Introduction to Partial Differential Equations

Lecture Notes for Math 4163

Weinan Wang
Department of Mathematics, University of Oklahoma
ww@ou.edu

April 2, 2026

Contents

Preface vii

I Foundations 1

1 What a Partial Differential Equation Is 3
- 1.1 Partial differential equations . . . 3
- 1.2 Linearity and superposition . . . 4
- 1.3 Where these equations come from . . . 5
- 1.4 Classification of second-order equations . . . 6
- 1.5 What we will ask, and the plan . . . 7

2 First-Order Equations and Characteristics 9
- 2.1 The transport equation . . . 9
- 2.2 The general first-order linear equation . . . 10
- 2.3 Variable coefficients . . . 10
- 2.4 The method of characteristics, summarized . . . 11
- 2.5 A glimpse of nonlinearity: shock formation . . . 12

3 Initial Conditions, Boundary Conditions, and Well-Posedness 15
- 3.1 Initial conditions . . . 15
- 3.2 Boundary conditions . . . 15
- 3.3 Boundary conditions for the three types . . . 16
- 3.4 Well-posedness . . . 17
- 3.5 Looking ahead . . . 18

II Waves and Diffusions 21

4 The Wave Equation 23
- 4.1 The general solution by factoring . . . 23
- 4.2 D'Alembert's formula . . . 24
- 4.3 Causality: domains of dependence and influence . . . 24
- 4.4 The energy method . . . 25
- 4.5 The inhomogeneous wave equation . . . 26
- 4.6 The wave equation on an interval, briefly . . . 26

5 The Diffusion Equation **29**
5.1 Derivation from conservation and Fourier's law . . . 29
5.2 The maximum principle . . . 30
5.3 Uniqueness by the energy method . . . 31
5.4 Two equations, two personalities . . . 32

6 The Diffusion Equation on the Whole Line **35**
6.1 A special solution from scaling . . . 35
6.2 Solving the initial-value problem by convolution . . . 36
6.3 The error function . . . 36
6.4 Smoothing and infinite propagation speed . . . 36
6.5 Looking ahead to the Fourier transform . . . 37

7 Reflections and Sources **39**
7.1 The half-line: reflection . . . 39
7.2 The method of images . . . 40
7.3 Inhomogeneous equations and Duhamel's principle . . . 40

III Fourier Series and Separation of Variables **43**

8 Separation of Variables **45**
8.1 The method on the diffusion equation . . . 45
8.2 The eigenvalue problem . . . 46
8.3 Superposition and the Fourier series . . . 46
8.4 A fully worked example . . . 47
8.5 Other boundary conditions, other series . . . 47
8.6 Inhomogeneous boundary conditions . . . 48
8.7 Separation of variables for the wave equation . . . 48

9 Fourier Series **51**
9.1 The Fourier coefficients via orthogonality . . . 51
9.2 Even, odd, and periodic extensions . . . 52
9.3 The complex form . . . 52
9.4 Convergence: in what sense, and how fast . . . 53
9.5 Differentiating and integrating Fourier series . . . 53
9.6 Completeness . . . 54
9.7 The Gibbs phenomenon . . . 54
9.8 Inhomogeneous boundary conditions . . . 54

10 Sturm–Liouville Theory **57**
10.1 The Sturm–Liouville problem . . . 57
10.2 The main theorem . . . 57
10.3 Why orthogonality holds . . . 58
10.4 The Rayleigh quotient . . . 58
10.5 Eigenfunction expansions and generalized Fourier series . . . 59
10.6 Singular problems and the special functions . . . 59

IV Transform Methods 63

11 The Fourier Transform 65
11.1 From series to transform . . . 65
11.2 The two properties that matter for PDEs . . . 65
11.3 Shifting, scaling, and other rules . . . 66
11.4 The Gaussian and its transform . . . 66
11.5 Plancherel's theorem . . . 67
11.6 Solving the diffusion equation on the line . . . 67
11.7 Solving the wave equation on the line . . . 68
11.8 The transform method in general . . . 68

12 The Laplace Transform 71
12.1 Definition and basic transforms . . . 71
12.2 Linearity and the derivative rule . . . 71
12.3 The shifting theorems and the Heaviside function . . . 72
12.4 The Dirac delta and impulse response . . . 72
12.5 The convolution theorem . . . 73
12.6 Solving initial-value problems for ODEs . . . 73
12.7 Inversion . . . 74
12.8 Application to PDEs . . . 74

