2n+1}; \tag{9.6.41}$$

bearing in mind that, by (9.6.38), only the $m=n$ terms contribute, this gives

$$\sum_{n=0}^{\infty} t^{2n}\int_{-1}^{+1} P_n^2(x)\,dx = \sum_{n=0}^{\infty} \frac{2t^{2n}}{2n+1}, \tag{9.6.42}$$

and consequently

$$\int_{-1}^{+1} P_n^2(x)\,dx = \frac{2}{2n+1}. \tag{9.6.43}$$

This together with (9.6.38) gives the Legendre polynomial orthonormality

$$\int_{-1}^{+1} P_n(x)\,P_m(x)\,dx = \frac{2}{2n+1}\,\delta_{nm}. \tag{9.6.44}$$

For an arbitrary function $f(x)$ defined on $[-1,+1]$, expanded in terms of the Legendre polynomials,

$$f(x) = \sum_{k=0}^{\infty} a_k P_k(x), \tag{9.6.45}$$

the expansion coefficients, by (9.6.44), read

$$a_n = \frac{2n+1}{2}\int_{-1}^{+1} f(x)\,P_n(x)\,dx. \tag{9.6.46}$$

On the other hand, again with $x\in[-1,+1]$, the equation

$$\left(1-x^2\right)y'' - 2xy' + \left[n(n+1) - \frac{m^2}{1-x^2}\right]y = 0 \tag{9.6.47}$$

or equivalently

$$\frac{d}{dx}\left[\left(1-x^2\right)y'\right]+\left[n(n+1)-\frac{m^2}{1-x^2}\right]y=0 \tag{9.6.48}$$

is called the *associated Legendre equation*. For $m=0$ it reduces to the Legendre equation (9.6.30). The solutions of (9.6.47) are called the *associated Legendre functions*; the first and second-kind solutions being denoted $P_n^m(x)$ and $Q_n^m(x)$, the general solution is

$$y=C_1P_n^m(x)+C_2Q_n^m(x). \tag{9.6.49}$$

With the function transformation

$$P_n^m(x)=\left(1-x^2\right)^{m/2}V(x), \tag{9.6.50}$$

inserted into (9.6.47),

$$\left(1-x^2\right)V''-2(m+1)xV'+\left[n(n+1)-m(m+1)\right]V=0. \tag{9.6.51}$$

On the other hand, the Legendre equation (9.6.30), corresponding to $m=0$, when differentiated m times reads

$$\left(1-x^2\right)\frac{d^{m+2}P_n}{dx^{m+2}}-2(m+1)x\frac{d^{m+1}P_n}{dx^{m+1}}+\left[n(n+1)-m(m+1)\right]\frac{d^mP_n}{dx^m}=0, \tag{9.6.52}$$

or

$$\left(1-x^2\right)\left(\frac{d^mP_n}{dx^m}\right)''-2(m+1)x\left(\frac{d^mP_n}{dx^m}\right)'+\left[n(n+1)-m(m+1)\right]\frac{d^mP_n}{dx^m}=0. \tag{9.6.53}$$

Comparing (9.6.53) with (9.6.51),

$$V(x)=\frac{d^mP_n}{dx^m}, \tag{9.6.54}$$

which by (9.6.50) gives

$$P_n^m(x)=\left(1-x^2\right)^{m/2}\frac{d^mP_n}{dx^m}, \tag{9.6.55}$$

and likewise by an analogous calculation

$$Q_n^m(x)=\left(1-x^2\right)^{m/2}\frac{d^mQ_n}{dx^m}. \tag{9.6.56}$$

Hence the associated Legendre functions are obtained from the Legendre functions by differentiation. The associated Legendre polynomials likewise satisfy the orthogonality relation

$$\int_{-1}^{+1}P_n^m(x)\,P_k^m(x)\,dx=\frac{2}{2n+1}\,\frac{(n+m)!}{(n-m)!}\,\delta_{nk}. \tag{9.6.57}$$

Worked Exercises

(VII.2) The flat faces of an insulated unit-radius homogeneous circular plate are held at zero temperature on its rim. Given the initial temperature distribution $F(\rho)$, determine the temperature distribution at any later instant.

Solution: Owing to the symmetry, the temperature distribution is independent of φ, and so the unknown is $u(\rho,t)$, the solution of

$$\begin{cases} \dfrac{\partial u}{\partial t} = k\left(\dfrac{\partial^2 u}{\partial \rho^2} + \dfrac{1}{\rho}\dfrac{\partial u}{\partial \rho}\right), \quad 0 \le \rho < 1, \\ u(1,t) = 0, \\ u(\rho,0) = F(\rho), \\ |u(\rho,t)| < M. \end{cases} \tag{9.6.58}$$

Applying the method of separation of variables,

$$u(\rho,t) = R(\rho)\,T(t). \tag{9.6.59}$$

With μ the separation constant, the equation gives

$$T' - k\mu T = 0, \qquad \rho^2 R'' + \rho R' - \mu\rho^2 R = 0, \tag{9.6.60}$$

which separates immediately into two ordinary differential equations. For $u(\rho,t)$ to be finite, $\mu = -\lambda^2$, and the two ODEs have solutions

$$T = C_1 e^{-k\lambda^2 t}, \qquad R(\rho) = A_1 J_0(\lambda\rho) + B_1 Y_0(\lambda\rho). \tag{9.6.61}$$

Since the temperature must be finite at $\rho = 0$, $B_1 \equiv 0$. Setting $A_1 C_1 \equiv A$,

$$u(\rho,t) = A e^{-k\lambda^2 t} J_0(\lambda\rho). \tag{9.6.62}$$

The boundary condition then requires

$$u(1,t) = A e^{-k\lambda^2 t} J_0(\lambda) = 0, \tag{9.6.63}$$

so $\lambda_1, \lambda_2, \ldots$ are the eigenvalues, given by the positive roots of $J_0(\lambda) = 0$. Hence the superposition of the admissible solutions is

$$u(\rho,t) = \sum_{m=1}^{\infty} A_m e^{-k\lambda_m^2 t} J_0(\lambda_m \rho) \qquad (m = 1,2,\ldots). \tag{9.6.64}$$

On the other hand, applying the initial condition,

$$u(\rho,0) = F(\rho) = \sum_{m=1}^{\infty} A_m J_0(\lambda_m \rho). \tag{9.6.65}$$

Multiplying both sides by $\rho J_0(\lambda_n\rho)$ and integrating over $[0,1]$,

$$\int_0^1 \rho F(\rho)J_0(\lambda_n\rho)\,d\rho = \sum_{m=1}^{\infty} A_m \int_0^1 \rho J_0(\lambda_m\rho)J_0(\lambda_n\rho)\,d\rho = \tfrac{1}{2}A_n\left[J_0'(\lambda_n)\right]^2 = \tfrac{1}{2}A_n J_1^2(\lambda_n), \tag{9.6.66}$$

so the expansion coefficients read

$$A_n = \frac{2}{J_1^2(\lambda_n)}\int_0^1 \rho F(\rho)J_0(\lambda_n\rho)\,d\rho. \tag{9.6.67}$$

The temperature distribution in the disk is therefore

$$u(\rho,t) = \sum_{m=1}^{\infty}\left[\frac{2}{J_1^2(\lambda_m)}\int_0^1 \rho F(\rho)J_0(\lambda_m\rho)\,d\rho\right] e^{-k\lambda_m^2 t}J_0(\lambda_m\rho). \tag{9.6.68}$$

(VII.3) A homogeneous, conducting cylinder of unit radius and unit height has the initial temperature $f(\rho,z)$. The entire outer surface of the cylinder is held at zero temperature, so that the temperature distribution in the cylinder, at any later instant, at any point in it, is sought.

Solution: Owing to the symmetry, the temperature distribution is again independent of φ, and so the problem reads

$$\begin{cases} \dfrac{\partial u}{\partial t} = k\left(\dfrac{\partial^2 u}{\partial \rho^2} + \dfrac{1}{\rho}\dfrac{\partial u}{\partial \rho} + \dfrac{\partial^2 u}{\partial z^2}\right), \quad 0 \le \rho < 1, \\ u(\rho,z,0) = f(\rho,z), \\ u(\rho,0,t) = u(\rho,1,t) = 0, \\ u(1,z,t) = 0, \\ |u(\rho,z,t)| < M. \end{cases} \tag{9.6.69}$$

For the solution of the initial-value problem we use separation of variables:

$$u(\rho,z,t) = R(\rho)\,Z(z)\,T(t). \tag{9.6.70}$$

With μ and λ separation constants, the equation gives

$$T' + k\lambda^2 T = 0, \qquad \rho^2 R'' + \rho R' + \mu^2\rho^2 R = 0, \qquad Z'' - \gamma^2 Z = 0 \;\; (\mu^2 - \lambda^2 = \gamma^2), \tag{9.6.71}$$

which separates immediately into three ordinary differential equations, with solutions

$$T = C_1 e^{-k\lambda^2 t}, \quad R = C_2 J_0(\mu\rho) + C_3 Y_0(\mu\rho), \quad Z(z) = C_4 e^{\gamma z} + C_5 e^{-\gamma z}. \tag{9.6.72}$$

In the solution of the zeroth-order Bessel equation, for the temperature to be finite at $\rho = 0$, $C_3 \equiv 0$. Hence

$$u(\rho,z,t) = e^{-k\lambda^2 t}J_0(\mu\rho)\left[Ae^{\gamma z} + Be^{-\gamma z}\right]. \tag{9.6.73}$$

The boundary condition with respect to z yields $A = -B$, and

$$u(\rho,z,t) = Ae^{-k\lambda^2 t}J_0(\mu\rho)\left[e^{\gamma z} - e^{-\gamma z}\right]. \tag{9.6.74}$$

Applying the second boundary condition with respect to z,

$$u(\rho,1,t) = Ae^{-k\lambda^2 t}J_0(\mu\rho)\left[e^{\gamma} - e^{-\gamma}\right] = 0, \tag{9.6.75}$$

since the equation $e^{\gamma} - e^{-\gamma} = 0$ would require $e^{2\gamma} = 1 = e^{2in\pi}$ $(n = 0,1,2,\ldots)$, that is $\gamma = in\pi$. Hence

$$u(\rho,z,t) = Ce^{-k\lambda^2 t}J_0(\mu\rho)\sin n\pi z. \tag{9.6.76}$$

The boundary condition in ρ gives the eigenvalues $\mu_1, \mu_2, \mu_3, \ldots$ as the positive roots of

$$J_0(\mu) = 0. \tag{9.6.77}$$

Hence

$$u_{mn}(\rho,z,t) = C_{mn}e^{-k\lambda^2 t}J_0(\mu_m\rho)\sin n\pi z \quad (m = 1,2,\ldots). \tag{9.6.78}$$

With $\lambda_{mn}^2 = \mu^2 - \gamma^2 = \mu_m^2 + n^2\pi^2$ $(m = 1,2,\ldots)$, the possible solutions satisfying every boundary condition save the initial one are, with $m = 0,1,2,\ldots, n = 0,1,2,\ldots$,

$$u_{mn}(\rho,z,t) = C_{mn}e^{-k(\mu_m^2 + n^2\pi^2)t}J_0(\mu_m\rho)\sin n\pi z, \tag{9.6.79}$$

and their superposition—bearing in mind that for $n = 0$ the solution vanishes—reads

$$u(\rho,z,t) = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} C_{mn}e^{-k(\mu_m^2 + n^2\pi^2)t}J_0(\mu_m\rho)\sin n\pi z. \tag{9.6.80}$$

Lastly, applying the initial condition,

$$u(\rho,z,0) = f(\rho,z) = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} C_{mn}J_0(\mu_m\rho)\sin n\pi z. \tag{9.6.81}$$

Multiplying both sides of this equation by $\rho J_0(\mu_p\rho)\sin q\pi z$ and integrating ρ and z over $[0,1]$, the expansion coefficients work out to be

$$C_{mn} = \frac{4}{J_1^2(\mu_m)}\int_0^1\int_0^1 \rho f(\rho,z)J_0(\mu_m\rho)\sin n\pi z\, d\rho\, dz. \tag{9.6.82}$$

The required solution is then obtained by inserting these constants into the general solution.

(VII.4) A unit-radius drum membrane is fastened tightly along its rim. The drum is struck a sharp blow that gives the membrane an initial velocity $F(\rho,\varphi)$, and the membrane is then released and left vibrating. Find the membrane displacement at any later instant.

Solution: The displacement $u(\rho,\varphi,t)$ of the membrane satisfies

$$\begin{cases} \dfrac{\partial^2 u}{\partial t^2} = \alpha^2\left(\dfrac{\partial^2 u}{\partial \rho^2} + \dfrac{1}{\rho}\dfrac{\partial u}{\partial \rho} + \dfrac{1}{\rho^2}\dfrac{\partial^2 u}{\partial \varphi^2}\right), \\ u(1,\varphi,t) = 0, \\ u_t(\rho,\varphi,0) = 0, \\ u(\rho,\varphi,0) = F(\rho,\varphi), \\ |u(\rho,\varphi,t)| < M. \end{cases} \tag{9.6.83}$$

For the solution of this initial-value problem we use separation of variables,

$$u(\rho,\varphi,t) = R(\rho)\,\Phi(\varphi)\,T(t), \tag{9.6.84}$$

λ being the separation constant. Then

$$\frac{T'}{\alpha^2 T} = \frac{R''}{R} + \frac{1}{\rho}\frac{R'}{R} + \frac{1}{\rho^2}\frac{\Phi''}{\Phi} = -\lambda^2. \tag{9.6.85}$$

The time-dependent equation reads

$$T'' + \alpha^2\lambda^2 T = 0. \tag{9.6.86}$$

Rearranging the second equality of (9.6.85), with μ a fresh separation constant,

$$\rho^2\frac{R''}{R} + \rho\frac{R'}{R} + \lambda^2\rho^2 = -\frac{\Phi''}{\Phi} = \mu^2, \tag{9.6.87}$$

giving

$$\Phi'' + \mu\Phi = 0, \qquad \rho^2 R'' + \rho R' + \left(\lambda^2\rho^2 - \mu^2\right)R = 0. \tag{9.6.88}$$

The solutions of (9.6.86) and (9.6.88) are

$$T = A_1\cos\alpha\lambda t + B_1\sin\alpha\lambda t, \quad \Phi = A_2\cos\mu\varphi + B_2\sin\mu\varphi, \quad R = A_3 J_\mu(\lambda\rho) + B_3 Y_\mu(\lambda\rho). \tag{9.6.89}$$

Since $u(\rho,\varphi,t)$ in the variable φ has periodicity 2π, $\Phi(\alpha) = \Phi(\alpha + 2\pi)$, that is $\mu = m = 0,1,2,\ldots$. Also, since the solution must be finite at $\rho = 0$, $B_3 \equiv 0$. Furthermore, the condition $u_t(\rho,\varphi,0) = 0$ gives $B_1 \equiv 0$, and hence

$$u_m(\rho,\varphi,t) = J_m(\lambda\rho)\cos\alpha\lambda t\,\left[A\cos m\varphi + B\sin m\varphi\right]. \tag{9.6.90}$$

Applying the first of these boundary conditions, $u(1,\varphi,t) = 0$ gives the eigenvalues $\lambda_{m1}, \lambda_{m2}, \ldots$ as the positive roots of $J_m(\lambda) = 0$. The general solution is then

$$u(\rho,\varphi,t) = \sum_{m=0}^{\infty}\sum_{k=1}^{\infty} J_m(\lambda_{mk}\rho)\cos\alpha\lambda_{mk}t\,\left[A_{mk}\cos m\varphi + B_{mk}\sin m\varphi\right]. \tag{9.6.91}$$

Applying the initial condition,

$$u(\rho,\varphi,0)=F(\rho,\varphi)=\sum_{m=0}^{\infty}\sum_{k=1}^{\infty}A_{mk}J_m(\lambda_{mk}\rho)\cos m\varphi+\sum_{m=0}^{\infty}\sum_{k=1}^{\infty}B_{mk}J_m(\lambda_{mk}\rho)\sin m\varphi. \tag{9.6.92}$$

The constants of the solution then split into

$$C_m\equiv\sum_{k=1}^{\infty}A_{mk}J_m(\lambda_{mk}\rho),\qquad D_m\equiv\sum_{k=1}^{\infty}B_{mk}J_m(\lambda_{mk}\rho), \tag{9.6.93}$$

and

$$F(\rho,\varphi)=\sum_{m=0}^{\infty}C_m\cos m\varphi+\sum_{m=0}^{\infty}D_m\sin m\varphi. \tag{9.6.94}$$

From this

$$C_m=\begin{cases}\dfrac{1}{\pi}\displaystyle\int_0^{2\pi}F(\rho,\varphi)\cos m\varphi\,d\varphi & (m=1,2,3\ldots),\\[2ex] \dfrac{1}{2\pi}\displaystyle\int_0^{2\pi}F(\rho,\varphi)\,d\varphi=C_0 & (m=0),\end{cases} \tag{9.6.95}$$

$$D_m=\frac{1}{\pi}\int_0^{2\pi}F(\rho,\varphi)\sin m\varphi\,d\varphi\qquad(m=1,2,3,\ldots), \tag{9.6.96}$$

and (9.6.95) and (9.6.96) fed into (9.6.93) give

$$\begin{aligned}A_{mk}&=\frac{2}{[J_{m+1}(\lambda_{mk})]^2}\int_0^1\rho J_m(\lambda_{mk}\rho)C_m\,d\rho\\&=\frac{2}{\pi[J_{m+1}(\lambda_{mk})]^2}\int_0^1\int_0^{2\pi}\rho F(\rho,\varphi)J_m(\lambda_{mk}\rho)\cos m\varphi\,d\rho\,d\varphi,\end{aligned} \tag{9.6.97}$$

$$\begin{aligned}B_{mk}&=\frac{2}{[J_{m+1}(\lambda_{mk})]^2}\int_0^1\rho J_m(\lambda_{mk}\rho)D_m\,d\rho\\&=\frac{2}{\pi[J_{m+1}(\lambda_{mk})]^2}\int_0^1\int_0^{2\pi}\rho F(\rho,\varphi)J_m(\lambda_{mk}\rho)\sin m\varphi\,d\rho\,d\varphi.\end{aligned} \tag{9.6.98}$$

The required solution is then obtained by inserting these constants into the general solution.

(VII.5) A thin plate, bounded by two concentric circles of radii a and b, has both flat faces insulated. The two bounding rims are kept respectively at the (continuous) temperatures $f(\varphi)$ and $g(\varphi)$. Find the steady temperature distribution at any point (ρ,φ).

Solution: The heat-conduction equation

$$k\nabla^2u=\frac{\partial u}{\partial t} \tag{9.6.99}$$

in the steady state reduces to the Laplace equation $\nabla^2u=0$. Hence the problem is

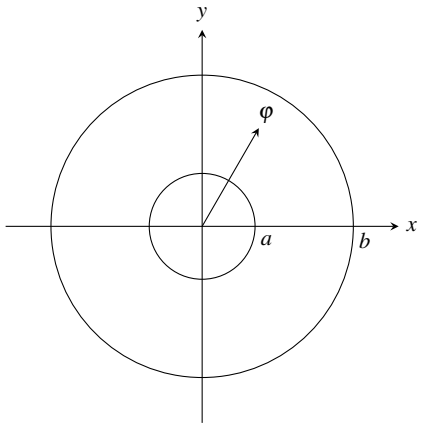


Fig. 9.1 Annular plate of inner radius a and outer radius b.

$$\begin{cases} \dfrac{\partial^2 u}{\partial \rho^2} + \dfrac{1}{\rho}\dfrac{\partial u}{\partial \rho} + \dfrac{1}{\rho^2}\dfrac{\partial^2 u}{\partial \varphi^2} = 0, \\ u(a,\varphi) = f(\varphi), \\ u(b,\varphi) = g(\varphi), \\ u(\rho,\alpha) = u(\rho,\alpha+2\pi), \\ |u(\rho,\varphi)| < M. \end{cases} \tag{9.6.100}$$

Separating, with λ the separation constant,

$$\rho^2 R'' + \rho R' - \lambda^2 R = 0, \qquad \Phi'' + \lambda^2 \Phi = 0, \tag{9.6.101}$$

the first being a Cauchy equation; with the periodicity $\Phi(\alpha) = \Phi(\alpha+2\pi)$ giving $\lambda = n = 0, 1, 2, \ldots$, the solutions are

$$R(\rho) = A_1 \rho^{\lambda} + B_1 \rho^{-\lambda}, \qquad \Phi(\varphi) = A_2 \cos\lambda\varphi + B_2 \sin\lambda\varphi. \tag{9.6.102}$$

By the periodicity in Φ, $\Phi_n(\varphi) = A_2 \cos n\varphi + B_2 \sin n\varphi$, so

$$R_n(\rho) = A_1 \rho^n + B_1 \rho^{-n}. \tag{9.6.103}$$

On the other hand, for $n = 0$, the equations (9.6.101)

$$\rho R'' + R' = 0, \qquad \Phi'' = 0, \tag{9.6.104}$$

have the general solutions

$$R_0(\rho) = A_0 + B_0 \ln\rho, \qquad \Phi_0(\varphi) = C_0\varphi + D_0. \tag{9.6.105}$$

By periodicity, $C_0 \equiv 0$. Setting $A_0 D_0 \equiv \alpha_0$, $B_0 D_0 \equiv \beta_0$, the superposition of admissible solutions reads

$$u(\rho,\varphi)=\left(\alpha_0+\beta_0\ln\rho\right)+\sum_{n=1}^{\infty}\left(\alpha_n\rho^n+\beta_n\rho^{-n}\right)\cos n\varphi+\sum_{n=1}^{\infty}\left(\gamma_n\rho^n+\delta_n\rho^{-n}\right)\sin n\varphi. \tag{9.6.106}$$

Applying the inhomogeneous boundary conditions,

$$u(a,\varphi)=f(\varphi)=\left(\alpha_0+\beta_0\ln a\right)+\sum_{n=1}^{\infty}\left(\alpha_n a^n+\beta_n a^{-n}\right)\cos n\varphi+\sum_{n=1}^{\infty}\left(\gamma_n a^n+\delta_n a^{-n}\right)\sin n\varphi, \tag{9.6.107}$$

which gives

$$\alpha_0+\beta_0\ln a=\frac{1}{2\pi}\int_0^{2\pi}f(\varphi)\,d\varphi, \tag{9.6.108}$$

$$\alpha_n a^n+\beta_n a^{-n}=\frac{1}{\pi}\int_0^{2\pi}f(\varphi)\cos n\varphi\,d\varphi, \tag{9.6.109}$$

$$\gamma_n a^n+\delta_n a^{-n}=\frac{1}{\pi}\int_0^{2\pi}f(\varphi)\sin n\varphi\,d\varphi, \tag{9.6.110}$$

and likewise

$$u(b,\varphi)=g(\varphi)=\left(\alpha_0+\beta_0\ln b\right)+\sum_{n=1}^{\infty}\left(\alpha_n b^n+\beta_n b^{-n}\right)\cos n\varphi+\sum_{n=1}^{\infty}\left(\gamma_n b^n+\delta_n b^{-n}\right)\sin n\varphi, \tag{9.6.111}$$

giving

$$\alpha_0+\beta_0\ln b=\frac{1}{2\pi}\int_0^{2\pi}g(\varphi)\,d\varphi, \tag{9.6.112}$$

$$\alpha_n b^n+\beta_n b^{-n}=\frac{1}{\pi}\int_0^{2\pi}g(\varphi)\cos n\varphi\,d\varphi, \tag{9.6.113}$$

$$\gamma_n b^n+\delta_n b^{-n}=\frac{1}{\pi}\int_0^{2\pi}g(\varphi)\sin n\varphi\,d\varphi. \tag{9.6.114}$$

The $4n+2$ relations (9.6.109) and (9.6.113) suffice to determine the $4n+2$ constants in the solution.

(VII.6) A hollow homogeneous spherical surface of unit radius is held at the surface potential $u_0\sin^2\theta\cos 2\varphi$. Find the potentials produced (a) inside, and (b) outside, by such a surface charge distribution.