V Laplace's Equation and Computation 77

13 Laplace's Equation and Harmonic Functions 79
13.1 Harmonic functions and where they arise . . . 79
13.2 The mean value property . . . 79
13.3 The maximum principle for harmonic functions . . . 80
13.4 Separation of variables in a rectangle . . . 81
13.5 Laplace's equation in a disk and Poisson's formula . . . 81
13.6 Green's functions, in brief . . . 82

14 Numerical Methods 85
14.1 Finite-difference approximations . . . 85
14.2 Truncation error and Big-O . . . 85
14.3 An explicit scheme for the diffusion equation . . . 86
14.4 Stability and the CFL condition . . . 87
14.5 Implicit schemes and Crank–Nicolson . . . 87
14.6 A scheme for the wave equation . . . 88

A Reference Tables and Formulas 91
A.1 Trigonometric and orthogonality identities . . . 91
A.2 Fourier series . . . 92
A.3 Fourier transform . . . 92
A.4 Laplace transform . . . 92
A.5 The three model equations at a glance . . . 93

B A Guide to Choosing a Method **95**
B.1 First, classify . . . 95
B.2 Then, match the method to the domain . . . 95
B.3 Handling complications . . . 96
B.4 When no formula exists . . . 96
B.5 A worked decision . . . 96

C Solution Sketches **97**

Further Reading **105**

Preface

These notes accompany a one-semester undergraduate introduction to partial differential equations, the course numbered Math 4163 at the University of Oklahoma. They follow the spirit and notation of Strauss's *Partial Differential Equations: An Introduction*, which remains the standard text for a course at this level, and they are meant to be read alongside it or on their own.

Who this is for. The notes assume multivariable calculus (partial derivatives, multiple integrals, the divergence theorem) and a first course in ordinary differential equations. Linear algebra is helpful, especially for the later chapters on eigenfunction expansions, but the needed ideas are recalled as they arise. No prior exposure to PDEs is assumed.

What a PDE course is about. A partial differential equation relates an unknown function of several variables to its partial derivatives. Three equations — the wave equation, the diffusion (heat) equation, and Laplace's equation — recur throughout the subject, both because they model fundamental physical processes and because they are the prototypes of the three great classes of second-order equations: hyperbolic, parabolic, and elliptic. Much of this course consists of learning what questions to ask of each equation (does a solution exist? is it unique? does it depend continuously on the data?) and the handful of powerful techniques — characteristics, energy methods, separation of variables, Fourier and Laplace transforms — that answer those questions.

Theory and computation, in balance. The notes hold two aims in tension throughout. One is *rigor*: PDE theory is built on proofs of uniqueness, the maximum principle, and convergence, and we present these carefully, because they are what distinguish a solution we can trust from a formula we merely wrote down. The other is *computation*: a student must be able to actually solve the heat equation on an interval, sum a Fourier series, invert a Laplace transform, and run a finite-difference scheme. Each chapter develops both, and the exercises mix the two deliberately.

On starred sections. Sections and results marked with a star ($\star$) are more demanding or more peripheral and may be skipped on a first reading without loss of continuity. They are included for students who want the underlying justification or a glimpse of where the subject leads.

Exercises. Each chapter ends with exercises, some computational and some proof-based, in proportions matching the text. Solution sketches are collected in an appendix; they indicate the intended approach rather than a complete write-up, which the reader should supply.

Notation. Subscripts denote partial derivatives: $u_t = \partial u/\partial t$, $u_{xx} = \partial^2 u/\partial x^2$. The diffusion equation is written $u_t = ku_{xx}$ with diffusion constant $k > 0$; the wave equation $u_{tt} = c^2 u_{xx}$ with wave speed c; Laplace's equation $\Delta u = 0$ with $\Delta = \partial_{xx} + \partial_{yy} + \cdots$ the Laplacian. A spatial interval has length l, so $0 < x < l$. The Fourier transform of f is written $\hat{f}(\xi)$ and the Laplace transform $\mathcal{L}\{f\}(s)$, with conventions fixed in Chapters 11 and 12.