Solution: The potential produced inside or outside an empty spherical surface satisfies

$$\nabla^2 u(r,\theta,\varphi)=0, \tag{9.6.115}$$

the Laplace equation, which in spherical coordinates reads

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial u}{\partial r}\right)+\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\,\frac{\partial u}{\partial\theta}\right)+\frac{1}{r^2\sin^2\theta}\frac{\partial^2 u}{\partial\varphi^2}=0. \tag{9.6.116}$$

With separation

$$u(r,\theta,\varphi) = R(r)\,\Theta(\theta)\,\Phi(\varphi), \tag{9.6.117}$$

λ and μ being separation constants, we obtain three ordinary differential equations:

$$\begin{cases} r^2R'' + 2rR' + \lambda^2 R = 0, \\ \Phi'' + \mu^2\Phi = 0, \\ \sin\theta\,\dfrac{d}{d\theta}\left(\sin\theta\,\dfrac{d\Theta}{d\theta}\right) - \left(\lambda^2\sin^2\theta + \mu^2\right)\Theta = 0. \end{cases} \tag{9.6.118}$$

The first of these is a Cauchy–Euler equation; with β an arbitrary constant, putting $R \sim r^\beta$ gives $\lambda^2 = -\beta(\beta+1)$, and the general solution is then

$$R(r) = Ar^\beta + \frac{B}{r^{\beta+1}}. \tag{9.6.119}$$

On account of the periodicity condition for the second equation of (9.6.118), μ takes the values $\mu = m = 0,1,2,\ldots$, that is positive integers. Hence

$$\Theta(\varphi) = C\cos m\varphi + D\sin m\varphi. \tag{9.6.120}$$

The third equation of (9.6.118), with the substitution $\xi = \cos\theta$, becomes

$$(1-\xi^2)\Theta'' - 2\xi\Theta' + \left[\beta(\beta+1) - \frac{m^2}{1-\xi^2}\right]\Theta = 0, \tag{9.6.121}$$

which is, as is well known, the associated Legendre equation, and since the z-axis points lie within our problem, only for $\beta = n = 0,1,2,\ldots$ can a finite solution exist. In this case the solution is

$$\Theta(\theta) = EP_n^m(\cos\theta) + FQ_n^m(\cos\theta). \tag{9.6.122}$$

At $\theta = 0,\pi$ ($\xi = \pm 1$) the Q_n^m are not finite, so on the points lying on the z-axis $F \equiv 0$ must hold. Hence

$$\Theta(\theta) = EP_n^m(\cos\theta). \tag{9.6.123}$$

With (9.6.119), (9.6.120) and (9.6.123) the admissible solutions are

$$u_{nm}(r,\theta,\varphi) = \left(A_n r^n + \frac{B_n}{r^{n+1}}\right)\left(C_m\cos m\varphi + D_m\sin m\varphi\right)E_{nm}P_n^m(\cos\theta). \tag{9.6.124}$$

The solutions thus depend on the distance r.

(a) For $0 \le r < 1$, since $r = 0$ may also occur, the solution should be finite there, so $B_n \equiv 0$. Renaming the constants, a superposition of the admissible solutions reads

$$u(r,\theta,\varphi) = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} r^n P_n^m(\cos\theta)\left(\alpha_{mn}\cos m\varphi + \beta_{mn}\sin m\varphi\right). \tag{9.6.125}$$

Applying the boundary condition,

$$u(1,\theta,\varphi) = u_0 \sin^2\theta \cos 2\varphi = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} P_n^m(\cos\theta)\big(\alpha_{mn}\cos m\varphi + \beta_{mn}\sin m\varphi\big), \tag{9.6.126}$$

or equivalently

$$u_0\big(1-\xi^2\big)\cos 2\varphi = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} P_n^m(\xi)\big(\alpha_{mn}\cos m\varphi + \beta_{mn}\sin m\varphi\big). \tag{9.6.127}$$

Multiplying both sides of this equation first by $P_k^q \sin q\varphi$ and integrating over $\xi \in [-1,+1]$ and $\varphi \in [0,2\pi]$, the left-hand side is identically zero, so for any value of m $\beta_{mn} = 0$. Once again, multiplying both sides of the equation by $P_k^q \cos q\varphi$ and integrating over φ and ξ,

$$\begin{aligned} u_0\int_{-1}^{+1}\big(1-\xi^2\big)P_k^q\,d\xi\int_0^{2\pi}\cos 2\varphi\cos q\varphi\,d\varphi &= \sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\alpha_{mn}\int_{-1}^{+1}P_k^m P_n^q\,d\xi\int_0^{2\pi}\cos m\varphi\cos q\varphi\,d\varphi \\ &= u_0\pi\int_{-1}^{+1}\big(1-\xi^2\big)P_k^2\,d\xi = \pi\sum_{n=0}^{\infty}\alpha_{qn}\int_{-1}^{+1}P_k^q P_n^q\,d\xi = \pi\alpha_{qk}\,\frac{2}{2k+1}\,\frac{(k+q)!}{(k-q)!}. \end{aligned} \tag{9.6.128}$$

For $q \neq 2$ the left-hand side vanishes, so for $q \neq 2$ all $\alpha_{qk} = 0$. The α_{2k} are

$$\alpha_{2k} = \frac{2k+1}{2}\,\frac{(k-2)!}{(k+2)!}\,u_0\int_{-1}^{+1}\big(1-\xi^2\big)P_k^2(\xi)\,d\xi = 0. \tag{9.6.129}$$

Now of all the P_k^2, only one is, up to a constant, equal to $(1-\xi^2)$ and the others differ from it. In fact, only

$$P_2^2 = 3\big(1-\xi^2\big). \tag{9.6.130}$$

Hence

$$\alpha_{2k} = \frac{2k+1}{2}\,\frac{(k-2)!}{(k+2)!}\,\frac{u_0}{3}\int_{-1}^{+1}P_2^2 P_k^2\,d\xi = \frac{u_0}{3}\,\delta_{k2}. \tag{9.6.131}$$

That is, $\alpha_{22} = u_0/3$ and $\alpha_{2k} = 0$ $(k \neq 2)$. The solution sought, with $P_2^2(\xi)$ inserted, becomes

$$u(r,\theta,\varphi) = u_0 r^2 \sin^2\theta\cos 2\varphi \qquad (0 \le r < 1). \tag{9.6.132}$$

(b) For $r > 1$, since $r \to \infty$ may also occur, for the solution to be finite, in (9.6.124) $A_n \equiv 0$ must hold. Hence

$$u(r,\theta,\varphi) = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{1}{r^{n+1}}P_n^m(\cos\theta)\big(\alpha_{mn}\cos m\varphi + \beta_{mn}\sin m\varphi\big), \tag{9.6.133}$$

and applying the boundary condition,

$$u(1,\xi,\varphi) = u_0\big(1-\xi^2\big)\cos 2\varphi = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty}P_n^m(\xi)\big(\alpha_{mn}\cos m\varphi + \beta_{mn}\sin m\varphi\big). \tag{9.6.134}$$

Following exactly the same procedure as in (a) one finds, in this case too, all $\beta_{mn} \equiv 0$ and $\alpha_{22} = u_0/3,\ \alpha_{mn} = 0\ (m \neq 2, n \neq 2)$. Hence for $r > 1$ the potential is

$$u(r,\theta,\varphi) = \frac{u_0}{r^3}\sin^2\theta\cos 2\varphi \qquad (1 < r \leq \infty). \tag{9.6.135}$$

Problems

(VII.1) For $n \geq 0$, show that the general solution of the Bessel equation

$$x^2y'' + xy' + (x^2 - n^2)y = 0,$$

solved by the Frobenius method, is, for $n \neq 0, 1, 2, \ldots$,

$$\begin{aligned} y &= Ax^n\left[1 - \frac{x^2}{2(2n+2)} + \frac{x^4}{2\cdot 4(2n+2)(2n+4)} - \cdots\right] \\ &\quad + Bx^{-n}\left[1 - \frac{x^2}{2(2-2n)} + \frac{x^4}{2\cdot 4(2-2n)(4-2n)} - \cdots\right] \\ &= CJ_n(x) + DJ_{-n}(x). \end{aligned}$$

(VII.2) E and F being arbitrary constants and n an integer, show that the general solution of the Bessel equation can be written

$$y = EJ_n(x) + FY_n(x) = EJ_n(x) + F\left[\frac{J_n(x)\cos n\pi - J_{-n}(x)}{\sin n\pi}\right],$$

in particular for $n = 0, 1, 2, \ldots$,

$$y = EJ_n(x) + F\lim_{p\to n}\left[\frac{J_p(x)\cos p\pi - J_{-p}(x)}{\sin p\pi}\right].$$

(VII.3) Solving by the Frobenius method,

$$\left(1 - x^2\right)y'' - 2xy' + n(n+1)y = 0,$$

show that the general solution of the Legendre equation is, with C_0 and C_1 arbitrary constants,

$$\begin{aligned} y &= C_0\left[1 - \frac{n(n+1)}{2!}x + \frac{n(n-2)(n+1)(n+3)}{4!}x^4 - \cdots\right] \\ &\quad + C_1\left[x - \frac{(n-1)(n+2)}{3!}x^3 + \frac{(n-1)(n-3)(n+2)(n+4)}{5!}x^5 - \cdots\right]. \end{aligned}$$

(VII.4) On the bottom face of a hollow cylinder of radius a and height ℓ, the surface charge distribution corresponding to the potential $f(\rho,\varphi)$ is given. The rest of the

cylinder is uncharged. Solving the initial-value problem

$$\begin{aligned}
\nabla^2 u = u_{\rho\rho} + \frac{1}{\rho}u_\rho + \frac{1}{\rho^2}u_{\varphi\varphi} + u_{zz} = 0, \quad 0 \le r \le a, \quad 0 \le z \le \ell,\\
u(\rho,\varphi,0) = f(\rho,\varphi) \quad (f(a,\varphi) = 0),\\
u(\rho,\varphi,\ell) = 0,\\
u(a,\varphi,z) = 0,\\
|u(\rho,\varphi,z)| < M,
\end{aligned}$$

determine the electrical potential within the cylinder.

(VII.5) In a very long, hollow homogeneous conducting cylinder of inner radius a and outer radius b, with thermal conductivity K, the inner and outer surfaces are kept continuously at zero temperature. Given the initial temperature distribution $f(\rho)$ in the cylinder, determine the temperature distribution at any later instant by solving

$$\begin{aligned}
\frac{\partial u}{\partial t} = K\left(\frac{\partial^2 u}{\partial \rho^2} + \frac{1}{\rho}\frac{\partial u}{\partial \rho}\right), \quad a \le \rho \le b, \quad t > 0,\\
u(\rho,0) = f(\rho),\\
u(a,t) = u(b,t) = 0,\\
|u(\rho,t)| < M.
\end{aligned}$$

(VII.6) Solve the initial-value problem

$$\begin{aligned}
\frac{\partial^2 u}{\partial t^2} = \alpha^2\left(\frac{\partial^2 u}{\partial \rho^2} + \frac{1}{\rho}\frac{\partial u}{\partial \rho} + \frac{1}{\rho^2}\frac{\partial^2 u}{\partial \varphi^2}\right), \quad a \le \rho \le b, \quad t > 0,\\
u(\rho,\varphi,0) = f(\rho,\varphi),\\
u_t(\rho,\varphi,0) = 0,\\
u(a,\varphi,t) = u(b,\varphi,t) = 0,\\
|u(\rho,\varphi,t)| < M,
\end{aligned}$$

in cylindrical coordinates by separation of variables.

(VII.7) A heavy uniform chain of length ℓ and constant linear mass density hangs vertically from a fixed point. The chain is initially perturbed slightly so that at $t = 0$ its shape is given by $f(x)$, and is then released from this state. Find the modes of small oscillation of the chain.

(VII.8) Determine the potential produced inside, at any point, by a hollow homogeneous spherical surface of radius a held at the surface potential $f(\theta,\varphi)$.

(VII.9) A hollow homogeneous spherical surface of unit radius is held at potential $u_0 =$ const. on one half of its surface, and at zero on the other. Find the potential produced (a) inside, and (b) outside, by such a surface charge distribution.

(VII.10) A homogeneous thin string of length ℓ, fastened at one end, is rotating freely about that point. Friction and gravity being neglected, the string spins about its fixed end with constant angular speed $\omega =$ const. in a horizontal plane. Slightly perturbed from this rectilinear configuration the string begins to vibrate. The dif-

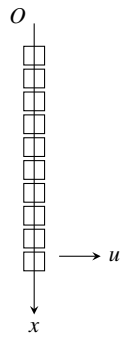


Fig. 9.2 Heavy chain of length ℓ suspended at O, with transverse displacement u.

ferential equation governing the displacement u as a function of the distance x from the fixed end and the time t is, with $a^2 \equiv \omega^2/2$,

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial}{\partial x}\left[\left(\ell^2 - x^2\right)\frac{\partial u}{\partial x}\right], \quad 0 \le x \le \ell.$$

Solve the initial-value problem

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

to determine the modes of vibration of the string.

Chapter 10
Finite Integral Transforms

Abstract When the initial-value problem for a PDE is not expressed in Cartesian coordinates the finite Fourier transforms of Chapter VI no longer apply. The appropriate transform kernel is determined from the coordinate system and the boundary conditions and is identified with the eigenfunctions of an associated Sturm–Liouville problem. The general theory is developed and then applied to canonical problems of mathematical physics on cylindrical and spherical domains.

10.1 Determination of the Integral Transform

Consider the most general second-order linear partial differential operator

$$\wp u(x_\alpha) = \sum_{\alpha,\beta=0}^{3} A_{\alpha\beta}\,\frac{\partial^2 u}{\partial x_\alpha\,\partial x_\beta} + \sum_{\alpha=0}^{3} B_\alpha\,\frac{\partial u}{\partial x_\alpha} + Cu, \tag{10.1.1}$$

with $A_{\alpha\beta} = A_{\beta\alpha}$ $(\alpha,\beta = 0,1,2,3)$, and the associated inhomogeneous PDE of second order

$$\wp u = f. \tag{10.1.2}$$

We wish to determine the appropriate integral-transform kernel for the initial-value problem associated with (10.1.2), and to show how the problem is then solved by that transform. With $K(x_i,m)$ the yet-to-be-determined kernel of an integral transform with respect to the variable $x_i \in [a,b]$, the transform is defined by

$$\bar{u}(m,x_\beta) = \int_a^b u(x_\alpha)\,K(x_i,m)\,dx_i \qquad (\beta \neq i). \tag{10.1.3}$$

Applied to (10.1.2), with $\bar{f}(m,x_\beta) = \int_a^b f(x_\alpha)\,K(x_i,m)\,dx_i$, this gives

$$\int_a^b \Big(\sum_{\alpha,\beta\neq i} A_{\alpha\beta}\,\frac{\partial^2 u}{\partial x_\alpha \partial x_\beta} + \sum_{\alpha\neq i} B_\alpha\,\frac{\partial u}{\partial x_\alpha}\Big) K(x_i,m)\,dx_i$$
$$+ \int_a^b \Big(A_{ii}\,\frac{\partial^2 u}{\partial x_i^2} + 2\sum_{\beta\neq i} A_{i\beta}\,\frac{\partial^2 u}{\partial x_i \partial x_\beta} + B_i\,\frac{\partial u}{\partial x_i} + Cu\Big) K(x_i,m)\,dx_i = \bar{f}(m,x_\beta). \tag{10.1.4}$$

Whenever, for $\alpha,\beta \neq i$, the coefficients $A_{\alpha\beta}$ and B_α are independent of x_i, the first integral on the left-hand side reduces to

$$\sum_{\alpha,\beta\neq i} A_{\alpha\beta}\,\frac{\partial^2 \bar{u}}{\partial x_\alpha \partial x_\beta} + \sum_{\alpha\neq i} B_\alpha\,\frac{\partial \bar{u}}{\partial x_\alpha}. \tag{10.1.5}$$

The first, second and third terms of the second integral in (10.1.4) are integrated by parts in turn. For the $\partial^2 u/\partial x_i^2$ term, two integrations by parts give

$$\int_a^b A_{ii}\,\frac{\partial^2 u}{\partial x_i^2} K\,dx_i = \Big[A_{ii}\,\frac{\partial u}{\partial x_i} K\Big]_a^b - \int_a^b \frac{\partial u}{\partial x_i}\,\frac{\partial (A_{ii}K)}{\partial x_i}\,dx_i,$$

and a further integration by parts gives

$$= \Big[\big(A_{ii}\tfrac{\partial u}{\partial x_i}\big)K - u\,\frac{\partial (A_{ii}K)}{\partial x_i}\Big]_a^b + \int_a^b u\,\frac{\partial^2 (A_{ii}K)}{\partial x_i^2}\,dx_i.$$

The mixed-derivative term yields

$$2\sum_{\beta\neq i}\int_a^b \Big(A_{i\beta}\,\frac{\partial^2 u}{\partial x_i \partial x_\beta}\Big) K\,dx_i = \Big[2\sum_{\beta\neq i} A_{i\beta}\,\frac{\partial u}{\partial x_\beta} K\Big]_a^b - \int_a^b 2\sum_{\beta\neq i}\frac{\partial u}{\partial x_\beta}\,\frac{\partial A_{i\beta}K}{\partial x_i}\,dx_i,$$

and the first-derivative term in x_i gives

$$\int_a^b B_i\,\frac{\partial u}{\partial x_i} K\,dx_i = \big[B_i u K\big]_a^b - \int_a^b u\,\frac{\partial (B_i K)}{\partial x_i}\,dx_i.$$

Substituting these into the second integral of (10.1.4) we obtain

$$\int_a^b \Big(A_{ii}\,\frac{\partial^2 u}{\partial x_i^2} + 2\sum_{\beta\neq i} A_{i\beta}\,\frac{\partial^2 u}{\partial x_i \partial x_\beta} + B_i\,\frac{\partial u}{\partial x_i} + Cu\Big) K(x_i,m)\,dx_i$$
$$= \Big\{ \Big[A_{ii}\,\frac{\partial u}{\partial x_i} + \Big(B_i - \frac{\partial A_{ii}}{\partial x_i}\Big)u\Big] K(x_i,m) - A_{ii}u\,\frac{\partial K}{\partial x_i} + 2\sum_{\beta\neq i} A_{i\beta}\,\frac{\partial u}{\partial x_\beta} K\Big\}_a^b \tag{10.1.6}$$
$$+ \int_a^b 2\sum_{\beta\neq i}\frac{\partial u}{\partial x_\beta}\,\frac{\partial A_{i\beta}K}{\partial x_i}\,dx_i + \int_a^b \Big(\frac{\partial^2 A_{ii}K}{\partial x_i^2} - \frac{\partial B_i K}{\partial x_i} + CK\Big)u\,dx_i.$$

If, for $i \neq \beta$, the off-diagonal coefficients vanish, $A_{i\beta} = 0$, then one of the integrals in (10.1.6) drops out:

$$\sum_{\beta\neq i}\int_a^b \frac{\partial u}{\partial x_\beta}\,\frac{\partial A_{i\beta}K}{\partial x_i}\,dx_i = 0. \tag{10.1.7}$$

Writing $\wp_i u \equiv A_{ii}\,\partial^2 u/\partial x_i^2 + B_i\,\partial u/\partial x_i + Cu$ for the part of $\wp$ that contains the derivatives with respect to x_i, and $\wp'$ for the remainder ($\wp = \wp_i u + \wp' u$), we now demand that the remaining integrand of the second integral in (10.1.6) vanish. With λ_m^2 a parameter independent of x_i, this requirement reads

$$\frac{\partial^2 A_{ii}K}{\partial x_i^2} - \frac{\partial B_i K}{\partial x_i} + CK = -\lambda_m^2 K, \tag{10.1.8}$$

so that

$$\int_a^b \Big(\frac{\partial^2 A_{ii}K}{\partial x_i^2} - \frac{\partial B_i K}{\partial x_i} + CK\Big)u\,dx_i = -\lambda_m^2\int_a^b uK\,dx_i = -\lambda_m^2\bar{u}. \tag{10.1.9}$$

The transform (10.1.3) therefore turns the PDE (10.1.2) into

$$\sum_{\alpha,\beta\neq i} A_{\alpha\beta}\,\frac{\partial^2\bar{u}}{\partial x_\alpha \partial x_\beta} + \sum_{\alpha\neq i} B_\alpha\,\frac{\partial\bar{u}}{\partial x_\alpha} - \lambda_m^2\bar{u} = \bar{f} - \Big[\Big(A_{ii}\,\frac{\partial u}{\partial x_i} + B_i u\Big)K - u\,\frac{\partial A_{ii}K}{\partial x_i}\Big]_a^b. \tag{10.1.10}$$

The right-hand boundary contribution involves the values of u, $\partial u/\partial x_i$, $K(x_i,m)$ and its derivative at $x_i = a$ and $x_i = b$, which are fixed by the boundary conditions of the original problem. In compact form,

$$\wp'\bar{u} - \lambda_m^2\bar{u} = \bar{f} - \Big[\Big(A_{ii}\,\frac{\partial u}{\partial x_i} + B_i u\Big)K - u\,\frac{\partial A_{ii}K}{\partial x_i}\Big]_a^b. \tag{10.1.11}$$

Once the boundary terms on the right-hand side are computed from the boundary conditions, (10.1.11) is an equation for $\bar{u}$ that no longer contains the variable x_i.

We now look for the most explicit form of the transform. Let N be a constant, $R(x_i)$ a weight function and $\bar{K}(x_i,m)$ a new kernel related to K by

$$K(x_i,m) = \frac{1}{N}\,R(x_i)\,\bar{K}(x_i,m). \tag{10.1.12}$$

Substituting (10.1.12) into (10.1.8) and using

$$\begin{aligned}\frac{d^2 A_{ii}R\bar{K}}{dx_i^2} &= \frac{d}{dx_i}\Big(\frac{dA_{ii}R}{dx_i}\bar{K} + A_{ii}R\,\frac{d\bar{K}}{dx_i}\Big)\\ &= \frac{d^2 A_{ii}R}{dx_i^2}\bar{K} + 2\,\frac{dA_{ii}R}{dx_i}\,\frac{d\bar{K}}{dx_i} + A_{ii}R\,\frac{d^2\bar{K}}{dx_i^2},\end{aligned}$$

we obtain

$$A_{ii}R\,\frac{d^2\bar{K}}{dx_i^2} + \Big(2\,\frac{dA_{ii}R}{dx_i} - B_i R\Big)\frac{d\bar{K}}{dx_i} + \Big(\frac{d^2 A_{ii}R}{dx_i^2} - \frac{dB_i R}{dx_i}\Big)\bar{K} + CR\bar{K} = -\lambda_m^2 R\bar{K}. \tag{10.1.13}$$

If $R(x_i)$ is now chosen so as to satisfy

$$\frac{dA_{ii}R}{dx_i} = B_i R, \tag{10.1.14}$$

and we set $A_{ii}R(x_i) \equiv P(x_i)$ and $-CR(x_i) \equiv Q(x_i)$, then (10.1.13) reduces to the Sturm–Liouville form

$$\frac{d}{dx_i}\Big[P(x_i)\,\frac{d\bar{K}}{dx_i}\Big] - Q(x_i)\bar{K} + \lambda_m^2 R(x_i)\bar{K} = 0. \tag{10.1.15}$$

The choice (10.1.8) together with (10.1.14) guarantees that the kernels $\bar{K}$ are eigenfunctions of a Sturm–Liouville problem of the form (10.1.15). Equation (10.1.14) can be rewritten as

$$R\,\frac{dA_{ii}}{dx_i} + A_{ii}\,\frac{dR}{dx_i} = B_i R, \qquad \frac{dR}{dx_i} = -\frac{1}{A_{ii}}\Big(\frac{dA_{ii}}{dx_i} - B_i\Big)R,$$

so that, in terms of the $\wp_i$ coefficients,

$$R(x_i) = \exp\Big[-\int \frac{1}{A_{ii}}\Big(\frac{dA_{ii}}{dx_i} - B_i\Big)dx_i\Big]. \tag{10.1.16}$$

The condition (10.1.14) therefore determines $R(x_i)$ from the coefficients A_{ii} and B_i, and once R is known, the kernel $\bar{K}$ is fixed as the eigenfunction of the Sturm–Liouville problem (10.1.15).

The eigenfunctions of a Sturm–Liouville problem with homogeneous boundary conditions form an orthogonal family. Hence the kernels $\bar{K}(x_i, m)$ form an orthogonal family if, in addition to (10.1.15), the boundary terms in the transformed (10.1.11) can be evaluated through the boundary conditions of the original problem. We shall therefore choose those boundary conditions for $\bar{K}$ that are the homogeneous counterparts of the boundary conditions on u in the initial-value problem for (10.1.2).

For the initial-value problem associated with (10.1.2), take the boundary conditions

$$\wp u = f, \qquad \Big[k_1\frac{\partial u}{\partial x_i} - h_1 u\Big]_{x_i=a} = \varphi_1(x_\beta), \qquad \Big[k_2\frac{\partial u}{\partial x_i} + h_2 u\Big]_{x_i=b} = \varphi_2(x_\beta) \quad (\beta \neq i). \tag{10.1.17}$$

Then the kernel of the appropriate integral transform is determined by the auxiliary problem

$$\frac{d}{dx_i}\Big[P\,\frac{d\bar{K}}{dx_i}\Big] - Q\bar{K} + \lambda_m^2 R\bar{K} = 0, \qquad \Big[k_1\frac{d\bar{K}}{dx_i} - h_1\bar{K}\Big]_{x_i=a} = 0, \qquad \Big[k_2\frac{d\bar{K}}{dx_i} + h_2\bar{K}\Big]_{x_i=b} = 0. \tag{10.1.18}$$

The eigenfunctions of (10.1.18) satisfy the orthogonality relation

$$\int_a^b R(x_i)\,\bar{K}(x_i, m)\,\bar{K}(x_i, n)\,dx_i = N\,\delta_{mn}, \tag{10.1.19}$$

with $R(x_i)$ as the weight function. Consequently $u(x_\alpha)$ admits an expansion in the kernels $\bar{K}(x_i, m)$,

$$u(x_\alpha) = \sum_{m=0}^{\infty} a_m \bar{K}(x_i, m), \tag{10.1.20}$$

with coefficients

$$a_m = \frac{1}{N} \int_a^b R(x_i)\, u(x_\alpha)\, \bar{K}(x_i, m)\, dx_i. \tag{10.1.21}$$

On the other hand, by (10.1.3) and (10.1.12),

$$\bar{u}(m, x_\beta) = \frac{1}{N} \int_a^b R(x_i)\, u(x_\alpha)\, \bar{K}(x_i, m)\, dx_i, \tag{10.1.22}$$

showing that the $\bar{u}(m, x_\beta)$ are precisely the expansion coefficients in (10.1.21). When $\bar{u}(m, x_\beta)$ is known, $u(x_\alpha)$ is recovered by the inversion formula

$$u(x_\alpha) = \sum_{m=0}^{\infty} \bar{u}(m, x_\beta)\, \bar{K}(x_i, m). \tag{10.1.23}$$

Returning to the boundary contribution in (10.1.11), with $dA_{ii}R/dx_i \equiv B_i R$ and $P \equiv A_{ii}R$, one finds

$$\left[\left(A_{ii}\frac{\partial u}{\partial x_i} + B_i u\right)K - u\frac{\partial A_{ii}K}{\partial x_i}\right]_a^b = \frac{1}{N}\left[P(x_i)\left(\frac{\partial u}{\partial x_i}\bar{K} - u\frac{d\bar{K}}{dx_i}\right)\right]_a^b = \frac{1}{N}\left[\frac{1}{k_2}(P\bar{K})_{x_i=b}\varphi_2 - \frac{1}{k_1}(P\bar{K})_{x_i=a}\varphi_1\right] \equiv F_b - F_a,$$

in which the boundary conditions (10.1.17) and the homogeneous boundary conditions in (10.1.18) have been used to eliminate the unknown derivatives of u at $x_i = a, b$. The transformed equation therefore reads

$$\mathscr{D}\bar{u} - \lambda_m^2 \bar{u} = \bar{f} + (F_b - F_a). \tag{10.1.24}$$

This equation, together with the initial conditions on the variables other than x_i, gives $\bar{u}(m, x_\beta)$, after which the solution of the original initial-value problem (10.1.17) follows from the inversion formula (10.1.23),

$$u(x_\alpha) = \sum_{m=0}^{\infty} \bar{u}(m, x_\beta)\, \bar{K}(x_i, m) \qquad (\beta \neq i,\ \alpha = 0, 1, 2, 3).$$

Worked Exercises

Exercise VIII.1 (Hanging string – Daniel Bernoulli problem).

A heavy string of length ℓ hangs from one end in the vertical direction. Let $\phi(x,t)$ denote a horizontal force per unit length acting on the string and write $F(x,t) = \phi(x,t)/\rho$, where ρ is the linear mass density of the string. With the x-axis directed along the equilibrium configuration of the string, the origin O at the lower end and the string in equilibrium hanging vertically, the equation governing the small transverse displace-

ments $u(x,t)$ from equilibrium is

$$\frac{\partial^2 u}{\partial t^2} = k^2 \frac{\partial}{\partial x}\Big(x \frac{\partial u}{\partial x}\Big) + F(x,t),$$

with initial data $u(x,0) = u_0(x)$, $u_t(x,0) = u_1(x)$. Determine the displacement of the string.

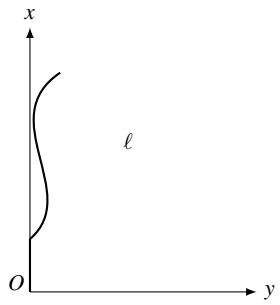


Fig. 10.1 Hanging string of length ℓ in the equilibrium and a deflected configuration.

Solution. The initial–boundary-value problem describing the displacement is

$$\begin{aligned} &\frac{\partial^2 u}{\partial t^2} = k^2 \frac{\partial}{\partial x}\Big(x \frac{\partial u}{\partial x}\Big) + F(x,t), \quad 0 \le x \le \ell,\ t > 0, \\ &u(x,0) = u_0(x), \qquad u_t(x,0) = u_1(x), \\ &u(\ell,t) = 0, \qquad |u(x,t)| < M. \end{aligned} \tag{10.1.25}$$

The x-derivative part of the operator is

$$\mathcal{O}_x u \equiv \frac{\partial}{\partial x}\Big(x \frac{\partial u}{\partial x}\Big) = x \frac{\partial^2 u}{\partial x^2} + \frac{\partial u}{\partial x},$$

so $A_{xx} = x$, $B_x = 1$, $C = 0$, and

$$\begin{aligned} &R(x) \equiv \exp\Big[-\int \frac{1}{A_{xx}}\Big(\frac{dA_{xx}}{dx} - B_x\Big)dx\Big] = 1, \\ &P(x) = A_{xx}R(x) = x, \qquad Q(x) = -CR(x) = 0. \end{aligned}$$

The auxiliary problem for the kernel of the integral transform is therefore

$$\frac{d}{dx}\Big[x \frac{d\bar{K}}{dx}\Big] + \lambda^2 \bar{K} = 0, \tag{10.1.26}$$

$$\big[\bar{K}\big]_{x=\ell} = 0, \tag{10.1.27}$$

$$\big|\bar{K}\big|_{x=0} < M. \tag{10.1.28}$$

With the change of variable $z = 2x^{1/2}$, equation (10.1.26) becomes the zero-order Bessel equation

$$z^2 \frac{d^2\bar{K}}{dz^2} + z\frac{d\bar{K}}{dz} + \lambda^2 z^2 \bar{K} = 0,$$

whose finite solution at $z = 0$ is $J_0(\lambda z) = J_0(2\lambda x^{1/2})$. The boundary condition (10.1.27) gives, with $2\lambda_n \equiv \beta_n$ $(n = 1, 2, \ldots)$,

$$J_0(2\lambda \ell^{1/2}) = 0,$$

which determines the eigenvalues. The normalisation constant is

$$N = \int_0^\ell J_0^2(\beta_n x^{1/2})\, dx = \frac{1}{2}\int_0^{2\sqrt{\ell}} z J_0^2(\lambda_n z)\, dz = \frac{1}{2}\left[2\ell J_1^2(\beta_n \ell^{1/2})\right],$$

and the kernel of the integral transform is

$$K(x,n) = \frac{1}{N} J_0(\beta_n x^{1/2}).$$

The transform itself is

$$\bar{u}(n,t) = \frac{1}{N}\int_0^\ell u(x,t)\, J_0(\beta_n x^{1/2})\, dx,$$

and applied to the equation (10.1.25) it gives

$$\frac{d^2\bar{u}}{dt^2} - \frac{k^2}{N}\int_0^\ell \frac{\partial}{\partial x}\Big(x\frac{\partial u}{\partial x}\Big) J_0(\beta_n x^{1/2})\, dx = \bar{F}(n,t), \tag{10.1.29}$$

where

$$\bar{F}(n,t) = \frac{1}{N}\int_0^\ell F(x,t) J_0(\beta_n x^{1/2})\, dx.$$

The remaining integral on the left of (10.1.29) can be written more explicitly as

$$\int_0^\ell \frac{\partial}{\partial x}\Big(x\frac{\partial u}{\partial x}\Big) K(n,x)\, dx = \int_0^\ell x\frac{\partial^2 u}{\partial x^2} K(n,x)\, dx + \int_0^\ell \frac{\partial u}{\partial x} K(n,x)\, dx, \tag{10.1.30}$$

and the first term on the right is integrated by parts as

$$\int_0^\ell x\frac{\partial^2 u}{\partial x^2} K_n\, dx = \Big[xK_n\frac{\partial u}{\partial x}\Big]_0^\ell - \int_0^\ell \frac{\partial u}{\partial x}\frac{d(xK_n)}{dx}\, dx,$$

where the first term on the right vanishes by virtue of (10.1.27). A second integration by parts gives

$$-\int_0^\ell \frac{\partial u}{\partial x}\frac{d(xK_n)}{dx}\, dx = -\Big[u\Big(x\frac{dK_n}{dx} + K_n\Big)\Big]_0^\ell + \int_0^\ell u\frac{d^2(xK_n)}{dx^2}\, dx = \big[uK_n\big]_{x=0} + \int_0^\ell u\frac{d^2(xK_n)}{dx^2}\, dx.$$

Hence

$$\int_0^\ell x\,\frac{\partial^2 u}{\partial x^2}K_n\,dx = \big[uK_n\big]_{x=0} + \int_0^\ell u\,\frac{d^2(xK_n)}{dx^2}\,dx. \tag{10.1.31}$$

The second term on the right of (10.1.30) integrates similarly,

$$\int_0^\ell \frac{\partial u}{\partial x}K_n\,dx = \big[uK_n\big]_0^\ell - \int_0^\ell u\,\frac{dK_n}{dx}\,dx = -\big[uK_n\big]_{x=0} - \int_0^\ell u\,\frac{dK_n}{dx}\,dx. \tag{10.1.32}$$

Substituting (10.1.31) and (10.1.32) into (10.1.30),

$$\int_0^\ell \frac{\partial}{\partial x}\Big(x\,\frac{\partial u}{\partial x}\Big)K_n\,dx = \int_0^\ell u\Big[\frac{d^2(xK_n)}{dx^2} - \frac{dK_n}{dx}\Big]dx,$$

in which the bracket on the right reduces to

$$\frac{d^2(xK_n)}{dx^2} - \frac{dK_n}{dx} = x\,\frac{d^2K_n}{dx^2} + \frac{dK_n}{dx} = \frac{d}{dx}\Big(x\,\frac{dK_n}{dx}\Big) = -\lambda_n^2 K_n,$$

on using (10.1.26). Therefore

$$\int_0^\ell \frac{\partial}{\partial x}\Big(x\,\frac{\partial u}{\partial x}\Big)K_n\,dx = -\lambda_n^2\int_0^\ell uK_n\,dx = -\lambda_n^2\bar{u}(n,t).$$

Inserting this into (10.1.29) and using $\beta_n = 2\lambda_n$, the transformed equation reads

$$\frac{d^2\bar{u}(n,t)}{dt^2} + \frac{\beta_n^2 k^2}{4}\,\bar{u}(n,t) = \bar{F}(n,t).$$

Its general solution is

$$\bar{u}(n,t) = A_n\cos\frac{k\beta_n}{2}t + B_n\sin\frac{k\beta_n}{2}t + \frac{2}{k\beta_n}\int_0^t \bar{F}(n,\tau)\sin\frac{k\beta_n}{2}(t-\tau)\,d\tau,$$

and the initial conditions

$$\bar{u}(n,0) = \int_0^\ell u(x,0)K_n\,dx = \bar{u}_0(n) \equiv A_n, \quad \Big(\frac{d\bar{u}}{dt}\Big)_{t=0} = \int_0^\ell u_t(x,0)K_n\,dx = \bar{u}_1(n) \equiv \frac{k\beta_n}{2}B_n$$

fix

$$\bar{u}(n,t) = \bar{u}_0(n)\cos\frac{k\beta_n}{2}t + \frac{2}{k\beta_n}\bar{u}_1(n)\sin\frac{k\beta_n}{2}t + \frac{2}{k\beta_n}\int_0^t \bar{F}(n,\tau)\sin\frac{k\beta_n}{2}(t-\tau)d\tau.$$

The solution is then recovered from the inversion formula

$$u(x,t) = \sum_{n=1}^{\infty}\bar{u}(n,t)\,J_0(\beta_n x^{1/2}).$$

Exercise VIII.2 (Rectangular elastic membrane).

A rectangular elastic membrane $0 \le x \le a$, $0 \le y \le b$, clamped at its edges, is subject to a force per unit area $P(x,y,t)$. With T the constant tension and $Z(x,y,t) \equiv P(x,y,t)/T$, the small transverse oscillations $u(x,y,t)$ are governed by

$$\begin{aligned}
&\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} - Z(x,y,t),\\
&u(0,y,t) = u(a,y,t) = 0, \quad u(x,0,t) = u(x,b,t) = 0,\\
&u(x,y,0) = u_0(x,y), \quad u_t(x,y,0) = u_1(x,y).
\end{aligned}$$

Solve by integral transform.
Solution. The basic operator that will eliminate x is $\mathscr{O}_x \equiv \partial^2/\partial x^2$, so $A_{xx} = 1$, $B_x = 0$, $C = 0$, $R = 1$, $P = 1$, $Q = 0$. The auxiliary problem

$$\frac{d^2\bar{K}}{dx^2} + \lambda^2 \bar{K} = 0, \qquad [\bar{K}]_{x=0} = [\bar{K}]_{x=a} = 0$$

has eigenvalues $\lambda_m = m\pi/a$ $(m = 1,2,\ldots)$ and eigenfunctions $\sin\lambda_m x$, with normalisation $\bar{N} = \int_0^a \sin^2\lambda_m x\,dx = a/2$. The corresponding finite sine transform on $[0,a]$ is

$$\bar{u}(m,y,t) = \frac{2}{a}\int_0^a u(x,y,t)\sin\frac{m\pi}{a}x\,dx. \tag{10.1.33}$$

Using the boundary conditions on x,

$$\frac{2}{a}\int_0^a \frac{\partial^2 u}{\partial x^2}\sin\frac{m\pi}{a}x\,dx = -\lambda_m^2\bar{u}(m,y,t).$$

Applying (10.1.33) to the equation gives the transformed problem

$$\begin{aligned}
&\frac{\partial^2\bar{u}}{\partial y^2} - \lambda_m^2\bar{u} = \frac{1}{c^2}\frac{\partial^2\bar{u}}{\partial t^2} - \bar{Z}(m,y,t),\\
&\bar{u}(m,0,t) = \bar{u}(m,b,t) = 0,\\
&\bar{u}(m,y,0) = \bar{u}_0(m,y), \quad \bar{u}_t(m,y,0) = \bar{u}_1(m,y).
\end{aligned} \tag{10.1.34}$$

For this problem, $\mathscr{O}_y\bar{u} \equiv \partial^2\bar{u}/\partial y^2$, so $R(y) = 1$, $P(y) = 1$, $Q(y) = 0$ and the auxiliary problem reads

$$\frac{d^2\hat{K}}{dy^2} + \mu^2\hat{K} = 0, \qquad [\hat{K}]_{y=0} = [\hat{K}]_{y=b} = 0,$$

giving eigenvalues $\mu_n = n\pi/b$ $(n = 1,2,\ldots)$ and eigenfunctions $\sin\mu_n y$. With $\hat{N} = \int_0^b \sin^2\mu_n y\,dy = b/2$, the finite sine transform on $[0,b]$,

$$\hat{u}(m,n,t) = \frac{2}{b}\int_0^b \bar{u}(m,y,t)\sin\frac{n\pi}{b}y\,dy,$$

applied to (10.1.34), gives, with $\frac{2}{b}\int_0^b \partial^2\bar{u}/\partial y^2 \sin(n\pi y/b)dy = -\mu_n^2\hat{u}$,

$$\frac{1}{c^2}\frac{d^2\hat{u}}{dt^2} + (\lambda_m^2+\mu_n^2)\hat{u} = \hat{Z}(m,n,t),$$

together with $\hat{u}(m,n,0) = \hat{u}_0(m,n)$, $\hat{u}_t(m,n,0) = \hat{u}_1(m,n)$. With $\omega_{mn}^2 = \lambda_m^2+\mu_n^2$, this elementary ordinary-differential initial-value problem has solution

$$\hat{u}(m,n,t) = A_{mn}\cos\omega_{mn}t + B_{mn}\sin\omega_{mn}t + \frac{c^2}{\omega_{mn}}\int_0^t \hat{Z}(m,n,\tau)\sin\omega_{mn}(t-\tau)d\tau.$$

Imposing the initial conditions, $A = \hat{u}_0(m,n)$ and $B = \hat{u}_1(m,n)/\omega_{mn}$, so that

$$\hat{u}(m,n,t) = \hat{u}_0(m,n)\cos\omega_{mn}t + \frac{1}{\omega_{mn}}\hat{u}_1(m,n)\sin\omega_{mn}t + \frac{c^2}{\omega_{mn}}\int_0^t \hat{Z}(m,n,\tau)\sin\omega_{mn}(t-\tau)d\tau. \tag{10.1.35}$$

The two inverse transforms then give

$$\bar{u}(m,y,t) = \sum_{n=1}^{\infty}\hat{u}(m,n,t)\sin\frac{n\pi}{b}y, \qquad u(x,y,t) = \sum_{m=1}^{\infty}\bar{u}(m,y,t)\sin\frac{m\pi}{a}x,$$

and on inserting (10.1.35) the solution sought is

$$u(x,y,t) = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\hat{u}(m,n,t)\sin\frac{m\pi}{a}x\sin\frac{n\pi}{b}y.$$

Exercise VIII.3 (Cylindrical rod).

A homogeneous cylindrical rod of radius a, with the heat loss at its ends neglected, has an initial temperature distribution $f(r,\varphi)$ along its body and loses heat through its lateral surface by convection to an external medium at zero temperature. Determine the temperature distribution inside the rod.

Solution. The initial–boundary-value problem is

$$\begin{aligned}
&\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial\varphi^2} = \frac{1}{k}\frac{\partial u}{\partial t}, \quad 0\le r\le a,\ 0\le\varphi\le 2\pi,\ t>0,\\
&u(r,\varphi,0) = f(r,\varphi),\\
&\left[\frac{\partial u}{\partial r} + hu\right]_{r=a} = 0,\\
&u(r,0,t) = u(r,2\pi,t), \quad |u(0,\varphi,t)| < \infty.
\end{aligned} \tag{10.1.36}$$

We first eliminate φ by an integral transform. The basic operator is $\mathscr{D}_\varphi u \equiv \partial^2 u/\partial\varphi^2$, so $R(\varphi)=1$, $P(\varphi)=1$, $Q(\varphi)=0$ and the auxiliary problem is

$$\frac{d^2\bar{K}}{d\varphi^2}+\mu^2\bar{K}=0, \qquad [\bar{K}]_{\varphi=0}=[\bar{K}]_{\varphi=2\pi},$$

whose two independent solutions are $\cos m\varphi$ and $\sin m\varphi$ with $\mu = m = 0,1,2,\ldots$. Set

$$\bar{K}(2m,\varphi)\equiv\cos m\varphi, \qquad \bar{K}(2m-1,\varphi)\equiv\sin m\varphi.$$

The normalisation, irrespective of the parity of γ, is

$$\bar{N}=\int_0^{2\pi}[\bar{K}(\gamma,\varphi)]^2d\varphi=\int_0^{2\pi}\sin^2 m\varphi\,d\varphi=\int_0^{2\pi}\cos^2 m\varphi\,d\varphi=\pi,$$

so

$$\frac{1}{\bar{N}}R(\varphi)\bar{K}(\gamma,\varphi)=\begin{cases}\dfrac{1}{\pi}\cos m\varphi & (\gamma=2m),\\ \dfrac{1}{\pi}\sin m\varphi & (\gamma=2m-1).\end{cases}$$

The transform on $[0,2\pi]$ is

$$\bar{u}(r,\gamma,t)=\frac{1}{\pi}\int_0^{2\pi}u(r,\varphi,t)\bar{K}(\gamma,\varphi)d\varphi.$$

Applied to (10.1.36), since

$$\gamma=2m:\ \frac{1}{\pi}\int_0^{2\pi}\frac{\partial^2 u}{\partial\varphi^2}\cos m\varphi\,d\varphi=-m^2\bar{u}, \quad \gamma=2m-1:\ \frac{1}{\pi}\int_0^{2\pi}\frac{\partial^2 u}{\partial\varphi^2}\sin m\varphi\,d\varphi=-m^2\bar{u},$$

i.e.

$$\frac{1}{\pi}\int_0^{2\pi}\frac{\partial^2 u}{\partial\varphi^2}\bar{K}(\gamma,\varphi)d\varphi=-m^2\bar{u},$$

and with $\bar{f}(\gamma,r)=\frac{1}{\pi}\int_0^{2\pi}f(r,\varphi)\bar{K}(\gamma,\varphi)d\varphi$, the transformed problem becomes

$$\begin{aligned}&\frac{\partial^2\bar{u}}{\partial r^2}+\frac{1}{r}\frac{\partial\bar{u}}{\partial r}-\frac{m^2}{r^2}\bar{u}=\frac{1}{k}\frac{\partial\bar{u}}{\partial t}, \quad 0\le r\le a,\ t>0,\\ &\bar{u}(r,\gamma,0)=\begin{cases}\bar{f}_{2m}(r),\\ \bar{f}_{2m-1}(r),\end{cases}\\ &\Big[\frac{\partial\bar{u}}{\partial r}+h\bar{u}\Big]_{r=a}=0,\\ &|\bar{u}(0,\gamma,t)|<\infty.\end{aligned} \tag{10.1.37}$$

We now seek the integral transform that will eliminate r. In (10.1.37),

$$\mathscr{L}_r\bar{u}\equiv\frac{\partial^2\bar{u}}{\partial r^2}+\frac{1}{r}\frac{\partial\bar{u}}{\partial r}-\frac{m^2}{r^2}\bar{u},$$

so $A_{rr}=1$, $B_r=1/r$, $C=-m^2/r^2$, and the coefficient functions of the Sturm–Liouville problem are

$$R(r)=\exp\Big[-\int\frac{1}{A_{rr}}\Big(\frac{dA_{rr}}{dr}-B_r\Big)dr\Big]=r,\quad P(r)=A_{rr}R(r)=r,\quad Q(r)=-CR(r)=\frac{m^2}{r}.$$

The Sturm–Liouville equation is

$$\frac{d}{dr}\Big(r\frac{d\widetilde{K}}{dr}\Big)-\frac{m^2}{r}\widetilde{K}+\lambda^2 r\widetilde{K}=0,$$

which when expanded becomes the auxiliary problem

$$\begin{aligned}&r^2\frac{d^2\widetilde{K}}{dr^2}+r\frac{d\widetilde{K}}{dr}+(\lambda^2r^2-m^2)\widetilde{K}=0,\\&\Big[\frac{d\widetilde{K}}{dr}+h\widetilde{K}\Big]_{r=a}=0,\\&|\widetilde{K}|_{r=0}<\infty.\end{aligned}\tag{10.1.38}$$

The solution finite at $r=0$ is the Bessel function $J_m(\lambda r)$, and the boundary condition at $r=a$ then determines the eigenvalues λ_n as the positive roots of

$$\lambda_n J_m'(\lambda_n a)+hJ_m(\lambda_n a)=0,\qquad n=1,2,3,\ldots\tag{10.1.39}$$

With $R(r)=r$ the eigenfunctions are

$$\frac{1}{\widetilde{N}}r\widetilde{K}(n,r)=\frac{1}{\widetilde{N}}rJ_m(\lambda_n r),$$

and the squared norm, on using the orthogonality relation for Bessel functions, is

$$\widetilde{N}=\int_0^a r[J_m(\lambda_n r)]^2dr=\frac{a^2}{2}\Big\{J_m'^2(\lambda_n a)+\Big(1-\frac{m^2}{\lambda_n^2a^2}\Big)J_m^2(\lambda_n a)\Big\}.$$