Part I

Foundations

Chapter 1

What a Partial Differential Equation Is

This chapter sets the stage. We say precisely what a partial differential equation is, introduce the three equations that will occupy us for the rest of the course, see briefly where they come from, and learn to sort second-order equations into the three great classes — hyperbolic, parabolic, and elliptic — that organize the whole subject. Nothing here is difficult, but the vocabulary and the classification are worth fixing carefully, because every later chapter speaks this language.

1.1 Partial differential equations

An *ordinary differential equation* (ODE) relates a function of one variable to its derivatives. A *partial differential equation* (PDE) relates a function of *several* variables to its *partial* derivatives. The unknown is a function u of two or more independent variables — typically a spatial variable x (sometimes several, x, y, z) and a time variable t — and the equation is a relation among u and its partial derivatives that is required to hold throughout some region.

We abbreviate partial derivatives by subscripts:

$$u_t = \frac{\partial u}{\partial t}, \quad u_x = \frac{\partial u}{\partial x}, \quad u_{xx} = \frac{\partial^2 u}{\partial x^2}, \quad u_{xt} = \frac{\partial^2 u}{\partial x\,\partial t},$$

and so on. With this notation the three equations at the heart of the course are compact to write:

$$\text{the } \textit{wave equation} \qquad u_{tt} = c^2 u_{xx}, \tag{1.1}$$

$$\text{the } \textit{diffusion (heat) equation} \qquad u_t = k u_{xx}, \tag{1.2}$$

$$\textit{Laplace's equation} \qquad u_{xx} + u_{yy} = 0. \tag{1.3}$$

Here $c > 0$ is a wave speed and $k > 0$ a diffusion constant. The wave and diffusion equations involve a time variable t and describe how a quantity evolves; Laplace's equation involves two spatial variables and describes a steady state, a configuration that has stopped changing.

Definition 1.1 (Order)**.** The *order* of a PDE is the order of the highest derivative that appears. All three equations (1.1)–(1.3) are of *second order*. The transport equation $u_t + cu_x = 0$, which we study in Chapter 2, is of *first order*.

A *solution* of a PDE on a region is a function that, when substituted into the equation along with its partial derivatives, satisfies it at every point of the region. Unlike an ODE, whose general solution involves arbitrary *constants*, the general solution of a PDE typically involves arbitrary *functions*, which is one reason PDEs support such a rich variety of behavior.

Example 1.2 (Arbitrary functions in the general solution)**.** Consider $u_{xy} = 0$ for a function $u(x, y)$. Integrating in y gives $u_x =$ (a function of x alone), say $u_x = f'(x)$; integrating in x gives

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

where f and g are *arbitrary* differentiable functions. Two arbitrary functions, not two arbitrary constants, parametrize the solutions. We will see the same phenomenon in d'Alembert's solution of the wave equation (Chapter 4), where the general solution is a sum of two arbitrary travelling waves.

Example 1.3 (Reading a PDE as a condition on a surface)**.** It helps to picture the solution $u(x, y)$ as a surface, the graph $z = u(x, y)$ over the (x, y) plane. A PDE is then a condition on the slopes and curvatures of this surface. Laplace's equation $u_{xx} + u_{yy} = 0$, for instance, says the two principal curvatures sum to zero: where the surface bends up in the x-direction it must bend down equally in the y-direction, so a harmonic graph is a saddle-like surface with no purely-bowl-shaped points. This geometric reading, which we develop in Chapter 13, already suggests that harmonic functions cannot have interior maxima — a fact we will prove as the maximum principle.

1.2 Linearity and superposition

The single most important structural distinction among PDEs is whether they are *linear*. Write the equation as $\mathcal{L}u = g$, where $\mathcal{L}$ is the operation performed on the unknown. For the diffusion equation, $\mathcal{L}u = u_t - ku_{xx}$ and $g = 0$; for the diffusion equation with a source, g is the given source function.

Definition 1.4 (Linear operator, linear equation)**.** The operator $\mathcal{L}$ is *linear* if for all functions u, v and all constants a, b,

$$\mathcal{L}(au + bv) = a\,\mathcal{L}u + b\,\mathcal{L}v.$$

The equation $\mathcal{L}u = g$ is *linear* if $\mathcal{L}$ is a linear operator. It is *homogeneous* if $g = 0$ and *inhomogeneous* otherwise. A PDE that is not linear is *nonlinear*.