Using (10.1.39) this simplifies to

$$\widetilde{N}=\frac{1}{2\lambda_n^2}\big[h^2a^2+\lambda_n^2a^2-m^2\big]J_m^2(\lambda_n a).$$

The integral transform is therefore

$$\widetilde{u}(n,\gamma,t)=\frac{2\lambda_n^2}{(h^2a^2+\lambda_n^2a^2-m^2)J_m^2(\lambda_n a)}\int_0^a r\,\bar{u}(r,\gamma,t)J_m(\lambda_n r)dr.$$

Applied to (10.1.37), the left-hand side integrates as

$$\int_0^a r\Big(\frac{\partial^2\bar{u}}{\partial r^2}+\frac{1}{r}\frac{\partial\bar{u}}{\partial r}-\frac{m^2}{r^2}\bar{u}\Big)\widetilde{K}\,dr=\Big[\frac{\partial\bar{u}}{\partial r}(r\widetilde{K})-\bar{u}\frac{d}{dr}(r\widetilde{K})+\bar{u}\widetilde{K}\Big]_{r=0}^{a}+\int_0^a\Big(\frac{d^2}{dr^2}(r\widetilde{K})-\frac{d\widetilde{K}}{dr}-\frac{m^2}{r}\widetilde{K}\Big)\bar{u}\,dr.$$

Rearranging the boundary terms,

$$\Big[r\Big(\frac{\partial\bar{u}}{\partial r}\widetilde{K}-\bar{u}\,\frac{\partial\widetilde{K}}{\partial r}\Big)\Big]_{r=0}^{a}=\int_0^a r\Big(\frac{d^2\widetilde{K}}{dr^2}+\frac{1}{r}\,\frac{d\widetilde{K}}{dr}-\frac{m^2}{r^2}\widetilde{K}\Big)\bar{u}\,dr,$$

and the boundary terms vanish by virtue of the boundary conditions in (10.1.37) and (10.1.38). Putting the second term into the bracket form,

$$\frac{1}{\widetilde{N}}\int_0^a r\Big(\frac{\partial^2\bar{u}}{\partial r^2}+\frac{1}{r}\,\frac{\partial\bar{u}}{\partial r}-\frac{m^2}{r^2}\bar{u}\Big)\widetilde{K}\,dr=\frac{1}{\widetilde{N}}\int_0^a r\Big(\frac{d^2\widetilde{K}}{dr^2}+\frac{1}{r}\,\frac{d\widetilde{K}}{dr}-\frac{m^2}{r^2}\widetilde{K}\Big)\bar{u}\,dr=-\frac{\lambda_n^2}{\widetilde{N}}\int_0^a r\widetilde{K}(n,r)\bar{u}(r,\gamma,t)dr=-\lambda_n^2\widetilde{u}(n,\gamma,t).$$

With $\widetilde{f}(\gamma,n)=\frac{1}{\widetilde{N}}\int_0^a r\,\bar{f}(\gamma,r)J_m(\lambda_n r)dr$, the transformed problem reads

$$\frac{d\widetilde{u}}{dt}+k\lambda_n^2\widetilde{u}=0,\qquad \widetilde{u}(n,\gamma,0)=\begin{cases}\widetilde{f}_{2m}(n),\\ \widetilde{f}_{2m-1}(n).\end{cases}$$

Its solution is

$$\widetilde{u}(n,\gamma,t)=Ce^{-k\lambda_n^2 t},$$

and the initial condition gives $C\equiv\widetilde{f}(\gamma,n)$, so

$$\widetilde{u}(n,\gamma,t)=\widetilde{f}(\gamma,n)e^{-k\lambda_n^2 t},\qquad \gamma=\begin{cases}2m\\ 2m-1.\end{cases}$$

Inversion gives

$$\bar{u}(r,\gamma,t)=\sum_{n=1}^{\infty}\widetilde{f}(\gamma,n)e^{-k\lambda_n^2 t}J_m(\lambda_n r),$$

and

$$u(r,\varphi,t)=\sum_{m=0}^{\infty}\big(\bar{u}_{2m}\bar{K}_{2m}+\bar{u}_{2m-1}\bar{K}_{2m-1}\big),$$

so the solution is

$$u(r,\varphi,t)=\sum_{m=0}^{\infty}\sum_{n=1}^{\infty}e^{-k\lambda_n^2 t}J_m(\lambda_n r)\big[\widetilde{f}(2m,n)\cos m\varphi+\widetilde{f}(2m-1,n)\sin m\varphi\big].$$

Exercise VIII.4 (Annular cylindrical pipe).

A homogeneous, infinitely long cylindrical pipe with inner radius a and outer radius b has both its inner and outer surfaces continuously maintained at zero temperature. Given an initial temperature distribution $f(r,\varphi)$ in the body of the pipe, determine the subsequent temperature distribution inside the pipe.

Solution. With the z-axis chosen along the axis of the cylinder, the initial–boundary-value problem is

$$\begin{aligned}
&\frac{\partial^2 u}{\partial r^2}+\frac{1}{r}\frac{\partial u}{\partial r}+\frac{1}{r^2}\frac{\partial^2 u}{\partial \varphi^2}=\frac{1}{k}\frac{\partial u}{\partial t}, \quad a\le r\le b,\ 0\le\varphi\le 2\pi,\ t\ge 0,\\
&u(r,\varphi,0)=f(r,\varphi),\\
&u(a,\varphi,t)=u(b,\varphi,t)=0, \quad |u(r,\varphi,t)|<M.
\end{aligned} \tag{10.1.40}$$

With $\wp_\varphi u\equiv \partial^2 u/\partial\varphi^2$, the auxiliary problem

$$\frac{d^2\bar{K}}{d\varphi^2}+\mu^2\bar{K}=0, \qquad \bar{K}(\mu,0)=\bar{K}(\mu,2\pi)$$

has solutions, with $\mu=m=0,1,2,\ldots$,

$$\bar{K}(m,\varphi)=A\cos m\varphi+B\sin m\varphi.$$

For each value of A and B the periodicity condition is satisfied, and the homogeneous boundary conditions are obeyed for any choice of A and B, so the family of functions admitted as kernels is $\{\cos m\varphi,\sin m\varphi\}$. Choosing $A=B=1$,

$$\bar{K}(m,\varphi)=\cos m\varphi+\sin m\varphi,$$

hence

$$\bar{N}=\int_0^{2\pi}[\bar{K}(m,\varphi)]^2 d\varphi=2\pi,$$

and the integral transform is

$$\bar{u}(r,m,t)=\frac{1}{2\pi}\int_0^{2\pi}u(r,\varphi,t)\bar{K}(m,\varphi)d\varphi,$$

with the inversion formula

$$u(r,\varphi,t)=\sum_{m=0}^{\infty}\bar{u}(r,m,t)\bar{K}(m,\varphi).$$

Multiplying the equation (10.1.40) by r^2 and applying the transform to all terms except $\partial^2 u/\partial\varphi^2$,

$$\frac{1}{\bar{N}}\int_0^{2\pi}\Big(r^2\frac{\partial^2 u}{\partial r^2}+r\frac{\partial u}{\partial r}\Big)\bar{K}(m,\varphi)d\varphi=r^2\frac{\partial^2\bar{u}}{\partial r^2}+r\frac{\partial\bar{u}}{\partial r}.$$

The transform of the second term is, by integration by parts,

$$\frac{1}{\bar{N}}\int_0^{2\pi}\frac{\partial^2 u}{\partial\varphi^2}\bar{K}\,d\varphi=\frac{1}{2\pi}\Big[\frac{\partial u}{\partial\varphi}\bar{K}-u\frac{d\bar{K}}{d\varphi}\Big]_0^{2\pi}+\frac{1}{2\pi}\int_0^{2\pi}u\frac{d^2\bar{K}}{d\varphi^2}\,d\varphi,$$

and the bracket on the right vanishes by virtue of the periodicity, so

$$\frac{1}{\bar{N}}\int_0^{2\pi}\frac{\partial^2 u}{\partial\varphi^2}\bar{K}\,d\varphi=\frac{1}{2\pi}\int_0^{2\pi}u\frac{d^2\bar{K}}{d\varphi^2}d\varphi=-\frac{\mu^2}{\bar{N}}\int_0^{2\pi}u\bar{K}\,d\varphi=-m^2\bar{u}.$$

Therefore the transform of (10.1.40) is

$$\begin{aligned}&\frac{\partial^2\bar{u}}{\partial r^2}+\frac{1}{r}\frac{\partial\bar{u}}{\partial r}-\frac{m^2}{r^2}\bar{u}=\frac{1}{k}\frac{\partial\bar{u}}{\partial t},\quad a\le r\le b,\ t>0,\\&\bar{u}(r,m,0)=\bar{f}(r,m),\\&\bar{u}(a,m,t)=\bar{u}(b,m,t)=0,\quad |\bar{u}(r,m,t)|<M.\end{aligned}\tag{10.1.41}$$

The auxiliary problem for the integral transform that will eliminate r is constructed from

$$\mathscr{O}_r\equiv\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}-\frac{m^2}{r^2},$$

so $A_{rr}=1$, $B_r=1/r$, $C=-m^2/r^2$, and the coefficient functions are

$$R(r)=\exp\Big[-\int\frac{1}{A_{rr}}\Big(\frac{dA_{rr}}{dr}-B_r\Big)dr\Big]=r,\quad P(r)=A_{rr}R(r)=r,\quad Q(r)=-CR(r)=\frac{m^2}{r}.$$

The auxiliary problem is therefore

$$\begin{aligned}&\frac{d}{dr}\Big[r\frac{d\hat{K}}{dr}\Big]+\Big(\lambda^2 r-\frac{m^2}{r}\Big)\hat{K}=0,\\&[\hat{K}]_{r=a}=[\hat{K}]_{r=b}=0,\quad |\hat{K}|<\infty.\end{aligned}\tag{10.1.42}$$

With D and E constants, the general solution is

$$\hat{K}=DJ_m(\lambda r)+EY_m(\lambda r),$$

which is finite throughout the body of the pipe since $r=0$, the z-axis, is not part of the domain. Applying the boundary conditions (10.1.42),

$$DJ_m(\lambda a)+EY_m(\lambda a)=0,\quad DJ_m(\lambda b)+EY_m(\lambda b)=0.\tag{10.1.43}$$

Non-trivial solutions D,E exist only when the determinant vanishes,

$$\begin{vmatrix}J_m(\lambda a) & Y_m(\lambda a)\\ J_m(\lambda b) & Y_m(\lambda b)\end{vmatrix}=J_m(\lambda a)Y_m(\lambda b)-J_m(\lambda b)Y_m(\lambda a)=0,\tag{10.1.44}$$

which determines the eigenvalues λ_n $(n=1,2,\ldots)$. From (10.1.43) we may write

$$\frac{D}{Y_m(\lambda_n r)}=-\frac{E}{J_m(\lambda_n b)}=\text{arbitrary constant},$$

and absorbing this arbitrary constant into the eigenfunction we obtain

$$\hat{K}=\big[Y_m(\lambda_n b)J_m(\lambda_n r)-J_m(\lambda_n b)Y_m(\lambda_n r)\big],$$

together with

$$\hat{N}\equiv\int_a^b r[\hat{K}]^2dr=\frac{2}{\pi^2}\frac{J_m^2(\lambda_n a)-J_m^2(\lambda_n b)}{J_m^2(\lambda_n a)}.$$

Hence the integral transform is

$$\hat{u}(n,m,t) = \frac{1}{\hat{N}} \int_a^b r\bar{u}(r,m,t)\hat{K}(n,r)dr, \tag{10.1.45}$$

with inversion formula

$$\bar{u}(r,m,t) = \sum_{n=1}^{\infty} \hat{u}(n,m,t)\hat{K}(n,r).$$

Returning to the problem (10.1.41), the sum of the first two terms of the equation can be written as

$$\frac{\partial^2 \bar{u}}{\partial r^2} + \frac{1}{r}\frac{\partial \bar{u}}{\partial r} = \frac{1}{r}\frac{\partial}{\partial r}\Big(r\frac{\partial \bar{u}}{\partial r}\Big),$$

and the transform (10.1.45) applied to that term gives

$$\frac{1}{\hat{N}}\int_a^b \Big(r\frac{\partial}{\partial r}\Big(r\frac{\partial \bar{u}}{\partial r}\Big)\Big)\frac{1}{r}\hat{K}\,dr = \frac{1}{\hat{N}}\Big[r\Big(\frac{\partial \bar{u}}{\partial r}\hat{K} - \bar{u}\frac{d\hat{K}}{dr}\Big)\Big]_a^b + \frac{1}{\hat{N}}\int_a^b \bar{u}\frac{d}{dr}\Big(r\frac{d\hat{K}}{dr}\Big)dr.$$

The boundary terms vanish by (10.1.41), and on using (10.1.42),

$$\frac{1}{\hat{N}}\int_a^b \Big(\frac{1}{r}\frac{\partial}{\partial r}\Big(r\frac{\partial \bar{u}}{\partial r}\Big) - \frac{m^2}{r^2}\bar{u}\Big)r\hat{K}\,dr = \frac{1}{\hat{N}}\int_a^b \bar{u}\Big[\frac{d}{dr}\Big(r\frac{d\hat{K}}{dr}\Big) - \frac{m^2}{r}\hat{K}\Big]dr = -\frac{\lambda_n^2}{\hat{N}}\int_a^b r\bar{u}\hat{K}\,dr = -\lambda_n^2\hat{u}(n,m,t).$$

Hence the transformed problem is

$$\frac{d\hat{u}}{dt} + k\lambda_n^2\hat{u} = 0, \qquad \hat{u}(n,m,0) = \hat{f}(n,m).$$

The solution of this elementary initial-value problem is, with G a constant,

$$\hat{u}(n,m,t) = Ge^{-k\lambda_n^2 t},$$

and the initial condition gives $G \equiv \hat{f}(n,m)$. Hence

$$\bar{u}(r,m,t) = \sum_{n=1}^{\infty} \hat{u}(n,m,t)\hat{K}(n,r) = \sum_{n=1}^{\infty} \hat{f}(n,m)e^{-k\lambda_n^2 t}\big[Y_m(\lambda_n b)J_m(\lambda_n r) - J_m(\lambda_n b)Y_m(\lambda_n r)\big],$$

and the solution sought is

$$u(r,\varphi,t) = \sum_{m=0}^{\infty}\sum_{n=1}^{\infty} \hat{f}(n,m)e^{-k\lambda_n^2 t}\big[Y_m(\lambda_n b)J_m(\lambda_n r) - J_m(\lambda_n b)Y_m(\lambda_n r)\big]\big[\cos m\varphi + \sin m\varphi\big].$$

Exercise VIII.5 (Sphere).

A homogeneous sphere of radius a has its surface continuously maintained at zero temperature. Given an initial temperature distribution $f(r,\theta,\varphi)$, find the temperature distribution inside the sphere.

Solution. With the origin of the coordinate system at the centre of the sphere and the change of variable $\mu = \cos\theta$, the relevant initial–boundary-value problem is

$$\begin{aligned}
&\frac{\partial^2 u}{\partial r^2} + \frac{2}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial}{\partial \mu}\Big[(1-\mu^2)\frac{\partial u}{\partial \mu}\Big] + \frac{1}{r^2(1-\mu^2)}\frac{\partial^2 u}{\partial \varphi^2} = \frac{1}{k}\frac{\partial u}{\partial t},\\
&u(r,\mu,\varphi,0) = f(r,\mu,0),\\
&u(a,\mu,\varphi,t) = 0, \quad u(r,\mu,0,t) = u(r,\mu,2\pi,t),\\
&|u(r,\mu,\varphi,t)| < \infty.
\end{aligned} \tag{10.1.46}$$

The integral transform that eliminates φ has, with $\mathscr{O}_\varphi u \equiv \partial^2 u/\partial\varphi^2$, the auxiliary problem

$$\frac{d^2\bar{K}}{d\varphi^2} + \beta^2\bar{K} = 0, \qquad [\bar{K}]_{\varphi=0} = [\bar{K}]_{\varphi=2\pi},$$

which is solved as in the previous problem. With $\beta = m = 0,1,2,\ldots$ the kernel is

$$\bar{K}(m,\varphi) = \cos m\varphi + \sin m\varphi,$$

$$\bar{N} = \int_0^{2\pi} [\bar{K}(m,\varphi)]^2 d\varphi = 2\pi,$$

and the transform is

$$\bar{u}(r,\mu,m,t) = \frac{1}{2\pi}\int_0^{2\pi} u(r,\mu,\varphi,t)\bar{K}(m,\varphi)d\varphi, \tag{10.1.47}$$

with inverse

$$\bar{u}(r,\mu,\varphi,t) = \sum_{m=0}^{\infty} \bar{u}(r,\mu,m,t)\bar{K}(m,\varphi).$$

Multiplying the equation (10.1.46) by $r^2(1-\mu^2)$ and applying the transform (10.1.47) to all terms except $\partial^2 u/\partial\varphi^2$ gives

$$\frac{1}{\bar{N}}\int_0^{2\pi}\Big\{r^2(1-\mu^2)\Big(\frac{\partial^2 u}{\partial r^2} + \frac{2}{r}\frac{\partial u}{\partial r}\Big) + (1-\mu^2)\frac{\partial}{\partial\mu}\Big[(1-\mu^2)\frac{\partial u}{\partial\mu}\Big]\Big\}\bar{K}(m,\varphi)d\varphi = r^2(1-\mu^2)\Big(\frac{\partial^2\bar{u}}{\partial r^2} + \frac{2}{r}\frac{\partial\bar{u}}{\partial r}\Big) + (1-\mu^2)\frac{\partial}{\partial\mu}\Big[(1-\mu^2$$

The transform of $\partial^2 u/\partial\varphi^2$, with the periodicity used in the integration by parts, is

$$\frac{1}{\bar{N}}\int_0^{2\pi}\frac{\partial^2 u}{\partial\varphi^2}\bar{K}(m,\varphi)d\varphi = -m^2\bar{u}(r,\mu,m,t).$$

Hence the transform of (10.1.46) is

$$\begin{aligned}
&\frac{\partial^2\bar{u}}{\partial r^2}+\frac{2}{r}\frac{\partial\bar{u}}{\partial r}+\frac{1}{r^2}\frac{\partial}{\partial\mu}\Big[(1-\mu^2)\frac{\partial\bar{u}}{\partial\mu}\Big]-\frac{m^2}{r^2(1-\mu^2)}\bar{u}=\frac{1}{k}\frac{\partial\bar{u}}{\partial t},\\
&\bar{u}(r,\mu,m,0)=\bar{f}(r,\mu,m),\\
&\bar{u}(a,\mu,m,t)=0,\quad |\bar{u}(r,\mu,m,t)|<\infty.
\end{aligned}\tag{10.1.48}$$

The integral transform that eliminates μ is determined from

$$\wp_\mu\bar{u}\equiv\frac{\partial}{\partial\mu}\Big[(1-\mu^2)\frac{\partial\bar{u}}{\partial\mu}\Big]-\frac{m^2}{1-\mu^2}\bar{u},$$

so $A_{\mu\mu}=1-\mu^2$, $B_\mu=-2\mu$, $C=-m^2/(1-\mu^2)$, and the coefficient functions of the Sturm–Liouville problem are

$$R(\mu)=\exp\Big[-\int\frac{1}{A_{\mu\mu}}\Big(\frac{dA_{\mu\mu}}{d\mu}-B_\mu\Big)d\mu\Big]=1,\quad P(\mu)=A_{\mu\mu}R(\mu)=1-\mu^2,\quad Q(\mu)=-CR(\mu)=\frac{m^2}{1-\mu^2}.$$

The auxiliary problem is therefore

$$\frac{\partial}{\partial\mu}\Big[(1-\mu^2)\frac{d\widetilde{K}}{d\mu}\Big]+\Big[\beta^2-\frac{m^2}{1-\mu^2}\Big]\widetilde{K}=0,\qquad |\widetilde{K}(r,\mu,m,t)|<\infty,$$

which is the associated Legendre equation. Its independent solutions finite on the z-axis ($\theta=0,\pi$ i.e. $\mu=\mp1$) when $\beta^2=n(n+1)$, with n an integer, are the associated Legendre polynomials $P_n^m(\mu)$. The eigenfunctions of this auxiliary problem are therefore $P_n^m(\mu)$ for $\beta_n^2=n(n+1)$, $n=0,1,2,\ldots$. Using the orthogonality of associated Legendre polynomials, the normalisation is

$$\widetilde{N}=\int_{-1}^{+1}[P_n^m(\mu)]^2d\mu=\frac{2}{2n+1}\frac{(n+m)!}{(n-m)!},$$

and the integral transform on $[-1,+1]$ is

$$\widetilde{u}(r,n,m,t)=\frac{2n+1}{2}\frac{(n-m)!}{(n+m)!}\int_{-1}^{+1}\bar{u}(r,\mu,m,t)P_n^m(\mu)d\mu.\tag{10.1.49}$$

Applied to (10.1.48),

$$\frac{1}{\widetilde{N}}\int_{-1}^{+1}\Big\{\frac{\partial}{\partial\mu}\Big[(1-\mu^2)\frac{\partial\bar{u}}{\partial\mu}\Big]-\frac{m^2}{1-\mu^2}\bar{u}\Big\}P_n^m d\mu=\frac{1}{\widetilde{N}}\int_{-1}^{+1}\bar{u}\Big\{\frac{d}{d\mu}\Big[(1-\mu^2)\frac{dP_n^m}{d\mu}\Big]-\frac{m^2}{1-\mu^2}P_n^m\Big\}d\mu=-\frac{\lambda_n^2}{\widetilde{N}}\int_{-1}^{+1}\bar{u}P_n^m d$$

in which an integration by parts that vanishes at the endpoints $\mu=\pm1$ has been used. The transformed problem is therefore

$$\begin{aligned}
&\frac{\partial^2\widetilde{u}}{\partial r^2}+\frac{2}{r}\frac{\partial\widetilde{u}}{\partial r}-\frac{n(n+1)}{r^2}\widetilde{u}=\frac{1}{k}\frac{\partial\widetilde{u}}{\partial t},\\
&\widetilde{u}(r,n,m,0)=\widetilde{f}(r,n,m),\\
&\widetilde{u}(a,n,m,t)=0,\quad |\widetilde{u}(r,n,m,t)|<\infty.
\end{aligned}\tag{10.1.50}$$

It remains to find the integral transform that eliminates r. We first transform the problem (10.1.50) by setting

$$\widetilde{u} = \frac{\widetilde{v}}{r^{1/2}}, \tag{10.1.51}$$

so that

$$\frac{\partial^2 \widetilde{v}}{\partial r^2} + \frac{1}{r}\frac{\partial \widetilde{v}}{\partial r} - \frac{(n+1/2)^2}{r^2}\widetilde{v} = \frac{1}{k}\frac{\partial \widetilde{v}}{\partial t},$$
$$\widetilde{v}(r,n,m,0) = r^{1/2}\widetilde{f}(r,n,m), \quad \widetilde{v}(a,n,m,t) = 0, \quad |\widetilde{v}(r,n,m,t)| < \infty. \tag{10.1.52}$$

We now seek the integral transform that eliminates r. In the operator

$$\mathscr{O}_r \widetilde{v} \equiv \frac{\partial^2 \widetilde{v}}{\partial r^2} + \frac{1}{r}\frac{\partial \widetilde{v}}{\partial r} - \frac{(n+1/2)^2}{r^2}\widetilde{v},$$

we have $A_{rr} = 1$, $B_r = 1/r$, $C = -(n+1/2)^2/r^2$, and the coefficient functions of the corresponding Sturm–Liouville problem are

$$R(r) = r, \quad P(r) = r, \quad Q(r) = \frac{(n+1/2)^2}{r}.$$

The auxiliary problem is

$$\begin{aligned} &r^2\frac{d^2\hat{K}}{dr^2} + r\frac{d\hat{K}}{dr} + \left[\lambda^2 r^2 - (n+1/2)^2\right]\hat{K} = 0, \\ &[\hat{K}]_{r=a} = 0, \quad |\hat{K}| < \infty, \end{aligned} \tag{10.1.53}$$

whose finite solutions at $r = 0$ are the half-integer Bessel functions $J_{n+1/2}(\lambda r)$. The boundary condition at $r = a$ in (10.1.52) fixes the eigenvalues λ_s ($s = 1, 2, \ldots$) as the positive roots of

$$J_{n+1/2}(\lambda a) = 0.$$

The squared norm is

$$\hat{N} = \int_0^a r[J_{n+1/2}(\lambda_s r)]^2 dr = \frac{a^2}{2}[J'_{n+1/2}(\lambda_s a)]^2,$$

hence the integral transform is

$$\hat{v}(r,n,m,t) = \frac{2}{a^2[J'_{n+1/2}(\lambda_s a)]^2}\int_0^a r\widetilde{v}(r,n,m,t)J_{n+1/2}(\lambda_s r)dr. \tag{10.1.54}$$

Applied to (10.1.52),

$$\hat{f}(s,n,m) = \frac{2}{a^2[J'_{n+1/2}(\lambda_s a)]^2}\int_0^a r^{3/2}\widetilde{f}(r,n,m)J_{n+1/2}(\lambda_s r)dr, \tag{10.1.55}$$

the transformed problem is

$$\frac{d\hat{v}}{dt} + k\lambda_s^2 \hat{v} = 0, \qquad \hat{v}(s,n,m,0) = \hat{f}(s,n,m).$$

Its solution is

$$\hat{v}(s,n,m,t) = \hat{f}(s,n,m)e^{-k\lambda_s^2 t},$$

and the inversion gives

$$\widetilde{v}(r,n,m,t) = \sum_{s=1}^{\infty} \hat{f}(s,n,m)e^{-k\lambda_s^2 t} J_{n+1/2}(\lambda_s r),$$

and substitution into the transformation (10.1.51) yields

$$\widetilde{u}(r,n,m,t) = \sum_{s=1}^{\infty} \hat{f}(s,n,m)e^{-k\lambda_s^2 t} \frac{J_{n+1/2}(\lambda_s r)}{r^{1/2}}. \tag{10.1.56}$$

The solution of (10.1.50) so obtained is finite at $r = 0$ for every point inside the sphere because $J_{n+1/2} \propto (r/2)^{1/2} J_n$ near the origin. Now the inverse of the transformation (10.1.49) gives

$$\bar{u}(r,\mu,m,t) = \sum_{n=0}^{\infty} \widetilde{u}(r,n,m,t) P_n^m(\mu),$$

and using (10.1.56),

$$\bar{u}(r,\mu,m,t) = \sum_{n=0}^{\infty} \sum_{s=1}^{\infty} \hat{f}(s,n,m)e^{-k\lambda_s^2 t} \frac{J_{n+1/2}(\lambda_s r)}{r^{1/2}} P_n^m(\mu).$$

The inverse of (10.1.47) then gives the solution of (10.1.46) as

$$u(r,\theta,\varphi,t) = \sum_{n=0}^{\infty} \sum_{s=1}^{\infty} \sum_{m=0}^{\infty} \hat{f}(s,n,m)e^{-k\lambda_s^2 t} \frac{J_{n+1/2}(\lambda_s r)}{r^{1/2}} P_n^m(\cos\theta)\big[\cos m\varphi + \sin m\varphi\big],$$

in which $\hat{f}(s,n,m)$, by (10.1.47), (10.1.49) and (10.1.55), is given in terms of $f(r,\mu,\varphi)$ by

$$\hat{f}(s,n,m) = \frac{2n+1}{\pi a^2} \frac{(n-m)!}{(n+m)!} \frac{1}{[J'_{n+1/2}(\lambda_s a)]^2} \int_0^a \int_{-1}^{+1} \int_0^{2\pi} r^{3/2} f(r,\mu,\varphi) \bar{K}(m,\varphi) P_n^m(\mu) J_{n+1/2}(\lambda_s r) dr\, d\mu\, d\varphi.$$

Problems

VIII.1.

Solve the initial-value problem

$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=0, \qquad 0\le x\le a,\ 0\le y\le b,$$

$$u(0,y)=u(a,y)=0, \quad u(x,b)=u(x,0)=0,$$

first by a single integral transform and then by a double integral transform, and show that the two solutions are identical.

VIII.2.

Solve the following initial-value problems by the finite integral transform method.

(a)

$$\frac{\partial^2 u}{\partial r^2}+\frac{2}{r}\frac{\partial u}{\partial r}=\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2}, \qquad 0\le r\le 1,\ t\ge 0,$$

$$u(r,0)=f(r), \quad \left(\frac{\partial u}{\partial t}\right)_{t=0}=g(r), \quad \left(\frac{\partial u}{\partial r}\right)_{r=a}=0, \quad |u(r,t)|<M.$$

(b)

$$\frac{\partial u}{\partial t}=k\Big(\frac{\partial^2 u}{\partial \rho^2}+\frac{1}{\rho}\frac{\partial u}{\partial \rho}\Big), \qquad 0\le \rho\le 1,\ t\ge 0,$$

$$u(1,t)=0, \quad u(\rho,0)=F(\rho), \quad |u(\rho,t)|<M.$$

(c)

$$\frac{\partial u}{\partial t}=k\Big(\frac{\partial^2 u}{\partial \rho^2}+\frac{1}{\rho}\frac{\partial u}{\partial \rho}+\frac{\partial^2 u}{\partial z^2}\Big), \qquad 0\le \rho\le 1,\ 0\le z\le 1,\ t\ge 0,$$

$$u(\rho,z,0)=f(\rho,z), \quad u(\rho,0,t)=u(\rho,1,t)=0, \quad u(1,z,t)=0, \quad |u(\rho,z,t)|<M.$$

VIII.3.

A homogeneous solid in the form of a rectangular parallelepiped of edge lengths a, b, c is held with all of its faces continuously maintained at zero temperature, and at $t=0$ has temperature distribution $F(x,y,z)$. Heat is generated inside the solid at the rate $g(x,y,z,t)$ for $t>0$. Determine the temperature distribution inside the solid by means of the finite integral transform method, given the initial–boundary-value problem

$$\frac{\partial u(\mathbf{r},t)}{\partial t}-k\nabla^2 u(\mathbf{r},t)=\beta g(\mathbf{r},t),$$

$$u(0,y,z,t)=u(a,y,z,t)=0, \quad u(x,0,z,t)=u(x,b,z,t)=0, \quad u(x,y,0,t)=u(x,y,c,t)=0,$$

$$u(x,y,z,0)=F(x,y,z).$$

VIII.4.

A homogeneous slab of width L, as shown in the figure, has its boundaries $x = 0$ and $x = L$ losing heat by convection (Newton's law of cooling) to a time-dependent external medium. The initial temperature is $F(x)$, and a heat source $g(x,t)$ is present in the slab for $t > 0$. The boundary and initial conditions for the slab are uniform; end heat-losses are neglected. Determine the temperature distribution by the finite integral transform method, given

$$\frac{\partial u}{\partial t} - k\frac{\partial^2 u}{\partial x^2} = \beta g(x,t), \quad 0 \le x \le L,\ t > 0,$$

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

$$\left[h_1 u - k_1 \frac{\partial u}{\partial x}\right]_{x=0} = f_1(t), \quad \left[h_2 u + k_2 \frac{\partial u}{\partial x}\right]_{x=L} = f_2(t).$$

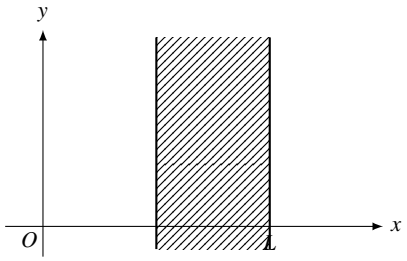


Fig. 10.2 Slab of width L between $x = 0$ and $x = L$.

VIII.5.

A homogeneous solid in the form of a rectangular parallelepiped of edge lengths a, b, c has its face $z = 0$ maintained at the constant temperature u_0 while its other faces lose heat by convection to an external medium at zero temperature. Determine the steady-state temperature distribution inside the solid by the finite integral transform method, given

$$\nabla^2 u(\mathbf{r}) = 0, \qquad 0 \le x \le a,\ 0 \le y \le b,\ 0 \le z \le c,$$

$$\left[\frac{\partial u}{\partial x} - h_1 u\right]_{x=0} = \left[\frac{\partial u}{\partial x} + h_2 u\right]_{x=a} = 0,$$

$$\left[\frac{\partial u}{\partial y} - h_1 u\right]_{y=0} = \left[\frac{\partial u}{\partial y} + h_2 u\right]_{y=b} = 0,$$

$$\left[\frac{\partial u}{\partial z} + h_2 u\right]_{z=c} = 0, \qquad u(x,y,0) = u_0.$$

VIII.6.

A homogeneous solid cylindrical rod of radius $r = a$ has its lateral surface losing heat by convection to a time-dependent external medium. The initial temperature distribution along the rod is $F(r)$. The end heat-losses of the rod are neglected, and the boundary and initial conditions along the rod are uniform. A heat source $g(r,t)$ is present in the body of the cylinder for $t > 0$. Determine the temperature distribution inside the rod by the finite integral transform method, given

$$\frac{\partial u(r,t)}{\partial t} - \alpha \nabla^2 u(r,t) = \lambda g(r,t), \quad 0 \le r \le a,\ t > 0,$$

$$u(r,0) = F(r),$$

$$\left[k\frac{\partial u}{\partial t} + hu\right]_{r=a} = f(t), \quad |u(r,t)| < \infty.$$

Chapter 11
Green's Functions and Fundamental Solutions

Abstract The Green's function is the analytic device by which an inhomogeneous linear PDE with prescribed auxiliary conditions is solved through a single convolution integral. We introduce the Dirac delta and its derivatives, construct the fundamental solutions of the three canonical operators of mathematical physics, derive the Green's-function representation of the Dirichlet problem, and exhibit the spectral expansion of the Green's function in terms of Sturm–Liouville eigenfunctions. Worked exercises and end-of-chapter problems illustrate the method on the heat, wave and Laplace equations.

11.1 The Dirac Delta

The notion of a point source—an idealised concentration of mass, charge or heat at a single point of space—is, strictly speaking, incompatible with the classical definition of a function. No ordinary function $\delta(x)$ of a real variable x vanishes identically except at $x=0$ and yet has unit total integral; the demand

$$\delta(x)=0\ (x\neq 0), \qquad \int_{-\infty}^{\infty}\delta(x)\,dx=1$$

exceeds the framework of Riemann or Lebesgue integration. Nevertheless, the formal manipulation of such an object is so productive that it has become an indispensable tool of mathematical physics, and its rigorous foundation—the theory of *distributions* due to Schwartz—was developed precisely to legitimise the calculations. Throughout this chapter we shall use the delta in the formal sense, trusting the reader to interpret integrals containing it as limits of sequences of ordinary integrals.

Sequential definition. Let $\eta_\varepsilon(x)$ $(\varepsilon>0)$ be a family of non-negative continuous functions with the properties

$$\int_{-\infty}^{\infty}\eta_\varepsilon(x)\,dx=1 \qquad \text{for every } \varepsilon>0, \tag{11.1.1}$$

$$\eta_\varepsilon(x) \longrightarrow 0 \qquad \text{as } \varepsilon \to 0 \quad \text{for every } x \neq 0, \tag{11.1.2}$$

the convergence in (11.1.2) being uniform on every closed set not containing the origin. The classical example is the Gaussian

$$\eta_\varepsilon(x) = \frac{1}{\sqrt{4\pi\varepsilon}} \exp\!\left(-\frac{x^2}{4\varepsilon}\right); \tag{11.1.3}$$

others include rectangular pulses, Lorentzian profiles and triangular spikes, all of unit area concentrated near the origin (Fig. 11.1). For every continuous test function $\varphi(x)$ that decays sufficiently rapidly at infinity,

$$\lim_{\varepsilon \to 0} \int_{-\infty}^{\infty} \eta_\varepsilon(x)\, \varphi(x)\, dx = \varphi(0). \tag{11.1.4}$$

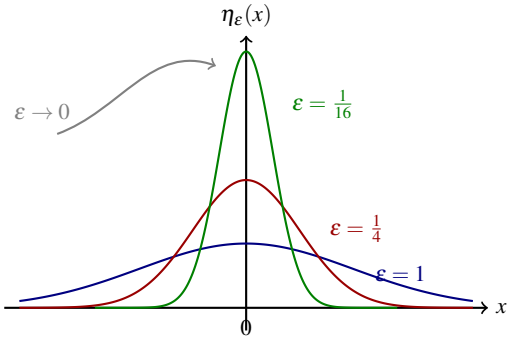


Fig. 11.1 The Gaussian family $\eta_\varepsilon(x)$ of (11.1.3), plotted for three successively smaller values of ε. As $\varepsilon \to 0$ each profile narrows and grows, its total area remaining equal to 1. The family converges, in the distributional sense of (11.1.4), to the Dirac delta.

The delta is then *defined* by (11.1.4): the symbol $\delta(x)$ stands for whatever object yields, when integrated against a test function, the value of that function at the origin. We write, formally,

$$\int_{-\infty}^{\infty} \delta(x)\, \varphi(x)\, dx = \varphi(0), \tag{11.1.5}$$

and, more generally,

$$\int_{-\infty}^{\infty} \delta(x - x_0)\, \varphi(x)\, dx = \varphi(x_0). \tag{11.1.6}$$

The relation (11.1.6) is sometimes called the *sifting property* of the delta; it expresses the role of $\delta(x - x_0)$ as a continuous analogue of the Kronecker symbol δ_{ij}.

Elementary identities. The formal calculus of the delta rests on a small number of identities, each established through (11.1.5).

- *Symmetry*: $\delta(-x) = \delta(x)$.
- *Scaling*: $\delta(ax) = |a|^{-1}\, \delta(x)$ for $a \neq 0$.

- *Composition with smooth f having simple zeros $x_1,\dots,x_n$*:

$$\delta\big(f(x)\big) = \sum_{k=1}^{n} \frac{\delta(x-x_k)}{|f'(x_k)|}. \tag{11.1.7}$$

- *Multiplication by a smooth function*: $f(x)\,\delta(x-x_0) = f(x_0)\,\delta(x-x_0)$.

The derivative $\delta'(x)$ is defined by integration by parts in the formal expression (11.1.5): for every test function φ,

$$\int_{-\infty}^{\infty} \delta'(x)\,\varphi(x)\,dx = -\varphi'(0). \tag{11.1.8}$$

Higher derivatives are introduced analogously,

$$\int_{-\infty}^{\infty} \delta^{(n)}(x)\,\varphi(x)\,dx = (-1)^n\,\varphi^{(n)}(0). \tag{11.1.9}$$

The Heaviside step function

$$H(x) = \begin{cases} 0, & x < 0, \\ 1, & x \geq 0, \end{cases}$$

satisfies, in this same formal sense,

$$H'(x) = \delta(x), \tag{11.1.10}$$

a relation that legitimises differentiating discontinuous functions and which will be used repeatedly below.

Multidimensional delta. For $\mathbf{r} = (x_1,\dots,x_n) \in \mathbb{R}^n$ we set

$$\delta(\mathbf{r}) \equiv \delta(x_1)\,\delta(x_2)\cdots\delta(x_n), \tag{11.1.11}$$

the product being interpreted in the same formal sense. The sifting property reads

$$\int_{\mathbb{R}^n} \delta(\mathbf{r}-\mathbf{r}_0)\,\varphi(\mathbf{r})\,dV = \varphi(\mathbf{r}_0), \tag{11.1.12}$$

and the change-of-variables formula in spherical coordinates yields the useful identity, for $\mathbf{r} \in \mathbb{R}^3$,

$$\delta(\mathbf{r}) = \frac{\delta(r)}{4\pi r^2}, \tag{11.1.13}$$

valid as a distribution on smooth radial test functions on $\mathbb{R}^3$.

11.2 Fundamental Solutions

Let L denote a constant-coefficient linear differential operator acting on functions $u(\mathbf{r})$ on $\mathbb{R}^n$.

Definition 11.1 (Fundamental solution) A *fundamental solution* of L is a function (or distribution) $E(\mathbf{r})$ satisfying

$$L[E(\mathbf{r})] = \delta(\mathbf{r}) \tag{11.2.1}$$

in the sense of distributions.

The importance of E rests on the following observation. If E satisfies (11.2.1) and $f(\mathbf{r})$ is a sufficiently well-behaved function on $\mathbb{R}^n$, the function

$$u(\mathbf{r}) = \int_{\mathbb{R}^n} E(\mathbf{r}-\mathbf{r}')\, f(\mathbf{r}')\, dV' \equiv (E * f)(\mathbf{r}) \tag{11.2.2}$$

formally satisfies $L[u] = f$, since

$$L[u] = \int L[E(\mathbf{r}-\mathbf{r}')]\, f(\mathbf{r}')\, dV' = \int \delta(\mathbf{r}-\mathbf{r}')\, f(\mathbf{r}')\, dV' = f(\mathbf{r}).$$

The fundamental solution thus inverts L by convolution; the free-space inhomogeneous problem $L[u] = f$ is reduced to the computation of an integral. The construction of E, however, is non-trivial, and depends sensitively on the operator. We compute the fundamental solutions of the three canonical operators of mathematical physics in turn.

The Laplacian. We seek a function $E(\mathbf{r})$ on $\mathbb{R}^3$ satisfying $\nabla^2 E = \delta(\mathbf{r})$ in the sense of distributions. The right-hand side is supported only at the origin, so E must be harmonic on $\mathbb{R}^3 \setminus \{0\}$; the rotational symmetry of the equation suggests that E is radial, $E = E(r)$ with $r = |\mathbf{r}|$. The Laplacian in spherical coordinates (1.1.7) acting on a radial function reduces to

$$\nabla^2 E(r) = \frac{1}{r^2}\frac{d}{dr}\left(r^2 \frac{dE}{dr}\right),$$

and setting this to zero gives, after one integration, $r^2 E'(r) = A$, hence $E(r) = -A/r + B$. The constant B may be set to zero (a constant is annihilated by ∇^2); the constant A is fixed by the requirement that $\nabla^2 E$ integrates to 1 over any neighbourhood of the origin. Apply Gauss's theorem (1.1.8) to the ball $|\mathbf{r}| < \rho$:

$$1 = \int_{|\mathbf{r}|<\rho} \nabla^2 E\, dV = \oint_{|\mathbf{r}|=\rho} \frac{\partial E}{\partial n}\, dS = \oint_{|\mathbf{r}|=\rho} \frac{A}{r^2}\, dS = 4\pi A,$$

so $A = 1/(4\pi)$. The fundamental solution of ∇^2 on $\mathbb{R}^3$ is therefore

$$\boxed{E(\mathbf{r}) = -\frac{1}{4\pi r}, \qquad r = |\mathbf{r}|.} \tag{11.2.3}$$

The corresponding solution of Poisson's equation $\nabla^2 u = -\rho/\varepsilon_0$ in $\mathbb{R}^3$ is, in view of (11.2.2),

$$u(\mathbf{r}) = \frac{1}{4\pi\varepsilon_0}\int_{\mathbb{R}^3}\frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,dV',$$

which is the Coulomb potential of a charge distribution ρ. The fundamental solution of the two-dimensional Laplacian, computed in exactly the same way, is

$$E_2(\mathbf{r}) = \frac{1}{2\pi}\ln r, \qquad \mathbf{r}\in\mathbb{R}^2, \tag{11.2.4}$$

and that of the n-dimensional Laplacian for $n \geq 3$,

$$E_n(\mathbf{r}) = -\frac{1}{(n-2)\,\omega_n\, r^{n-2}}, \tag{11.2.5}$$

where $\omega_n = 2\pi^{n/2}/\Gamma(n/2)$ is the surface area of the unit sphere in $\mathbb{R}^n$.

The heat operator. We seek $E(\mathbf{r},t)$ on $\mathbb{R}^n\times\mathbb{R}$ satisfying

$$\partial_t E - \alpha\nabla^2 E = \delta(\mathbf{r})\,\delta(t), \tag{11.2.6}$$

with $E\equiv 0$ for $t<0$ (a causality requirement). The spatial Fourier transform $\hat{E}(\boldsymbol{\xi},t) = \int E(\mathbf{r},t)\,e^{-i\boldsymbol{\xi}\cdot\mathbf{r}}\,dV$ converts the Laplacian into multiplication by $-|\boldsymbol{\xi}|^2$, and (11.2.6) becomes

$$\partial_t\hat{E} + \alpha|\boldsymbol{\xi}|^2\hat{E} = \delta(t),$$

the unique causal solution of which is, in view of (11.1.10),

$$\hat{E}(\boldsymbol{\xi},t) = H(t)\,e^{-\alpha|\boldsymbol{\xi}|^2 t}.$$

Inversion of the Fourier transform gives the celebrated *Gaussian heat kernel*

$$\boxed{E(\mathbf{r},t) = \frac{H(t)}{(4\pi\alpha t)^{n/2}}\exp\!\left(-\frac{|\mathbf{r}|^2}{4\alpha t}\right).} \tag{11.2.7}$$

The factor $H(t)$ encodes causality: heat does not propagate backwards in time. As $t\to 0^+$ the kernel concentrates at the origin and acts as a delta sequence in the sense of (11.1.4). For an initial-value problem $u_t = \alpha\nabla^2 u$ on $\mathbb{R}^n$ with $u(\mathbf{r},0) = f(\mathbf{r})$, the solution is therefore

$$u(\mathbf{r},t) = \frac{1}{(4\pi\alpha t)^{n/2}}\int_{\mathbb{R}^n}\exp\!\left(-\frac{|\mathbf{r}-\mathbf{r}'|^2}{4\alpha t}\right)f(\mathbf{r}')\,dV' \qquad (t>0). \tag{11.2.8}$$

The integral (11.2.8) expresses the smoothing action of the heat equation: even an initial datum with discontinuities is, after arbitrarily small time, infinitely differentiable. This regularising property, characteristic of the parabolic type, is here made quantitative.

The wave operator. The construction for the wave operator $\partial_t^2 - c^2\nabla^2$ is more delicate, and the result depends qualitatively on the spatial dimension. We sketch it for

$n = 3$, the case in which the answer is most familiar and physically the most transparent.

We seek $E(\mathbf{r},t)$ on $\mathbb{R}^3 \times \mathbb{R}$ vanishing for $t < 0$ and satisfying

$$\partial_t^2 E - c^2 \nabla^2 E = \delta(\mathbf{r})\,\delta(t). \tag{11.2.9}$$

Spatial Fourier transformation yields

$$\partial_t^2 \hat{E} + c^2 |\xi|^2 \hat{E} = \delta(t),$$

the causal solution of which is

$$\hat{E}(\xi,t) = H(t)\,\frac{\sin(c|\xi|t)}{c|\xi|}.$$

The inverse Fourier transform of $\sin(c|\xi|t)/(c|\xi|)$ in three dimensions can be evaluated by writing the integral in spherical coordinates and using the residue theorem; the result is

$$\boxed{E(\mathbf{r},t) = \frac{H(t)}{4\pi c^2 t}\,\delta(|\mathbf{r}| - ct)\,.} \tag{11.2.10}$$

The solution of the inhomogeneous wave equation $u_{tt} - c^2\nabla^2 u = f(\mathbf{r},t)$ on $\mathbb{R}^3 \times \mathbb{R}_+$ with vanishing initial data is therefore

$$u(\mathbf{r},t) = \frac{1}{4\pi c^2} \int_{\mathbb{R}^3} \frac{f\big(\mathbf{r}', t - |\mathbf{r}-\mathbf{r}'|/c\big)}{|\mathbf{r}-\mathbf{r}'|}\, dV', \tag{11.2.11}$$

the celebrated *retarded potential* of Liénard and Wiechert. The structure of (11.2.11) encodes the *Huygens principle*: the value of u at $(\mathbf{r},t)$ depends only on the values of the source f on the backward light cone of $(\mathbf{r},t)$ (Fig. 11.2), that is on data that have had exactly enough time to propagate from $\mathbf{r}'$ to $\mathbf{r}$ at the speed c. In two space dimensions Huygens's principle fails, and the retarded potential is replaced by an integral over the entire interior of the backward light cone—a phenomenon whose physical interpretation ("two-dimensional waves leave a tail behind") is one of the more striking dimensional anomalies in mathematical physics.

11.3 Green's Functions on Bounded Domains

The fundamental solutions of Section 11.2 solve the free-space problem. To accommodate boundary conditions one passes to the *Green's function*, the analogue of E adapted to the geometry of the domain.