All three model equations are linear and homogeneous. The operator $\mathcal{L}u = u_t - ku_{xx}$ is linear because differentiation is linear: $(au + bv)_t = au_t + bv_t$ and similarly for u_{xx}. By contrast, the equation $u_t + uu_x = 0$ (the inviscid Burgers equation) is nonlinear, because the term uu_x is not linear in u — doubling u quadruples that term.

A finer distinction within the nonlinear world is sometimes useful. An equation is *semilinear* if the highest-order derivatives appear linearly (only lower-order terms are nonlinear), *quasilinear* if the highest-order derivatives appear linearly but with coefficients depending on u and its lower derivatives, and *fully nonlinear* otherwise. Burgers' equation is quasilinear (the coefficient of u_x is u itself); the eikonal equation $u_x^2 + u_y^2 = 1$ is fully nonlinear. We will work almost entirely with linear equations, where the following principle holds.

Linearity buys us the *superposition principle*, which is the foundation of nearly every solution method in this course.

Proposition 1.5 (Superposition)**.** *If $\mathcal{L}$ is linear and $u_1, u_2, \ldots$ each solve the homogeneous equation $\mathcal{L}u = 0$, then every linear combination $\sum_n c_n u_n$ also solves it (granting convergence when the sum is infinite). More generally, if $\mathcal{L}u_1 = g_1$ and $\mathcal{L}u_2 = g_2$, then $\mathcal{L}(c_1u_1 + c_2u_2) = c_1g_1 + c_2g_2$.*

Proof. Immediate from the definition of linearity: $\mathcal{L}(\sum_n c_n u_n) = \sum_n c_n \mathcal{L}u_n = \sum_n c_n \cdot 0 = 0$. □

Superposition is what lets us build complicated solutions from simple ones. In the method of separation of variables (Chapter 8) we find infinitely many simple "product" solutions of the heat equation and then superpose them — with coefficients chosen to match the initial data — into a Fourier series. The legitimacy of that construction rests entirely on Proposition 1.5, and it is exactly why linear equations are so much more tractable than nonlinear ones.

Remark 1.6 (The structure of the solution set)**.** Superposition gives the solution set of a linear equation an algebraic structure familiar from linear algebra. The solutions of a homogeneous linear equation $\mathcal{L}u = 0$ form a *vector space*: sums and scalar multiples of solutions are again solutions. The solutions of an inhomogeneous equation $\mathcal{L}u = g$ form an *affine space*: any two differ by a solution of the homogeneous equation, so the general solution is one particular solution plus the general homogeneous solution — exactly as for linear ODEs and linear systems $Ax = b$. This analogy, with $\mathcal{L}$ playing the role of a matrix, is worth keeping in mind; the eigenvalue problems of Chapter 8 push it considerably further.

1.3 Where these equations come from

The three equations are not arbitrary; each arises from a simple physical principle. We sketch the derivations briefly, both for motivation and because the physical reading of each term aids intuition later. Fuller derivations appear in the dedicated chapters.

Conservation and the transport equation. Imagine a substance of concentration $u(x,t)$ carried along a pipe at constant speed c. In a time Δt the entire profile shifts a distance $c\,\Delta t$ to the right, so $u(x, t + \Delta t) = u(x - c\,\Delta t, t)$. Expanding to first order in Δt gives the *transport equation*

$$u_t + cu_x = 0.$$

It is the simplest expression of *conservation*: the substance is neither created nor destroyed, only moved. We solve it geometrically in Chapter 2.

The vibrating string and the wave equation. Consider a taut elastic string, displaced slightly from rest, with $u(x,t)$ its vertical displacement. Newton's second law applied to a small element of string, under tension T and with mass density ρ, balances the transverse acceleration ρu_{tt} against the net transverse component of tension, which for small displacements is Tu_{xx}. The result is the wave equation $u_{tt} = c^2 u_{xx}$ with $c^2 = T/\rho$. The constant c is the speed at which disturbances travel along the string — a fact we will make precise in Chapter 4. Notice the appearance of *two* time derivatives: Newton's law involves acceleration, u_{tt}, which is why the wave equation is second order in time and will require two initial conditions.