Definition 11.2 (Green's function) Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with sufficiently smooth boundary $\partial\Omega$, and let L be a linear differential operator acting on functions $u : \Omega \to \mathbb{R}$. The *Green's function* of L on Ω with homogeneous Dirichlet boundary conditions is the function $G(\mathbf{r},\mathbf{r}')$, defined on $\Omega \times \Omega$, satisfying

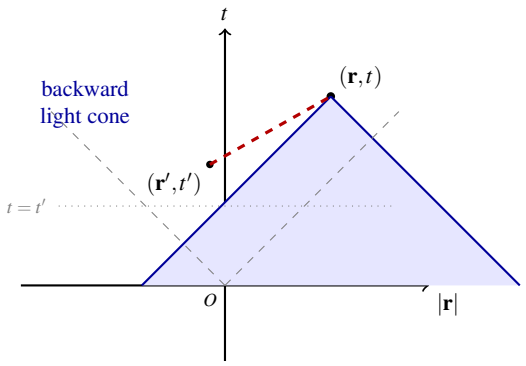


Fig. 11.2 The backward light cone of the observation point $(\mathbf{r},t)$ in three-dimensional spacetime (drawn here with one spatial dimension suppressed). The light-cone surface is given by $|\mathbf{r}-\mathbf{r}_0| = c\,(t-t_0)$. By the Huygens principle, the value of $u(\mathbf{r},t)$ in formula (11.2.11) is determined by the values of the source f on the cone's surface alone—only those source events whose signal has had exactly enough time to reach $(\mathbf{r},t)$ at the speed c. The dashed segment from $(\mathbf{r}',t')$ to $(\mathbf{r},t)$ has length $c\,(t-t')$.

$$L_{\mathbf{r}}\,G(\mathbf{r},\mathbf{r}') = \delta(\mathbf{r}-\mathbf{r}') \quad \text{in } \Omega, \qquad G(\mathbf{r},\mathbf{r}') = 0 \quad \text{for } \mathbf{r}\in\partial\Omega, \tag{11.3.1}$$

where $L_{\mathbf{r}}$ indicates differentiation with respect to the unprimed variable. Green's functions for Neumann or Robin conditions are defined analogously, with the corresponding homogeneous condition imposed on $\partial\Omega$.

The role played by G is made precise by the following representation theorem.

Theorem 11.1 (Green's representation) *Let u be the solution of the inhomogeneous Dirichlet problem*

$$L[u] = f \ \text{ in } \Omega, \qquad u = g \ \text{ on } \partial\Omega, \tag{11.3.2}$$

where $L = -\nabla^2$. Then

$$u(\mathbf{r}) = \int_\Omega G(\mathbf{r},\mathbf{r}')\,f(\mathbf{r}')\,dV' - \oint_{\partial\Omega} g(\mathbf{r}')\,\frac{\partial G(\mathbf{r},\mathbf{r}')}{\partial n'}\,dS', \tag{11.3.3}$$

where $\partial/\partial n'$ denotes differentiation with respect to $\mathbf{r}'$ in the direction of the outward normal to $\partial\Omega$.

Sketch of proof. Apply Green's second identity, which is obtained by subtracting two copies of Green's first identity (1.1.10),

$$\int_\Omega (v\nabla^2 u - u\nabla^2 v)\,dV = \oint_{\partial\Omega}\Big(v\,\frac{\partial u}{\partial n} - u\,\frac{\partial v}{\partial n}\Big)\,dS,$$

to the pair $u(\mathbf{r}')$ and $v(\mathbf{r}') = G(\mathbf{r},\mathbf{r}')$. The boundary term involving $\partial u/\partial n'$ vanishes because $G = 0$ on $\partial\Omega$; the volume term involving $u\nabla^2_{\mathbf{r}'}G$ collapses, by (11.3.1), to $-u(\mathbf{r})$. Rearranging gives (11.3.3). □ □

The volume term of (11.3.3) expresses the contribution of the inhomogeneous source f, while the surface term expresses that of the inhomogeneous boundary data g. The Green's function thus provides a single object from which the entire solution operator of the boundary-value problem can be reconstructed.

Symmetry of the Green's function. For the Laplace operator on a bounded domain, the Green's function is *symmetric*,

$$G(\mathbf{r},\mathbf{r}') = G(\mathbf{r}',\mathbf{r}). \tag{11.3.4}$$

This is a consequence of the self-adjointness of $-\nabla^2$: applying Green's second identity to two copies of G, $G(\mathbf{r}_1,\mathbf{r}')$ and $G(\mathbf{r}_2,\mathbf{r}')$ regarded as functions of $\mathbf{r}'$, the boundary terms vanish (both Green's functions are zero on $\partial\Omega$), and the volume terms reduce to $G(\mathbf{r}_1,\mathbf{r}_2) - G(\mathbf{r}_2,\mathbf{r}_1)$. The relation (11.3.4) expresses a reciprocity principle of physical significance: the response at $\mathbf{r}$ to a unit source placed at $\mathbf{r}'$ equals the response at $\mathbf{r}'$ to a unit source placed at $\mathbf{r}$.

Construction by the method of images. For domains with sufficiently symmetric geometry, the Green's function can be constructed from the free-space fundamental solution by adding fictitious source terms outside Ω that cancel the boundary values. The simplest example is the upper half-space $\{z>0\}$ with Dirichlet condition on $\{z=0\}$. The fundamental solution (11.2.3) placed at (x_0,y_0,z_0) with $z_0>0$ does not vanish on the boundary; placing an opposite "image" source at $(x_0,y_0,-z_0)$ produces the combination

$$G(\mathbf{r},\mathbf{r}') = -\frac{1}{4\pi}\left[\frac{1}{|\mathbf{r}-\mathbf{r}'|} - \frac{1}{|\mathbf{r}-\mathbf{r}'_*|}\right], \tag{11.3.5}$$

where $\mathbf{r}'_*$ denotes the reflection of $\mathbf{r}'$ across the plane $z=0$. By symmetry the two terms cancel exactly when $z=0$, and the requisite boundary condition is satisfied (Fig. 11.3). The same idea, with appropriate modifications, produces the Green's function of the Laplacian for the ball, the half-plane, the strip, and a number of other classical domains.

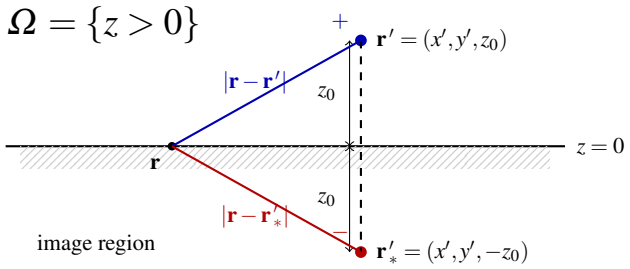


Fig. 11.3 The method of images for the upper half-space with Dirichlet boundary condition on $\{z=0\}$. A unit positive source placed at $\mathbf{r}' = (x',y',z_0)$ inside Ω is paired with an opposite "image" source at the reflected point $\mathbf{r}'_* = (x',y',-z_0)$. For any point $\mathbf{r}$ on the boundary plane $z=0$ the two distances $|\mathbf{r}-\mathbf{r}'|$ and $|\mathbf{r}-\mathbf{r}'_*|$ coincide, so the two fundamental solutions in (11.3.5) cancel exactly: the homogeneous Dirichlet boundary condition is enforced by symmetry.

11.4 The Spectral Representation

The Sturm–Liouville theory of Chapter 9 provides an alternative construction of the Green's function as an eigenfunction expansion, complementary to the integral construction of the preceding section. This representation is the analytic bridge between the bounded-domain methods of Chapters 7–10 and the free-space methods of the present chapter.

Let $\{\varphi_n\}_{n=1}^{\infty}$ be a complete orthonormal family of eigenfunctions of the operator $L = -\nabla^2$ on Ω with homogeneous Dirichlet conditions, and let $\{\lambda_n\}$ be the corresponding (positive) eigenvalues:

$$-\nabla^2 \varphi_n = \lambda_n \varphi_n \ \text{ in } \Omega, \qquad \varphi_n = 0 \ \text{ on } \partial\Omega, \qquad \int_\Omega \varphi_m \varphi_n \, dV = \delta_{mn}.$$

By the Sturm–Liouville completeness theorem 9.1, every sufficiently regular function on Ω admits an expansion in this basis. Applying it to the Green's function, for each fixed $\mathbf{r}'$,

$$G(\mathbf{r}, \mathbf{r}') = \sum_{n=1}^{\infty} c_n(\mathbf{r}') \, \varphi_n(\mathbf{r}).$$

Substituting into the defining relation (11.3.1),

$$-\nabla^2 G(\mathbf{r}, \mathbf{r}') = \sum_{n=1}^{\infty} c_n(\mathbf{r}') \, \lambda_n \, \varphi_n(\mathbf{r}) = \delta(\mathbf{r} - \mathbf{r}'),$$

and using the completeness relation $\delta(\mathbf{r}-\mathbf{r}') = \sum_n \varphi_n(\mathbf{r}) \, \varphi_n(\mathbf{r}')$, one reads off $c_n(\mathbf{r}') = \varphi_n(\mathbf{r}')/\lambda_n$. Thus

$$\boxed{G(\mathbf{r}, \mathbf{r}') = \sum_{n=1}^{\infty} \frac{\varphi_n(\mathbf{r}) \, \varphi_n(\mathbf{r}')}{\lambda_n} \, .} \tag{11.4.1}$$

The expansion (11.4.1) expresses G as a sum of contributions from each Sturm–Liouville mode, weighted inversely by the corresponding eigenvalue. Symmetry $G(\mathbf{r}, \mathbf{r}') = G(\mathbf{r}', \mathbf{r})$ is now manifest: each summand of (11.4.1) is symmetric under $\mathbf{r} \leftrightarrow \mathbf{r}'$.

The spectral representation has a striking corollary: the smallest eigenvalue λ_1 controls the size of the Green's function. For domains in which λ_1 is large (small or thin domains, by the Faber–Krahn inequality), the Green's function is correspondingly small, expressing the rapid decay of the influence of the source.

Connection with the heat kernel. The Green's function and the heat kernel are intimately related on a bounded domain. With $K(\mathbf{r}, \mathbf{r}'; t)$ the heat kernel on Ω with homogeneous Dirichlet boundary conditions, that is the solution of

$$\partial_t K = \alpha \nabla^2 K \text{ in } \Omega, \qquad K|_{\partial\Omega} = 0, \qquad K(\mathbf{r}, \mathbf{r}'; 0) = \delta(\mathbf{r} - \mathbf{r}'),$$

the spectral representation gives

$$K(\mathbf{r},\mathbf{r}';t)=\sum_{n=1}^{\infty}e^{-\alpha\lambda_n t}\,\varphi_n(\mathbf{r})\,\varphi_n(\mathbf{r}'),$$

and hence, by integrating against t from 0 to ∞,

$$G(\mathbf{r},\mathbf{r}')=\alpha\int_0^{\infty}K(\mathbf{r},\mathbf{r}';t)\,dt. \tag{11.4.2}$$

The Green's function is, up to a constant, the time integral of the heat kernel: the static problem is the time integral of the parabolic one, an observation of considerable depth that connects elliptic and parabolic regularity.

Worked Exercises

Exercise 11.1 Using the method of images, derive the explicit solution of the Dirichlet problem

$$\nabla^2 u=0 \ \text{ in } \{y>0\},\qquad u(x,0)=\varphi(x),$$

known as the Poisson integral formula for the upper half-plane.

Solution. The fundamental solution of ∇^2 on $\mathbb{R}^2$ is (11.2.4), $E(\mathbf{r})=(2\pi)^{-1}\ln r$. The Green's function on the upper half-plane $\{y>0\}$ with Dirichlet condition on $y=0$ is, in analogy with (11.3.5),

$$G(\mathbf{r},\mathbf{r}')=\frac{1}{2\pi}\left[\ln|\mathbf{r}-\mathbf{r}'|-\ln|\mathbf{r}-\mathbf{r}'_{*}|\right],$$

where $\mathbf{r}=(x,y)$, $\mathbf{r}'=(x',y')$, $\mathbf{r}'_{*}=(x',-y')$ is the reflection of $\mathbf{r}'$ in the line $y=0$. The outward normal on $y=0$ points in the negative y-direction; explicit computation gives, for $\mathbf{r}'=(x',0)$ on the boundary,

$$\left.\frac{\partial G}{\partial n'}\right|_{y'=0}=-\left.\frac{\partial G}{\partial y'}\right|_{y'=0}=-\frac{1}{\pi}\,\frac{y}{(x-x')^2+y^2}.$$

Substituting into Theorem 11.1 (with $f\equiv 0$),

$$u(x,y)=\frac{y}{\pi}\int_{-\infty}^{\infty}\frac{\varphi(x')}{(x-x')^2+y^2}\,dx', \tag{11.4.3}$$

the celebrated *Poisson integral formula* for the upper half-plane. The kernel $y/[\pi((x-x')^2+y^2)]$ is the *Poisson kernel*; as $y\to 0^+$ it converges, in the distributional sense, to $\delta(x-x')$, so that the boundary value $u(x,0^+)=\varphi(x)$ is recovered.

Exercise 11.2 By the method of images, derive the heat kernel and the resulting solution formula for the initial-boundary value problem on the half-line,

$$u_t=\alpha u_{xx}\ \ (x>0,t>0),\qquad u(0,t)=0,\qquad u(x,0)=f(x).$$

Solution. The fundamental solution of the one-dimensional heat operator is (11.2.7) with $n=1$,

$$E(x,t)=\frac{H(t)}{\sqrt{4\pi\alpha t}}\exp\!\left(-\frac{x^2}{4\alpha t}\right).$$

On the half-line $\{x>0\}$ with Dirichlet condition at $x=0$, the image method pairs each source at $x'>0$ with an opposite-sign source at $-x'<0$. The resulting heat kernel is

$$K(x,x';t)=\frac{1}{\sqrt{4\pi\alpha t}}\Big[e^{-(x-x')^2/(4\alpha t)}-e^{-(x+x')^2/(4\alpha t)}\Big]\qquad (x,x'>0,\,t>0).$$

The factor in brackets vanishes when $x=0$, satisfying the boundary condition by symmetry. The solution of the initial-boundary value problem is therefore

$$u(x,t)=\int_0^\infty K(x,x';t)\,f(x')\,dx'. \tag{11.4.4}$$

For $f\equiv 1$ (uniform initial temperature), the integral evaluates to $u(x,t)=\operatorname{erf}\!\big(x/\sqrt{4\alpha t}\big)$, the Gaussian error function—a classical result.

Exercise 11.3 A point source emits radiation according to $f(\mathbf{r},t)=\delta(\mathbf{r})\,\sin\omega t\,H(t)$ in three space dimensions. Find the resulting wave field $u(\mathbf{r},t)$ from the retarded-potential formula (11.2.11).

Solution. Substituting the source into (11.2.11),

$$u(\mathbf{r},t)=\frac{1}{4\pi c^2}\int_{\mathbb{R}^3}\frac{\delta(\mathbf{r}')\,\sin\omega(t-|\mathbf{r}-\mathbf{r}'|/c)\,H(t-|\mathbf{r}-\mathbf{r}'|/c)}{|\mathbf{r}-\mathbf{r}'|}\,dV'.$$

The volume integral collapses on the delta to give

$$u(\mathbf{r},t)=\frac{H(t-r/c)}{4\pi c^2 r}\,\sin\omega(t-r/c),\qquad r=|\mathbf{r}|. \tag{11.4.5}$$

The factor $H(t-r/c)$ encodes the finite speed of propagation: at the spatial point $\mathbf{r}$ no signal is received until time $t=r/c$ has elapsed since the onset of the source. Once the wave has arrived, the field oscillates at the source frequency ω but lagged by r/c, with an amplitude that decreases as $1/r$, as required by energy conservation in three dimensions.

Exercise 11.4 Construct, by means of the spectral representation (11.4.1), the Green's function of $-d^2/dx^2$ on the interval $(0,\ell)$ with Dirichlet boundary conditions $u(0)=u(\ell)=0$.

Solution. The eigenvalue problem

$$-\varphi''=\lambda\varphi,\qquad \varphi(0)=\varphi(\ell)=0,$$

has eigenvalues $\lambda_n=(n\pi/\ell)^2$ and orthonormal eigenfunctions $\varphi_n(x)=\sqrt{2/\ell}\,\sin(n\pi x/\ell)$ ($n=1,2,\dots$). The spectral representation (11.4.1) gives

$$G(x,x') = \frac{2}{\ell}\sum_{n=1}^{\infty} \frac{\sin(n\pi x/\ell)\,\sin(n\pi x'/\ell)}{(n\pi/\ell)^2}. \tag{11.4.6}$$

This series may be summed in closed form. The result, easily verified by direct construction (matching continuity of G and a unit jump in G_x at $x = x'$), is

$$G(x,x') = \begin{cases} \dfrac{x(\ell - x')}{\ell}, & 0 < x < x' < \ell, \\[2mm] \dfrac{x'(\ell - x)}{\ell}, & 0 < x' < x < \ell. \end{cases} \tag{11.4.7}$$

The two representations (11.4.6) and (11.4.7) illustrate the dual character of the Green's function: as an infinite series of Sturm–Liouville modes and as an explicit piecewise-linear function with a prescribed jump in its derivative.

Exercise 11.5 Show that the fundamental solution of the time-dependent Schrödinger operator $i\hbar\,\partial_t + \dfrac{\hbar^2}{2m}\nabla^2$ on $\mathbb{R}^3 \times \mathbb{R}_+$, with the causal condition $E \equiv 0$ for $t < 0$, is the free-space Schrödinger propagator

$$E(\mathbf{r},t) = H(t)\Big(\frac{m}{2\pi i\hbar t}\Big)^{3/2} \exp\!\left(\frac{im\,|\mathbf{r}|^2}{2\hbar t}\right). \tag{11.4.8}$$

Solution. Spatial Fourier transformation reduces the equation $i\hbar\partial_t E + (\hbar^2/2m)\nabla^2 E = \delta(\mathbf{r})\,\delta(t)$ to the ODE

$$i\hbar\partial_t \hat{E} - (\hbar^2|\xi|^2/2m)\hat{E} = \delta(t),$$

the causal solution of which is

$$\hat{E}(\xi,t) = H(t)\,\frac{1}{i\hbar}\,e^{-i\hbar|\xi|^2 t/(2m)}.$$

Inversion (a Gaussian integral with imaginary parameter, evaluated by analytic continuation from the heat-kernel calculation) yields the formula (11.4.8). The Schrödinger propagator is, formally, the heat kernel with the diffusion constant α replaced by $i\hbar/(2m)$; the analytic continuation from real to imaginary time that connects the two is the basis of the Wick rotation used in quantum field theory.

Problems

11.1 Show, by direct computation, that the function $E(\mathbf{r}) = -1/(4\pi r)$ satisfies $\nabla^2 E = 0$ on $\mathbb{R}^3 \setminus \{0\}$. Then verify, by integrating $\nabla^2 E$ against a smooth radial test function compactly supported near the origin, that the distributional identity $\nabla^2 E = \delta(\mathbf{r})$ holds.

11.2 Derive the fundamental solution (11.2.4) of the two-dimensional Laplacian by the same method employed in the text for the three-dimensional case. Show that, in two

dimensions, the constant of integration is fixed by Gauss's theorem applied to a disc of radius ρ, yielding the factor $1/(2\pi)$.

11.3 By the method of images, construct the Green's function of the Laplacian on the quarter-plane $\{x>0, y>0\}$ with homogeneous Dirichlet conditions on the boundary axes. (*Hint*: a single image is insufficient; three images are required.)

11.4 Show that the heat kernel (11.2.7) satisfies, for every $t_1, t_2 > 0$, the *semigroup property*

$$E(\mathbf{r}_1 - \mathbf{r}_2, t_1 + t_2) = \int_{\mathbb{R}^n} E(\mathbf{r}_1 - \mathbf{r}', t_1)\, E(\mathbf{r}' - \mathbf{r}_2, t_2)\, dV',$$

which expresses, by means of a Gaussian integral, the composition of heat propagation over two successive time intervals.

11.5 Let u satisfy the inhomogeneous heat equation

$$u_t - \alpha \nabla^2 u = f(\mathbf{r}, t) \ \ (\mathbf{r} \in \mathbb{R}^n, t > 0), \qquad u(\mathbf{r}, 0) = u_0(\mathbf{r}).$$

Using the heat kernel (11.2.7), write down an explicit integral representation of $u(\mathbf{r}, t)$ in terms of u_0 and f (this is the *Duhamel formula* for the heat equation).

11.6 Construct the Green's function of $-d^2/dx^2$ on the interval $(0, \ell)$ with the mixed boundary conditions $u(0) = 0$, $u'(\ell) = 0$, both as a spectral series and in piecewise closed form. Compare the two representations.

11.7 A spherical shell of radius a is held at zero temperature on its inner surface and at the constant temperature T_0 on its outer surface. Using the radial Green's function for the operator $-(1/r^2)\, d/dr(r^2\, d/dr)$ on (a_1, a_2) with Dirichlet conditions, derive the steady-state temperature distribution within the shell.

11.8 Show that the retarded potential (11.2.11) satisfies $\partial_t^2 u - c^2 \nabla^2 u = f$ in the sense of distributions, by verifying that the four-dimensional integral over the source $f(\mathbf{r}', t')$ collapses, on the delta of the wave operator, to give the requisite identity. (*Hint*: change variables to centre on the backward light cone.)

11.9 The biharmonic operator is $\nabla^4 \equiv \nabla^2 \nabla^2$. Using the fundamental solution of the Laplacian, show that the fundamental solution of ∇^4 on $\mathbb{R}^3$ is

$$E_{\nabla^4}(\mathbf{r}) = -\frac{r}{8\pi} \qquad (r = |\mathbf{r}|).$$

Deduce the integral representation of the solution of the biharmonic equation $\nabla^4 u = f$ on $\mathbb{R}^3$ for compactly supported f.

11.10 Let $\Omega \subset \mathbb{R}^n$ be a bounded domain. Show, using the spectral representation (11.4.1), that the Green's function of $-\nabla^2 + m^2$ ($m > 0$) on Ω with Dirichlet conditions satisfies

$$0 \le G_m(\mathbf{r}, \mathbf{r}') \le G_0(\mathbf{r}, \mathbf{r}')$$

for all $\mathbf{r}, \mathbf{r}' \in \Omega$, where G_0 is the Green's function of $-\nabla^2$. Interpret this inequality physically: the Helmholtz operator $-\nabla^2 + m^2$ damps the response, so its Green's function is bounded above by that of the Laplacian.

Chapter 12
Variational Methods and Energy Estimates

Abstract The Laplace, heat and wave equations all admit variational characterisations as the Euler–Lagrange equations of integral functionals. The same functionals provide *energy identities* from which the uniqueness of solutions to the canonical boundary-value and Cauchy problems follows directly, without explicit construction of the solutions. We close with the Rayleigh–Ritz approximation method, the variational basis of finite-element analysis.

12.1 Functionals and the Euler–Lagrange Equation

The variational method of Section 9.3 recast the search for the eigenvalues of a Sturm–Liouville operator as an extremal problem for an integral functional. We here recapitulate the result in a form sufficiently general to apply to partial differential equations.

Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with smooth boundary, and let $u(\mathbf{r})$ be a sufficiently smooth function on Ω satisfying prescribed boundary conditions on $\partial\Omega$. Consider an integral functional of the form

$$J[u] \;=\; \int_{\Omega} F\big(\mathbf{r}, u(\mathbf{r}), \operatorname{grad} u(\mathbf{r})\big)\, dV, \tag{12.1.1}$$

in which $F = F(\mathbf{r}, u, \mathbf{p})$ is a smooth function of its $n+2$ arguments. The variational problem associated with (12.1.1) is the determination of those u at which J is stationary against all admissible variations,

$$u(\mathbf{r}) \;\longmapsto\; u(\mathbf{r}) + \varepsilon\, \eta(\mathbf{r}),$$

where η is an arbitrary smooth function vanishing on $\partial\Omega$ (so that the boundary conditions are preserved). The condition that J be stationary at u reads

$$\delta J[u] \;\equiv\; \frac{d}{d\varepsilon} J[u + \varepsilon\eta]\bigg|_{\varepsilon=0} = 0 \qquad \text{for every admissible } \eta. \tag{12.1.2}$$

Computing the derivative under the integral sign,

$$\delta J[u] = \int_\Omega \left[\frac{\partial F}{\partial u}\,\eta + \sum_{i=1}^{n} \frac{\partial F}{\partial p_i}\,\frac{\partial \eta}{\partial x_i} \right] dV,$$

and integrating the second term by parts—the boundary contributions vanishing because $\eta = 0$ on $\partial\Omega$—one obtains

$$\delta J[u] = \int_\Omega \left[\frac{\partial F}{\partial u} - \sum_{i=1}^{n} \frac{\partial}{\partial x_i}\left(\frac{\partial F}{\partial p_i} \right) \right] \eta \, dV.$$

Since this expression must vanish for every admissible η, the bracket itself must vanish identically: u satisfies the *Euler–Lagrange equation*

$$\boxed{\frac{\partial F}{\partial u} - \sum_{i=1}^{n} \frac{\partial}{\partial x_i}\left(\frac{\partial F}{\partial p_i} \right) = 0 \text{ in } \Omega.} \tag{12.1.3}$$

The equation (12.1.3) is the partial-differential analogue of the ordinary-differential Euler–Lagrange equation derived in (9.3.10). It extends with little modification to integrands F involving second derivatives of u, in which case the boundary contributions yield the appropriate *natural boundary conditions*.