Heat flow and the diffusion equation. Let $u(x,t)$ be the temperature in a thin rod. Two physical laws combine. *Conservation of heat* says the rate of change of heat stored in a segment equals the net heat flux across its ends. *Fourier's law* says heat flows from hot to cold at a rate proportional to the temperature gradient, with flux $-Ku_x$. Combining them — the rate of change of stored heat, proportional to u_t, equals the divergence of the flux, proportional to u_{xx} — yields the diffusion equation $u_t = ku_{xx}$. The same equation governs the diffusion of a chemical concentration, with Fourier's law replaced by Fick's law; hence the two names, heat equation and diffusion equation, for the same object. Only *one* time derivative appears, because conservation involves the rate of change of a stored quantity, not its acceleration — which is why the diffusion equation is first order in time and needs only one initial condition.

Steady states and Laplace's equation. If the temperature in a region settles to a steady state, no longer changing in time, then $u_t = 0$, and the diffusion equation $u_t = k\Delta u$ reduces to $\Delta u = 0$ — Laplace's equation. Its solutions, the *harmonic functions*, describe equilibrium temperature distributions, electrostatic potentials in charge-free regions, and steady incompressible flows. We study them in Chapter 13.

Remark 1.7 (One process, three equations)**.** It is striking that diffusion gives rise to all three model equations at once. The full time-dependent process is the diffusion equation; its long-time equilibrium is Laplace's equation; and if the storage term had been an acceleration rather than a rate (as for an elastic rather than a diffusive medium) it would have been the wave equation. The three equations are thus not unrelated curiosities but the three faces of how a quantity stored in a medium responds to imbalance: spread (parabolic), settle (elliptic), or oscillate (hyperbolic). The classification of Section 1.4 makes this trichotomy precise.

1.4 Classification of second-order equations

The three model equations look different, and they behave differently: waves propagate and persist, diffusions smooth out and decay, and harmonic functions sit in placid equilibrium. This difference is not accidental. Every linear second-order PDE in two variables falls into one of three types, and the model equations are the prototypes.

Consider the general linear second-order equation in two variables,

$$a\,u_{xx} + 2b\,u_{xy} + c\,u_{yy} + (\text{lower-order terms}) = 0, \tag{1.4}$$

where a, b, c are constants (or, more generally, functions of the variables). The *type* is determined by the sign of the *discriminant* $b^2 - ac$, in exact analogy with the conic sections $ax^2 + 2bxy + cy^2 = \text{const}$, which are ellipses, parabolas, or hyperbolas according to the same sign.

Definition 1.8 (Classification)**.** The equation (1.4) is

- *elliptic* where $b^2 - ac < 0$,
- *parabolic* where $b^2 - ac = 0$,
- *hyperbolic* where $b^2 - ac > 0$.

Let us classify the three model equations. For **Laplace's equation** $u_{xx} + u_{yy} = 0$ we read off $a = 1$, $b = 0$, $c = 1$, so $b^2 - ac = -1 < 0$: *elliptic*. For the **wave equation** $u_{tt} - c^2 u_{xx} = 0$, treating (x, t) as the two variables with u_{tt} playing the role of u_{yy}, we have leading coefficients 1 (on u_{tt}) and $-c^2$ (on u_{xx}) and no mixed term, giving discriminant $0 - (1)(-c^2) = c^2 > 0$: *hyperbolic*. For the **diffusion equation** $u_t - ku_{xx} = 0$, the only second-order term is u_{xx}, so in the (x, t) variables $a = -k$ (coefficient of u_{xx}), $b = 0$, and $c = 0$ (no u_{tt}), whence $b^2 - ac = 0$: *parabolic*. The three prototypes thus exhibit the three types, and this is why they are studied as representatives.

Where the names come from: characteristics. The deeper meaning of the classification lies in the *characteristic curves*, the curves along which the equation degenerates — along which, loosely, information propagates. For (1.4) the characteristic directions satisfy $a\,(dy)^2 - 2b\,dx\,dy + c\,(dx)^2 = 0$, a quadratic whose discriminant is again $b^2 - ac$. A hyperbolic equation has *two* real families of characteristics (the quadratic has two real roots) — these are the lines $x \pm ct = \text{const}$ for the wave equation, along which signals travel. A parabolic equation has *one* family (a repeated root) — the lines $t = \text{const}$ for the heat equation, reflecting its instantaneous spreading. An elliptic equation has *no* real characteristics (complex roots) — which is why Laplace's equation has no preferred directions of propagation and its solutions are so smooth and rigid. The count of real characteristic families — two, one, or zero — is the geometric content of the discriminant's sign.