12.2 The Dirichlet Principle

The Laplace equation is the most fundamental example of a PDE with a variational origin. Consider, on a bounded domain $\Omega \subset \mathbb{R}^n$, the *Dirichlet energy*

$$D[u] \;=\; \frac{1}{2}\int_\Omega |\operatorname{grad} u|^2 \, dV \;=\; \frac{1}{2}\sum_{i=1}^{n} \int_\Omega \left(\frac{\partial u}{\partial x_i} \right)^2 dV, \tag{12.2.1}$$

and let g be a prescribed continuous function on $\partial\Omega$. The *Dirichlet variational problem* asks for the function that minimises $D[u]$ among all sufficiently smooth u with $u = g$ on $\partial\Omega$.

Applying (12.1.3) to the integrand $F = \tfrac{1}{2}|\mathbf{p}|^2$,

$$\frac{\partial F}{\partial u} = 0, \qquad \frac{\partial F}{\partial p_i} = p_i = \frac{\partial u}{\partial x_i}, \qquad \sum_{i=1}^{n} \frac{\partial}{\partial x_i}\left(\frac{\partial u}{\partial x_i} \right) = \nabla^2 u,$$

the Euler–Lagrange equation reduces to

$$\nabla^2 u = 0 \text{ in } \Omega, \tag{12.2.2}$$

which is precisely the Laplace equation. Thus, among admissible functions taking the prescribed boundary values g on $\partial\Omega$, the harmonic one is uniquely characterised as the function of least Dirichlet energy.

Theorem 12.1 (Dirichlet's principle) *Suppose u is a sufficiently smooth function on $\bar{\Omega}$ with $u = g$ on $\partial\Omega$. Then u minimises the Dirichlet energy $D[\cdot]$ over the class of admissible functions if and only if u is harmonic in Ω, that is, satisfies $\nabla^2 u = 0$.*

Proof. That a minimiser satisfies (12.2.2) has just been shown. Conversely, suppose u is harmonic with $u = g$ on $\partial\Omega$, and let v be any other admissible function. Write $v = u + \eta$ with $\eta \equiv 0$ on $\partial\Omega$. Then

$$D[v] - D[u] = \int_\Omega \operatorname{grad} u \cdot \operatorname{grad} \eta \, dV + \frac{1}{2}\int_\Omega |\operatorname{grad}\eta|^2 \, dV.$$

The first integral, by Green's first identity (1.1.10) applied with $f = u$ and $g = \eta$, equals

$$-\int_\Omega \eta \, \nabla^2 u \, dV + \oint_{\partial\Omega} \eta \, \frac{\partial u}{\partial n} \, dS = 0,$$

the volume term vanishing because $\nabla^2 u = 0$ and the boundary term because $\eta = 0$ on $\partial\Omega$. Hence

$$D[v] - D[u] = \frac{1}{2}\int_\Omega |\operatorname{grad}\eta|^2 \, dV \geq 0,$$

with equality only when $\operatorname{grad}\eta \equiv 0$, that is, η is constant; the boundary condition $\eta = 0$ forces this constant to be zero. □ □

Uniqueness for the Dirichlet problem. Theorem 12.1 immediately yields the uniqueness of the Dirichlet problem $\nabla^2 u = 0$ in Ω, $u = g$ on $\partial\Omega$: if u_1 and u_2 both solve the problem, the difference $w = u_1 - u_2$ is harmonic with $w = 0$ on $\partial\Omega$, hence minimises $D[\cdot]$ over the class of functions with zero boundary values. The minimum value of $D[\cdot]$ over that class is 0, attained at $w \equiv 0$, so $u_1 = u_2$ throughout Ω.

The same argument, adapted to admit a non-vanishing right-hand side, proves the uniqueness of the Poisson problem $\nabla^2 u = h$, $u = g$ on $\partial\Omega$.

Maximum principle and Dirichlet's principle. The Dirichlet principle provides one route to the classical *maximum principle* for harmonic functions: a non-constant harmonic function on a bounded domain attains its maximum and minimum on the boundary, never in the interior. We do not develop this corollary here, since the proof through the mean-value theorem (sketched in the worked examples below) is more direct.

12.3 Hamilton's Principle and the Wave Equation

The wave equation arises from the variational principle of classical mechanics, namely *Hamilton's principle* of stationary action. Consider a one-dimensional vibrating string of length ℓ, with linear mass density ρ and constant tension T. Let $u(x,t)$ denote the transverse displacement, and let

$$\mathscr{L} = \frac{1}{2}\rho\, u_t^2 - \frac{1}{2} T\, u_x^2 \tag{12.3.1}$$

denote the *Lagrangian density*: the kinetic-energy density minus the potential-energy density. The *action functional* associated with a motion $u(x,t)$ on the time interval $[t_1,t_2]$ is

$$S[u] \;=\; \int_{t_1}^{t_2}\!\!\int_0^{\ell} \mathscr{L}(u_x,u_t)\,dx\,dt. \tag{12.3.2}$$

Hamilton's principle asserts that the physical motion is the one that renders $S[u]$ stationary among all admissible variations that vanish at the initial and final times.

The Euler–Lagrange equation associated with the Lagrangian (12.3.1) is, by direct application of (12.1.3) adapted to two independent variables,

$$\frac{\partial}{\partial t}\left(\frac{\partial \mathscr{L}}{\partial u_t}\right) + \frac{\partial}{\partial x}\left(\frac{\partial \mathscr{L}}{\partial u_x}\right) - \frac{\partial \mathscr{L}}{\partial u} = 0,$$

namely

$$\rho\, u_{tt} - T\, u_{xx} = 0,$$

or, with $c = \sqrt{T/\rho}$,

$$u_{tt} = c^2\, u_{xx}, \tag{12.3.3}$$

the wave equation. The same derivation, generalised to n space dimensions, produces $u_{tt} = c^2\nabla^2 u$. The wave equation is thus the Euler–Lagrange equation of the action functional of a homogeneous elastic medium.

Energy conservation. The variational origin of the wave equation immediately yields a conservation law. Define the *total energy* at time t,

$$E(t) \;=\; \frac{1}{2}\int_0^{\ell} \bigl(\rho\, u_t^2 + T\, u_x^2\bigr)\,dx, \tag{12.3.4}$$

the sum of kinetic and potential parts. Differentiating,

$$\frac{dE}{dt} = \int_0^{\ell} \bigl(\rho\, u_t u_{tt} + T\, u_x u_{xt}\bigr)\,dx,$$

and integrating the second term by parts,

$$\int_0^{\ell} T\, u_x u_{xt}\,dx = \bigl[T\, u_x u_t\bigr]_0^{\ell} - \int_0^{\ell} T\, u_{xx} u_t\,dx,$$

the boundary term vanishing for the standard fixed-end ($u(0,t)=u(\ell,t)=0$, hence $u_t = 0$ at the endpoints) or free-end ($u_x = 0$) conditions. Therefore

$$\frac{dE}{dt} = \int_0^{\ell} u_t\bigl(\rho u_{tt} - T u_{xx}\bigr)\,dx = 0,$$

and the total energy is conserved:

$$E(t) = E(0) \qquad \text{for all } t \geq 0. \tag{12.3.5}$$

The conservation law (12.3.5) expresses, in the present context, the time-translation invariance of the Lagrangian (12.3.1), by Noether's theorem.

Uniqueness for the wave equation. Energy conservation provides an immediate proof of the uniqueness of the initial-boundary value problem for the wave equation. Suppose u_1 and u_2 are two solutions of

$$u_{tt} = c^2 u_{xx}, \;\; u(0,t) = u(\ell,t) = 0, \;\; u(x,0) = f(x), \;\; u_t(x,0) = g(x).$$

Their difference $w = u_1 - u_2$ satisfies the same equation with *vanishing* boundary and initial data:

$$w_{tt} = c^2 w_{xx}, \;\; w(0,t) = w(\ell,t) = 0, \;\; w(x,0) = 0, \;\; w_t(x,0) = 0.$$

The energy associated with w,

$$E_w(t) = \frac{1}{2}\int_0^\ell (\rho\, w_t^2 + T\, w_x^2)\, dx,$$

is conserved by (12.3.5). At $t = 0$ both w_t and w_x vanish, so $E_w(0) = 0$; and since $E_w(t)$ is non-negative and constant, $E_w(t) \equiv 0$ for all t. The non-negativity of each term forces $w_t \equiv 0$ and $w_x \equiv 0$, hence w is constant; the initial condition $w(x,0) = 0$ fixes this constant to be zero. Therefore $u_1 \equiv u_2$.

12.4 Energy Methods for the Heat Equation

The heat equation, being first-order in time, is not the Euler–Lagrange equation of a manifestly conservative action functional in the spirit of Hamilton's principle. It does, however, admit an energy-type analysis, which produces uniqueness, stability and the maximum principle in a unified way.

Consider the initial-boundary value problem for the heat equation on a bounded one-dimensional interval,

$$u_t = \alpha\, u_{xx}, \;\; 0 < x < \ell, t > 0, \quad u(0,t) = u(\ell,t) = 0, \quad u(x,0) = f(x), \tag{12.4.1}$$

and define the *thermal energy*

$$E_T(t) \;=\; \frac{1}{2}\int_0^\ell u^2(x,t)\, dx. \tag{12.4.2}$$

Differentiating and using the heat equation,

$$\frac{dE_T}{dt} = \int_0^\ell u\, u_t\, dx = \alpha \int_0^\ell u\, u_{xx}\, dx,$$

and integrating by parts, with the boundary conditions discarding the endpoint contributions,

$$\frac{dE_T}{dt} = -\alpha \int_0^\ell u_x^2 \, dx \;\le\; 0. \tag{12.4.3}$$

The thermal energy is therefore monotone non-increasing in time, and it strictly decreases unless $u_x \equiv 0$, that is, unless u is spatially constant. Combined with the boundary conditions, this means $u \equiv 0$. We have proved:

Theorem 12.2 (Energy decay and uniqueness for the heat equation) *Solutions of the homogeneous Dirichlet problem for the heat equation satisfy*

$$\int_0^\ell u^2(x,t)\,dx \;\le\; \int_0^\ell u^2(x,0)\,dx \;=\; \int_0^\ell f^2(x)\,dx$$

for every $t \ge 0$, with strict inequality unless $f \equiv 0$. In particular, the problem (12.4.1) *admits at most one solution.*

The proof of uniqueness mirrors that for the wave equation: the difference w of two solutions satisfies the homogeneous problem with vanishing initial datum, and its thermal energy at $t = 0$ is zero; non-negativity and monotonicity of E_T then force $w \equiv 0$.

Continuous dependence on the initial datum. Theorem 12.2 also yields, as an immediate corollary, the third of Hadamard's criteria from Section 5.2: the solution depends continuously on the data. If u_1 and u_2 are the solutions corresponding to initial data f_1 and f_2, the difference $w = u_1 - u_2$ satisfies

$$\int_0^\ell |u_1(x,t) - u_2(x,t)|^2\,dx \;\le\; \int_0^\ell |f_1(x) - f_2(x)|^2\,dx,$$

which is the desired stability statement in the L^2 norm. Small changes in the initial datum produce no larger changes in the solution—the heat equation is, on a bounded domain, well posed in Hadamard's sense.

Backward heat equation. The same energy identity (12.4.3) demonstrates the ill-posedness of the *backward* heat equation: integrating from 0 to $-T$, the energy grows in proportion to the spatial gradient of u, and the inequality (12.4.3) is replaced by an unbounded amplification. Small errors in the data are amplified into arbitrarily large errors in the solution—the structural ill-posedness anticipated in Section 5.3.

12.5 The Rayleigh–Ritz Method

The variational characterisation of the Laplace, heat and wave equations is more than conceptual; it furnishes a constructive numerical procedure—the *Rayleigh–Ritz method*—which forms the basis of finite-element analysis.

We illustrate the method on the Dirichlet problem

$$-\nabla^2 u = f \text{ in } \Omega, \qquad u = 0 \text{ on } \partial\Omega, \tag{12.5.1}$$

on a bounded domain $\Omega \subset \mathbb{R}^n$. The variational characterisation of (12.5.1), an extension of Dirichlet's principle, is the minimisation of

$$J[u] \;=\; \frac{1}{2}\int_\Omega |\operatorname{grad} u|^2 \, dV \;-\; \int_\Omega f\,u\,dV \tag{12.5.2}$$

over all admissible u with $u = 0$ on $\partial\Omega$. The Euler–Lagrange equation of J is precisely (12.5.1).

Approximation by trial functions. Let $\{\varphi_1, \varphi_2, \dots, \varphi_N\}$ be a finite system of admissible *trial functions*, each vanishing on $\partial\Omega$. Restricting J to linear combinations

$$u_N(\mathbf{r}) = \sum_{j=1}^{N} c_j\,\varphi_j(\mathbf{r}), \tag{12.5.3}$$

producing a real-valued function of the N coefficients $c_1, \dots, c_N$:

$$J[u_N] = \frac{1}{2}\sum_{i,j=1}^{N} c_i\,c_j\,K_{ij} - \sum_{i=1}^{N} c_i\,F_i,$$

with the *stiffness matrix* and *load vector*

$$K_{ij} = \int_\Omega \operatorname{grad}\varphi_i \cdot \operatorname{grad}\varphi_j \, dV, \qquad F_i = \int_\Omega f\,\varphi_i\,dV.$$

Setting $\partial J[u_N]/\partial c_k = 0$ produces the linear algebraic system

$$\sum_{j=1}^{N} K_{kj}\,c_j = F_k \qquad (k = 1, 2, \dots, N). \tag{12.5.4}$$

The system (12.5.4) is the Rayleigh–Ritz approximation to the boundary-value problem (12.5.1): a single PDE has been replaced by a system of N linear algebraic equations, whose solution determines the coefficients of the best (in the sense of minimum J) trial-function approximation.

Convergence of the method. If the family $\{\varphi_j\}$ is dense in the function space of admissible functions—that is, if every admissible function can be approximated arbitrarily closely by linear combinations of finitely many of the φ_j—then $u_N \to u$ as $N \to \infty$, in a sense made precise by the theory of Sobolev spaces. The convergence is quantified by Cea's lemma, which bounds the error in the grad-norm by the best approximation available in the trial-function family. The choice of trial functions is therefore the principal practical question; the modern theory of finite elements takes the φ_j to be piecewise-polynomial functions on a triangulation of Ω, securing flexibility, sparseness of K_{ij}, and controlled approximation error.

The Rayleigh quotient and eigenvalue estimates. Applied to the Sturm–Liouville problem $-\nabla^2\varphi = \lambda\varphi$ on Ω with homogeneous Dirichlet conditions, the variational formulation (12.5.2) produces, for the smallest eigenvalue, the celebrated *Rayleigh*

quotient

$$\lambda_1 = \inf_{\substack{u \in H_0^1(\Omega) \\ u \neq 0}} \frac{\int_\Omega |\operatorname{grad} u|^2 \, dV}{\int_\Omega u^2 \, dV}, \tag{12.5.5}$$

where $H_0^1(\Omega)$ is the natural function space of admissible u. The Rayleigh–Ritz method applied to (12.5.5) yields upper bounds on λ_1 from any choice of trial functions: substituting an admissible u_N produces a number that majorises λ_1, and the infimum over such trial functions converges to λ_1 as the trial space is enlarged. The same principle underlies the asymptotic estimates of the Sturm–Liouville eigenvalues established in Section 9.4.

Worked Exercises

Exercise 12.1 Derive the Euler–Lagrange equation associated with the functional

$$J[u] = \int_\Omega \Big[\tfrac{1}{2} |\operatorname{grad} u|^2 + \tfrac{1}{2} q(\mathbf{r})\, u^2 - f(\mathbf{r})\, u \Big] \, dV,$$

defined on functions with $u = 0$ on $\partial\Omega$, and identify the corresponding boundary-value problem.

Solution. With $F(\mathbf{r}, u, \mathbf{p}) = \tfrac{1}{2}|\mathbf{p}|^2 + \tfrac{1}{2} q\, u^2 - f u$,

$$\frac{\partial F}{\partial u} = q\, u - f, \qquad \frac{\partial F}{\partial p_i} = p_i, \qquad \sum_i \frac{\partial}{\partial x_i} \frac{\partial F}{\partial p_i} = \nabla^2 u.$$

The Euler–Lagrange equation (12.1.3) reads

$$q\, u - f - \nabla^2 u = 0, \quad \text{i.e.} \quad -\nabla^2 u + q\, u = f \text{ in } \Omega,$$

the inhomogeneous *Helmholtz* (or screened-Poisson) equation, together with the imposed Dirichlet condition $u = 0$ on $\partial\Omega$. The variational characterisation of this equation underlies the energy-method proofs of uniqueness and stability for the Helmholtz problem; in particular, the term $\tfrac{1}{2} q\, u^2$ in the functional becomes, with $q = m^2 > 0$, an additional positive contribution to the energy that strengthens uniqueness arguments.

Exercise 12.2 Show that the energy associated with the wave equation $u_{tt} = c^2 \nabla^2 u$ on a bounded domain $\Omega \subset \mathbb{R}^n$ with Dirichlet conditions $u = 0$ on $\partial\Omega$, namely

$$E(t) = \frac{1}{2} \int_\Omega \big(u_t^2 + c^2 |\operatorname{grad} u|^2 \big) \, dV,$$

is conserved.

Solution. Differentiating under the integral sign,

$$\frac{dE}{dt} = \int_\Omega \bigl(u_t u_{tt} + c^2 \operatorname{grad} u \cdot \operatorname{grad} u_t\bigr)\, dV.$$

Apply Green's first identity (1.1.10) to the second term, with $g = u_t$ and $f = u$:

$$\int_\Omega \operatorname{grad} u \cdot \operatorname{grad} u_t \, dV = -\int_\Omega u_t \, \nabla^2 u \, dV + \oint_{\partial\Omega} u_t \, \frac{\partial u}{\partial n}\, dS.$$

The boundary term vanishes (since $u = 0$ on $\partial\Omega$ implies $u_t = 0$ there). Substituting,

$$\frac{dE}{dt} = \int_\Omega u_t \bigl(u_{tt} - c^2 \nabla^2 u\bigr)\, dV = 0,$$

because $u_{tt} - c^2\nabla^2 u = 0$. Hence $E(t) = E(0)$ for all $t \geq 0$.

Exercise 12.3 Apply the Rayleigh–Ritz method to estimate the smallest eigenvalue λ_1 of $-\nabla^2$ on the unit disc $\Omega = \{(x,y) : x^2 + y^2 < 1\}$ with Dirichlet conditions, using the single trial function $\varphi(r) = 1 - r^2$ ($r = \sqrt{x^2+y^2}$).

Solution. The trial function $\varphi = 1 - r^2$ vanishes on the unit circle and is therefore admissible. Computing the numerator of (12.5.5),

$$\int_\Omega |\operatorname{grad}\varphi|^2 \, dV = \int_0^{2\pi}\!\!\int_0^1 |\varphi'(r)|^2 \, r\, dr\, d\theta = 2\pi \int_0^1 (2r)^2 \, r\, dr = 2\pi.$$

The denominator,

$$\int_\Omega \varphi^2 \, dV = 2\pi \int_0^1 (1-r^2)^2 \, r\, dr = 2\pi \int_0^1 (r - 2r^3 + r^5)\, dr = 2\pi\bigl[\tfrac{1}{2} - \tfrac{1}{2} + \tfrac{1}{6}\bigr] = \frac{\pi}{3}.$$

Therefore

$$\lambda_1 \leq \frac{2\pi}{\pi/3} = 6.$$

The exact value is $\lambda_1 = j_{0,1}^2 \approx 5.7831$, where $j_{0,1} \approx 2.4048$ is the first positive zero of J_0. The trial function $1 - r^2$ thus yields a Rayleigh-quotient upper bound within 4% of the exact eigenvalue—a remarkably good estimate from a single, very simple trial function. The agreement reflects the close shape resemblance between the trial function and the true fundamental eigenfunction (Fig. 12.1): both vanish at $r = 1$, both are radial, both attain their maximum at the centre, and both decrease monotonically.

Exercise 12.4 Prove the maximum principle for the heat equation: a solution $u(x,t)$ of $u_t = \alpha u_{xx}$ on $0 \leq x \leq \ell$, $0 \leq t \leq T$, attains its maximum and minimum either on the initial line $\{t = 0\}$ or on the spatial boundary $\{x = 0\} \cup \{x = \ell\}$.

Solution. Suppose, for a contradiction, that the maximum is attained at an interior point (x_0, t_0) with $0 < x_0 < \ell$, $0 < t_0 \leq T$. Consider the auxiliary function $v(x,t) = u(x,t) + \varepsilon(x - x_0)^2$, with $\varepsilon > 0$ small. At an interior maximum of v one would have $v_{xx} \leq 0$ and $v_t \geq 0$. Computing,

$$v_{xx} = u_{xx} + 2\varepsilon, \qquad v_t = u_t = \alpha\, u_{xx}.$$

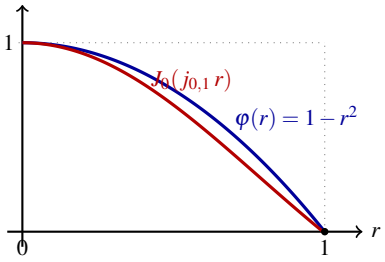


Fig. 12.1 Comparison of the Rayleigh–Ritz trial function $\varphi(r) = 1 - r^2$ (blue) and the exact fundamental Dirichlet eigenfunction $J_0(j_{0,1}\,r)$ (red) on the unit disc, both normalised to the value 1 at the centre. The two profiles are visually almost indistinguishable; the small difference accounts for the 4% discrepancy between the Rayleigh-quotient estimate $\lambda_1 \le 6$ and the exact value $\lambda_1 = j_{0,1}^2 \approx 5.78$.

Substituting,

$$v_t - \alpha\, v_{xx} = \alpha\, u_{xx} - \alpha\, u_{xx} - 2\alpha\varepsilon = -2\alpha\varepsilon < 0.$$

This contradicts $v_t \ge \alpha v_{xx}$, which holds at any interior maximum (since $v_t \ge 0$ and $v_{xx} \le 0$). Hence v has no interior maximum, and its maximum lies on the parabolic boundary. Letting $\varepsilon \to 0$ recovers the result for u itself.

Problems

12.1 Show that the Euler–Lagrange equation associated with the action functional

$$S[u] = \int_{t_1}^{t_2} \int_\Omega \Big[\tfrac{1}{2} u_t^2 - \tfrac{1}{2} c^2 |\operatorname{grad} u|^2 - V(\mathbf{r})\, u \Big]\, dV\, dt,$$

on functions $u(\mathbf{r},t)$ with prescribed values at $t = t_1$ and $t = t_2$, is the inhomogeneous wave equation $u_{tt} = c^2 \nabla^2 u - V(\mathbf{r})$.

12.2 Derive the energy conservation law for the inhomogeneous wave equation $u_{tt} = c^2\nabla^2 u + f(\mathbf{r},t)$ on a bounded domain with homogeneous Dirichlet conditions, and use it to estimate the rate of energy growth in terms of the source f. (*Hint*: the energy is no longer conserved, but dE/dt has a manageable expression.)