Remark 1.9 (Why the classification matters)**.** The type of an equation predicts its qualitative behavior and dictates what kind of data make it well-posed. Hyperbolic equations propagate signals at finite speed along characteristics, preserve features (a sharp pulse stays sharp), and are naturally paired with *initial* data. Parabolic equations smooth data instantly, propagate "information" at infinite speed, and decay toward equilibrium; they take initial data in time and boundary data in space. Elliptic equations have no time variable, describe equilibrium, are exceptionally smooth, and take *boundary* data all around a region. Much of the course is the elaboration of these three personalities, and a recurring first question about any new equation is simply: which type is it?

Example 1.10 (Classifying a general constant-coefficient equation)**.** Classify $3u_{xx} + 2u_{xy} + u_{yy} = 0$. Matching to (1.4), $a = 3$, $2b = 2$ so $b = 1$, $c = 1$. Then $b^2 - ac = 1 - 3 = -2 < 0$: elliptic. By contrast $u_{xx} + 4u_{xy} + u_{yy} = 0$ has $a = 1$, $b = 2$, $c = 1$, $b^2 - ac = 4 - 1 = 3 > 0$: hyperbolic. The presence of a mixed-derivative term does not by itself decide the type; only the sign of the discriminant does, and a change in a single coefficient can flip an equation from one type to another.

Example 1.11 (A variable-type equation*)**.** When a, b, c depend on position, an equation can change type from region to region. The *Tricomi equation* $u_{xx} + x\,u_{yy} = 0$ has $a = 1$, $b = 0$, $c = x$, so $b^2 - ac = -x$: it is elliptic where $x > 0$, parabolic on the line $x = 0$, and hyperbolic where $x < 0$. Such mixed-type equations arise in transonic flow, where the governing equation is elliptic in subsonic regions and hyperbolic in supersonic ones. We will not study them, but the example shows that the classification is local, a property of the equation at each point.

1.5 What we will ask, and the plan

Writing down a PDE is the easy part; the science is in understanding its solutions. Three questions recur, and together they constitute what it means for a problem to be *well-posed*, a notion we treat carefully in Chapter 3:

1. **Existence:** does a solution exist (for given initial and/or boundary data)?
2. **Uniqueness:** is it the only one?
3. **Stability:** does it depend continuously on the data, so that a small change in the data produces only a small change in the solution?

A problem that has a unique solution depending continuously on its data is well-posed; one that fails any of the three is ill-posed, and ill-posed problems — though they arise in real applications — demand special care.

The plan of the notes follows the structure of the subject. Part I (this chapter and the next two) lays foundations: classification, the method of characteristics for first-order equations, and the apparatus of initial and boundary conditions and well-posedness. Part II treats the wave and diffusion equations directly — d'Alembert's formula, energy methods, the heat kernel, reflections, and sources. Part III develops separation of variables and the Fourier series it requires, the workhorse technique for problems on a finite interval, together with the Sturm–Liouville theory that explains why it works. Part IV introduces the Fourier and Laplace transforms for problems on unbounded domains. Part V returns to Laplace's equation and the harmonic functions, and closes with the numerical methods used when no formula is available.

Exercises

Exercise 1.1 (*Order and linearity*). For each equation state its order, and whether it is linear or nonlinear; if linear, whether homogeneous or inhomogeneous. (a) $u_t = ku_{xx}$; (b) $u_t + uu_x = 0$; (c) $u_{tt} = c^2 u_{xx} + \sin t$; (d) $u_x^2 + u_y^2 = 1$; (e) $u_{xx} + u_{yy} = u$.

Exercise 1.2 (*Semilinear, quasilinear, fully nonlinear*). Classify each nonlinear equation as semilinear, quasilinear, or fully nonlinear: (a) $u_t + uu_x = u_{xx}$; (b) $u_t = u_{xx} + u^3$; (c) $u_x u_y = 1$; (d) $u_{tt} = (1 + u_x^2)u_{xx}$.