12.3 Apply the Rayleigh–Ritz method to estimate the smallest eigenvalue of $-\nabla^2$ on the square $\Omega = \{(x,y) : 0 < x < 1,\, 0 < y < 1\}$ with Dirichlet conditions, using the trial function $\varphi(x,y) = x(1-x)\,y(1-y)$. Compare the result with the exact value $\lambda_1 = 2\pi^2 \approx 19.74$.

12.4 Prove the uniqueness of the Neumann problem

$$\nabla^2 u = f \text{ in } \Omega, \qquad \frac{\partial u}{\partial n} = g \text{ on } \partial\Omega,$$

up to an additive constant, by an energy argument analogous to that of Section 12.4. (Recall from Problem 5.6 that this problem requires a compatibility condition for existence; the present uniqueness statement holds under that condition.)

12.5 Show that the Schrödinger equation

$$i\hbar\, u_t = -\frac{\hbar^2}{2m}\nabla^2 u + V(\mathbf{r})\, u$$

arises as the Euler–Lagrange equation of the complex Lagrangian density

$$\mathscr{L} = \frac{i\hbar}{2}(u^* u_t - u\, u_t^*) - \frac{\hbar^2}{2m}|\operatorname{grad} u|^2 - V(\mathbf{r})\,|u|^2,$$

treating u and u^* as independent variables.

12.6 Define the *Dirichlet energy functional*

$$D[u] = \frac{1}{2}\int_\Omega |\operatorname{grad} u|^2\, dV$$

on a bounded domain Ω. Show that, under the constraint $\int_\Omega u^2\, dV = 1$, the functional $D[u]$ is minimised by the first eigenfunction φ_1 of the Dirichlet Laplacian, with the minimum value equal to $\frac{1}{2}\lambda_1$. Generalise to the higher eigenvalues through the min-max principle.

12.7 A non-uniform string of length ℓ has linear mass density $\rho(x)$ and tension $T(x)$, with both functions positive and continuous on $[0,\ell]$. Show that the Euler–Lagrange equation of the action functional

$$S[u] = \int_{t_1}^{t_2}\int_0^\ell \Big[\tfrac{1}{2}\rho u_t^2 - \tfrac{1}{2}T u_x^2\Big]\, dx\, dt$$

is $\rho u_{tt} = (T u_x)_x$, and prove energy conservation in this non-uniform case.

12.8 Show, using the energy method of Section 12.4, that the solution of the heat equation $u_t = \alpha\nabla^2 u$ on a bounded domain with Dirichlet conditions decays exponentially in L^2:

$$\int_\Omega u^2(\mathbf{r},t)\, dV \le e^{-2\alpha\lambda_1 t}\int_\Omega u^2(\mathbf{r},0)\, dV,$$

where λ_1 is the smallest Dirichlet eigenvalue of $-\nabla^2$. (*Hint*: combine (12.4.3) with the Rayleigh quotient (12.5.5).)

Chapter 13
Beyond the Linear Theory: An Outlook

Abstract This concluding chapter sketches the directions in which the linear theory is extended in modern mathematical analysis. The themes treated are nonlinear equations and conservation laws, higher-order and tensor-valued equations, the probabilistic interpretation of diffusion, weak and viscosity solutions, and the numerical methods that descend from the variational viewpoint of Chapter 12. Each section closes with references for further study.

13.1 Nonlinear Partial Differential Equations

The restriction of this book to linear equations was deliberate, and the preface acknowledged it. The advantages of the restriction are considerable: linearity permits the principle of superposition; it admits a representation of the solution operator as an integral against a Green's function or a kernel; and it places at one's disposal the linear-algebraic vocabulary of eigenfunctions, operators, and spectra. Each of these features disappears in the nonlinear setting, and in compensation a richer phenomenology emerges—one that has been a central preoccupation of applied analysis since the second half of the twentieth century.

Conservation laws and shocks. The simplest nonlinear PDE is the *inviscid Burgers equation*,

$$u_t + u\,u_x = 0, \tag{13.1.1}$$

the prototype of a one-dimensional conservation law. The method of characteristics, developed in Section 6.3 for linear hyperbolic equations, applies formally to (13.1.1): the characteristics are the lines $x = u_0 t + x_0$ along which u is constant, but the slope of each characteristic depends on the value of u that it carries. Faster-moving characteristics overtake slower-moving ones, the solution becomes multivalued in finite time, and a *shock wave*—a discontinuous propagating front—forms. The classical solution theory breaks down at the shock and must be replaced by the theory of *weak solutions*, in which the equation is satisfied in an integrated rather than a pointwise sense. The

Rankine–Hugoniot jump conditions then determine the speed of propagation of the shock, while an additional *entropy condition* singles out the physically relevant weak solution among the infinitely many that the integrated equation admits. These notions, classical in gas dynamics, generalise to multidimensional systems of conservation laws and form the mathematical core of the theory of compressible flow.

Solitons and integrable equations. A different nonlinear regime is occupied by the *Korteweg–de Vries equation*,

$$u_t + 6\,u\,u_x + u_{xxx} = 0, \tag{13.1.2}$$

which describes weakly nonlinear, weakly dispersive long waves on a shallow fluid. Despite its nonlinearity, equation (13.1.2) is integrable: it admits an infinite hierarchy of conservation laws, its initial-value problem can be solved exactly by the *inverse-scattering transform*, and its solutions include *solitons*, localised waves that propagate without changing shape and that pass through one another with only a phase shift. The discovery of solitons in the 1960s, and the subsequent unfolding of integrable systems—the *sine-Gordon* equation $u_{tt} - u_{xx} + \sin u = 0$, the *nonlinear Schrödinger* equation $iu_t + u_{xx} + |u|^2 u = 0$, and many others—constitute one of the most striking developments of late-twentieth-century applied mathematics, with connections to algebraic geometry, Lie theory, and quantum field theory.

The Navier–Stokes equations. The most consequential nonlinear PDE of mathematical physics is the Navier–Stokes system governing the motion of an incompressible viscous fluid,

$$\partial_t \mathbf{u} + (\mathbf{u}\cdot\nabla)\mathbf{u} = -\nabla p + \nu\,\nabla^2\mathbf{u}, \qquad \nabla\cdot\mathbf{u} = 0. \tag{13.1.3}$$

The nonlinear advective term $(\mathbf{u}\cdot\nabla)\mathbf{u}$ generates phenomena—vortex stretching, energy cascade, turbulence—that have no linear analogue. Whether the three-dimensional initial-value problem for (13.1.3) admits a unique smooth solution for all time is, at present, a famous open question, the Navier–Stokes regularity problem being one of the seven Clay Mathematical Institute Millennium Prize problems. The body of analytic technique developed for its study—weak solutions of Leray, Sobolev embeddings, vorticity equations, energy methods—has nonetheless extended the linear theory of this book into one of the most active areas of contemporary mathematics.

Reaction-diffusion equations. Among the most physically diverse nonlinear PDEs are the *reaction-diffusion equations*, of the general form

$$u_t = \alpha\,\nabla^2 u + R(u), \tag{13.1.4}$$

in which the linear diffusion of Chapter 4 is augmented by a nonlinear reaction term $R(u)$. With $R(u) = u(1-u)$ one obtains the *Fisher–KPP* equation, governing the spread of an advantageous gene in a population; with $R(u) = u - u^3$, an equation modelling pattern formation in a chemical or biological system; with $R(u) = u^2$, an equation exhibiting finite-time blow-up. The qualitative theory of these equations—travelling waves, pattern formation, blow-up phenomena—combines the linear analytic methods

of the present book with techniques from dynamical systems and ordinary differential equation theory.

Further reading. A first introduction to nonlinear PDE theory at the level of this book is provided by Logan, *An Introduction to Nonlinear Partial Differential Equations* (Wiley, 2008); for conservation laws and shocks, Smoller, *Shock Waves and Reaction-Diffusion Equations* (Springer, 1994); for solitons and integrable systems, Drazin and Johnson, *Solitons: an Introduction* (Cambridge, 1989); for Navier–Stokes, Galdi, *An Introduction to the Mathematical Theory of the Navier–Stokes Equations* (Springer, 2011).

13.2 Higher-Order Equations and Systems

The restriction to second-order equations was, like the restriction to linear equations, deliberate. Higher-order equations and systems of equations occur, however, with sufficient frequency that we sketch them briefly.

The biharmonic equation. The simplest fourth-order linear PDE is the *biharmonic equation*

$$\nabla^4 u \equiv \nabla^2 \nabla^2 u = 0, \tag{13.2.1}$$

arising in the theory of plate bending and in slow viscous (Stokes) flow. Its theory parallels that of the Laplace equation, with representation theorems generalising those of Section 11.3 and a corresponding fundamental solution, computed in Problem 11.9. The boundary conditions for $\nabla^4 u = 0$ are richer, however, than those for $\nabla^2 u = 0$: the natural *clamped* condition prescribes both u and $\partial u/\partial n$ on $\partial\Omega$, while the natural *simply supported* condition prescribes u and $\nabla^2 u$. Each leads to a well-posed boundary-value problem, and the corresponding Green's functions are classical objects in elasticity theory.

Maxwell's equations. The Maxwell equations of classical electromagnetism,

$$\operatorname{rot}\mathbf{E} = -\partial_t \mathbf{B}, \quad \operatorname{rot}\mathbf{B} = \mu_0\varepsilon_0\,\partial_t \mathbf{E} + \mu_0\,\mathbf{J}, \quad \operatorname{div}\mathbf{E} = \rho/\varepsilon_0, \quad \operatorname{div}\mathbf{B} = 0,$$

constitute a system of *four first-order* linear PDEs for the six components of $\mathbf{E}$ and $\mathbf{B}$. As noted in the Introduction, elimination of one of the fields recovers a second-order wave equation for each component of the other; this reduction hides, however, important structural features of the system. The intrinsic geometry of Maxwell's equations is most transparent in the language of differential forms on Minkowski space, in which they collapse to the elegant pair $dF = 0$, $d\star F = \mu_0 J$. This formulation extends, with appropriate generalisations, to the Yang–Mills equations of non-abelian gauge theory and ultimately to the Einstein field equations of general relativity.

The equations of elasticity. The wave equation $\rho\,\mathbf{u}_{tt} = (\lambda+\mu)\,\operatorname{grad}\operatorname{div}\mathbf{u} + \mu\,\nabla^2\mathbf{u}$ derived in Section 3.3 for an isotropic elastic medium is one of the simplest *tensor-valued* PDE systems of mathematical physics. Anisotropic and inhomogeneous extensions, the

incorporation of geometric and material nonlinearities, and the imposition of contact and friction conditions on $\partial\Omega$ have generated a vast technical literature, and the mathematical theory of nonlinear elasticity remains one of the deepest and most demanding branches of applied analysis.

Further reading. For higher-order equations and the biharmonic, Selvadurai, *Partial Differential Equations in Mechanics 1: Fundamentals, Laplace's Equation, Diffusion Equation, Wave Equation* (Springer, 2000); for Maxwell's equations, Jackson, *Classical Electrodynamics* (Wiley, 3rd edn., 1998), and from a mathematical viewpoint, Sander, *Maxwell's Equations* (Princeton, 2017); for elasticity, Gurtin, *An Introduction to Continuum Mechanics* (Academic, 1981).

13.3 Probabilistic and Modern Functional-Analytic Perspectives

The analytic methods of this book stand within a broader spectrum of techniques. Two perspectives, profoundly different from those developed here, deserve particular mention.

Brownian motion and the heat equation. The heat equation $u_t = \tfrac{1}{2}\nabla^2 u$ on $\mathbb{R}^n$ is not merely a parabolic PDE; it is the *Kolmogorov backward equation* of the standard n-dimensional Brownian motion. The connection allows the Gaussian heat kernel (11.2.7) to be interpreted as the transition probability density of the Wiener process: $E(\mathbf{r}-\mathbf{r}',t)$ is the density at $\mathbf{r}$ at time t of a Brownian particle that started at $\mathbf{r}'$ at time zero. The solution formula (11.2.8) for the Cauchy problem,

$$u(\mathbf{r},t) = \mathbb{E}\big[f(\mathbf{r}+\sqrt{2t}\,\mathbf{Z})\big],$$

where $\mathbf{Z}$ is a standard n-dimensional Gaussian random variable, expresses the temperature at $(\mathbf{r},t)$ as the expected value of the initial datum, evaluated along Brownian trajectories that have had time t to spread out from their starting positions. This *Feynman–Kac formula*, in its general form, extends the representation to a Schrödinger-like equation with potential, the heat kernel being replaced by a Wiener integral over Brownian paths weighted by the exponential of the action.

The probabilistic perspective transforms the analysis of the heat and elliptic equations into the theory of stochastic processes, with phenomena—harmonic measure, Brownian motion on manifolds, recurrent versus transient diffusion—that have no immediate analogue in the purely analytic theory of this book. It also provides a constructive Monte-Carlo method for solving boundary-value problems that is in principle independent of the variational and Green's function methods developed here.

Sobolev spaces and weak solutions. The variational method of Chapter 12 required of admissible functions only that they have square-integrable gradient and satisfy the boundary conditions; classical smoothness was not invoked. The modern theory of PDEs makes this precise through the introduction of *Sobolev spaces*, complete normed spaces of functions whose distributional derivatives lie in L^p for some $p \geq 1$. The

Sobolev space $H^1(\Omega) = W^{1,2}(\Omega)$ consists of functions in $L^2(\Omega)$ whose first-order distributional partial derivatives also lie in $L^2(\Omega)$; its closed subspace $H^1_0(\Omega)$ consists of those that vanish, in the appropriate sense, on $\partial\Omega$.

The Dirichlet problem for $-\nabla^2$ admits, in this language, a clean existence theory: by the *Lax–Milgram theorem*, applied to the bilinear form $a(u,v) = \int \operatorname{grad} u \cdot \operatorname{grad} v \, dV$ on $H^1_0(\Omega)$, the equation $-\nabla^2 u = f$ with $u = 0$ on $\partial\Omega$ admits, for every $f \in L^2(\Omega)$, a unique *weak solution* $u \in H^1_0(\Omega)$. Higher regularity of u, when available, follows from the regularity theorems of Calderón–Zygmund and Schauder. The weak-solution viewpoint extends to evolution equations through Bochner spaces of vector-valued functions, and forms the standard framework of present-day PDE theory.

Viscosity solutions. A distinct extension to nonlinear elliptic and parabolic equations is provided by the theory of *viscosity solutions*, developed by Crandall and Lions. A continuous function u is a viscosity solution of $F(\mathbf{r}, u, \operatorname{grad} u, D^2 u) = 0$ if certain inequalities hold for every smooth test function that touches u from above or below. The notion is broad enough to admit non-differentiable solutions of fully nonlinear equations such as Hamilton–Jacobi–Bellman equations of optimal control theory, and at the same time strong enough to support uniqueness theorems. Together with the Sobolev framework, it constitutes the basis of the modern theory of fully nonlinear PDEs.

Further reading. For Sobolev spaces and weak solutions, the standard references are Evans, *Partial Differential Equations* (AMS, 2nd edn., 2010), and Brezis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations* (Springer, 2011); for the probabilistic viewpoint, Øksendal, *Stochastic Differential Equations* (Springer, 6th edn., 2003); for viscosity solutions, the user-friendly account of Crandall, Ishii and Lions, *User's guide to viscosity solutions of second-order partial differential equations* (Bull. AMS, 1992).

13.4 Numerical Methods

The variational characterisation of Chapter 12 furnishes not only an analytic tool but the foundation of the modern numerical treatment of partial differential equations. The Rayleigh–Ritz procedure—restricting the variational principle to a finite-dimensional subspace and minimising over the coefficients—generalises directly to the *finite-element method*, which chooses the trial functions as piecewise polynomials on a triangulation of the spatial domain. The resulting linear systems are sparse, the approximation error is controlled by the mesh size and the polynomial degree, and the method extends with manageable modifications to elliptic, parabolic and hyperbolic equations alike.

Two further families of numerical methods complement the finite-element approach. The *finite-difference method* discretises the differential operator itself, replacing partial derivatives by their lattice difference quotients on a grid; though conceptually simpler than the finite-element method, it is less flexible on irregular domains. The *spectral method* expands the unknown in a global basis of eigenfunctions—typically Fourier modes for periodic problems and Chebyshev polynomials for non-periodic ones—and

obtains exponentially small errors when the solution is smooth, at the cost of dense linear systems and reduced flexibility on complex domains.

A complete development of these methods is beyond the scope of the present book; they are mentioned here to indicate the bridge from the analytic methods of this volume to the computational methods on which contemporary applied mathematics rests. The connection of the finite-element method to the variational principle, and of the spectral method to the eigenfunction expansions of Chapters 8–10, should now be evident: the analytic and numerical approaches are not opposed but complementary, each illuminating the other.

Further reading. For finite elements, Brenner and Scott, *The Mathematical Theory of Finite Element Methods* (Springer, 3rd edn., 2008); for finite differences, LeVeque, *Finite Difference Methods for Ordinary and Partial Differential Equations* (SIAM, 2007); for spectral methods, Trefethen, *Spectral Methods in MATLAB* (SIAM, 2000).

13.5 Closing Remarks

The twelve preceding chapters have led the reader from the most elementary properties of partial differential equations through the classical analytical methods of mathematical physics to the modern analytic devices of Green's functions and variational principles. The emphasis throughout has been on *methods*—procedures by which the standard equations can be solved explicitly, and through which the structural properties of those solutions can be understood. The restriction to linear, second-order equations admitted, in exchange for its narrowness, a degree of completeness that would be impossible in a more comprehensive treatment.

The reader who has worked through the present book is, the author hopes, in a position to read with profit the modern literature sketched in the preceding sections. Three further texts deserve particular mention as natural successors. Evans, *Partial Differential Equations*, develops the modern weak-solution theory of linear and nonlinear equations within a unified analytic framework, and may be regarded as a sequel to the present book. Folland, *Introduction to Partial Differential Equations* (Princeton, 2nd edn., 1995), occupies an intermediate position, covering both the classical and the modern theory at a graduate level. John, *Partial Differential Equations* (Springer, 4th edn., 1982), takes a perspective close to that of the present volume but with a stronger emphasis on the qualitative theory of hyperbolic equations, and is particularly recommended to readers of physical inclination.

It remains, in closing, to express the author's hope that the methods, identities and viewpoints assembled here will serve the reader in whatever direction his or her further studies take. The classical theory developed in these pages is, after all, neither out-of-date nor closed; it is the foundation on which every modern extension is built, and the well from which the working mathematician or physicist continues to draw.

References

1. R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Vols. I and II, Wiley-Interscience, New York, 1962.
2. P. R. Garabedian, *Partial Differential Equations*, Wiley, New York, 1964.
3. N. S. Koshlyakov, M. M. Smirnov, and E. B. Gliner, *Differential Equations of Mathematical Physics*, North-Holland, Amsterdam, 1964.
4. P. M. Morse and H. Feshbach, *Methods of Theoretical Physics*, Vols. I and II, McGraw-Hill, New York, 1953.
5. T. Myint-U, *Partial Differential Equations of Mathematical Physics*, Elsevier, New York, 1973.
6. I. N. Sneddon, *Elements of Partial Differential Equations*, McGraw-Hill, New York, 1957.
7. A. Sommerfeld, *Partial Differential Equations in Physics*, Academic Press, New York, 1949.
8. A. N. Tikhonov and A. A. Samarskii, *Equations of Mathematical Physics*, Pergamon Press, Oxford, 1963.
9. V. S. Vladimirov, *Equations of Mathematical Physics*, Marcel Dekker, New York, 1971.
10. L. C. Evans, *Partial Differential Equations*, 2nd edn., Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, RI, 2010.

Index

A

acoustics 1
action functional 194
adjoint operator 71
asymptotic eigenvalues 129
auxiliary conditions 53

B

Bessel function 10, 75
Bessel functions 134
biharmonic equation 205
boundary conditions 53
boundary-value problem 54
Brownian motion 206

C

canonical form 36
  elliptic 40
  hyperbolic 38
Cauchy data 63
Cauchy problem 62
Cauchy–Kowalevski theorem 65
characteristic curves 38
characteristic equation 38
  ODE 8
completeness 134
conservation law 203
curl 6

D

d'Alembert formula 66
d'Alembertian operator 2
differential equation 1
diffusion equation 2
Dirac delta
  multidimensional 179
  sequential definition 177
  sifting property 178
Dirac equation 3
Dirichlet condition 53
Dirichlet energy 192
Dirichlet variational problem 192
Dirichlet's principle 193
divergence 5
Duhamel formula 189

E

elasticity equations 205
electrodynamics 1
electromagnetic theory 1
electrostatics 1
energy conservation
  wave equation 194
Euler–Lagrange equation 192
Euler–Lagrange equations 129

F

Feynman–Kac formula 206
finite-difference method 207
finite-element method 207
fluid mechanics 3
Fourier series 9
  cosine 9
  sine 9
Fourier's law 30
fundamental solution 180
  heat operator 181
  Laplacian 180
  wave operator 181

G

Gauss's theorem 6
Gaussian heat kernel 181
gradient 5
gravitation, classical 1
Green's function 182
  and heat kernel 185
Green's representation theorem 183

H

Hadamard's example 55
Hadamard, J. 54
Hamilton's principle 193
harmonic function 33
heat kernel 181
heat-conduction equation 2
Heaviside step function 179
Helmholtz equation 2, 122
homogeneous PDE 15
Hooke's law 28
Huygens principle 182
hydrodynamics 1

I

inhomogeneous PDE 15
initial conditions 53
initial–boundary-value problem 54
initial-value problem 3, 54, 62
integrable equation 204

K

KdV equation 204
kinetic theory of gases 2
Klein–Gordon equation 3

L

Lagrangian density 194
Lamé constants 28
Laplace equation 1, 33
Laplacian operator 6
  cylindrical coordinates 6
  spherical coordinates 6
Law of Inertia 44
Legendre equation 124
Legendre polynomials 10, 137
linear operator 18
linear PDE 14
load vector 197

M

magnetostatics 1
master equation 1
mathematical physics 1
maximum principle 193
Maxwell equations 3, 205
method
  Riemann's 70
method of characteristics 65
method of images 184
Minkowski space 2

N

Navier–Stokes equations 204
Neumann condition 54
neutron transport 2
nonlinear equation 3
nonlinear PDE 14

O

operator
  linear differential 17
order
  of a PDE 14
ordinary differential equation 7
  constant coefficients 8
orthogonality
  of Sturm–Liouville eigenfunctions 126

P

partial derivative 13
partial differential equation 1, 14
Poisson equation 1
Poisson integral formula 186

Q

quantum mechanics 2
quasilinear PDE 14

R

Rayleigh quotient 197
Rayleigh–Ritz method 196
reaction-diffusion equation 204
retarded potential 182
Riemann function 72
Riemann's method 70
Robin condition 54

S

Schrödinger equation 2
self-adjoint 71
semigroup property 189
shock wave 203
sine-Gordon equation 204
Sobolev space 206
soliton 204
special relativity 2
spectral method 207
stiffness matrix 197
Stokes's theorem 6
Sturm–Liouville equation 124
Sturm–Liouville expansion theorem 126
Sturm–Liouville problem ix
superposition principle ix, 19

T

telegrapher's equation 73
telegraphers' equation 1, 73
thermal energy 195
total energy 194
trial functions 197
Tricomi equation 46
type
  of a PDE 35

U

uniqueness theorem
  Dirichlet problem 193
  heat equation 196
  wave equation 195

V

variational method
  Sturm–Liouville 127
viscosity solution 207

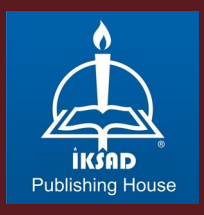
W

wave equation 1
  one-dimensional 25
weak solution 203, 206
well-posedness 55
Wronskian 8

About the Author

Göksel Daylan Esmer received her undergraduate, master's and doctoral degrees in physics, with a specialisation in mathematical physics and partial differential equations. She is a faculty member of the Department of Physics at Istanbul University, where, over the course of more than two decades, she has taught the graduate and advanced-undergraduate courses on mathematical physics from which the present material is drawn.

Her research interests include the analytical theory of linear partial differential equations, special functions of mathematical physics, the Sturm–Liouville theory and its applications to boundary-value problems on cylindrical and spherical domains, and the variational and energy methods at the interface between classical analysis and modern PDE theory. The single-author volume presented here grew out of an earlier joint manuscript co-authored with Professor Şehsuvar Zebitay, in continuous classroom use for many years, and represents a revised and substantially expanded treatment.


İKSAD
Publishing House