Exercise 1.3 (*Verifying solutions*). Verify that each function solves the indicated equation. (a) $u(x,t) = e^{-k\xi^2 t}\sin(\xi x)$ solves $u_t = ku_{xx}$ for any constant ξ. (b) $u(x,t) = f(x - ct) + g(x + ct)$ solves $u_{tt} = c^2 u_{xx}$ for any twice-differentiable f, g. (c) $u(x,y) = x^2 - y^2$ solves $u_{xx} + u_{yy} = 0$.

Exercise 1.4 (*Superposition fails for nonlinear equations*). Show that $u_1(x,t) = 1$ and any nonzero constant both solve the nonlinear equation $u_t + uu_x = 0$, but exhibit a linear combination of two solutions of $u_t + uu_x = 0$ that fails to solve it. Why does the superposition principle not apply?

Exercise 1.5 (*Particular plus homogeneous*). Verify that $u(x,y) = \frac{1}{2}x^2$ solves the inhomogeneous equation $u_{xx} + u_{yy} = 1$. Using the affine structure of the solution set, describe the general solution in terms of this particular solution and the harmonic functions.

Exercise 1.6 (*Arbitrary functions*). Find the general solution of $u_{xy} = 0$ (done in the text), and then of $u_{xx} = 0$, regarding u as a function of x and y. How many arbitrary functions appear in each?

Exercise 1.7 (*Classifying equations*). Classify each as elliptic, parabolic, or hyperbolic: (a) $u_{xx} + u_{yy} = 0$; (b) $u_{tt} - 4u_{xx} = 0$; (c) $u_t - u_{xx} = 0$; (d) $3u_{xx} + 2u_{xy} + u_{yy} = 0$; (e) $u_{xx} + 4u_{xy} + u_{yy} = 0$.

Exercise 1.8 (*The discriminant and conics*). Explain the analogy between the classification of (1.4) by the sign of $b^2 - ac$ and the classification of the conic $ax^2 + 2bxy + cy^2 = 1$ as an ellipse, parabola, or hyperbola. Why is the same algebraic quantity decisive in both?

Exercise 1.9 (*Counting characteristics*). For each type — hyperbolic, parabolic, elliptic — state how many real families of characteristic curves the equation has, and name them for the wave and heat equations. How does this count relate to the sign of the discriminant?

Exercise 1.10 (*A mixed-type equation*). For the equation $y\, u_{xx} + u_{yy} = 0$, determine the regions of the plane where it is elliptic, parabolic, and hyperbolic.

Exercise 1.11 (*Reducing to a steady state*). Starting from the diffusion equation in two space variables $u_t = k(u_{xx} + u_{yy})$, explain why a time-independent (steady-state) solution must satisfy Laplace's equation. What physical situation does such a steady state describe?

Chapter 2

First-Order Equations and Characteristics

Before the second-order equations that occupy most of the course, we treat first-order equations, because they are simpler and because they introduce an idea — the *method of characteristics* — that recurs when we solve the wave equation. The guiding insight is geometric: a first-order PDE can be solved by reducing it to a family of ordinary differential equations along special curves, the characteristics, on which the PDE collapses to a statement about how the unknown changes along the curve.

2.1 The transport equation

The simplest PDE of all is the constant-coefficient *transport equation*

$$u_t + c\,u_x = 0, \qquad -\infty < x < \infty,\ t > 0, \tag{2.1}$$

with c a constant. We met it in Section 1.3 as the expression of a substance carried along at speed c. Let us solve it properly.

The key observation is that the left-hand side of (2.1) is a *directional derivative*. Consider a straight line in the (x,t) plane moving at speed c, namely $x = ct + x_0$ for a constant x_0, and ask how u changes as we travel along it. Writing $u(t) = u(ct + x_0, t)$ and differentiating with the chain rule,

$$\frac{d}{dt}\,u(ct + x_0, t) = u_x \cdot c + u_t = u_t + cu_x = 0.$$

The transport equation says exactly that u does not change along such a line. These lines are the *characteristics* of (2.1).

Definition 2.1 (Characteristics of the transport equation)**.** The *characteristic curves* of $u_t + cu_x = 0$ are the lines $x - ct = \text{const}$ in the (x,t) plane. Along each, the solution u is constant.

Since u is constant on each line $x - ct = x_0$, its value there equals its value at $t = 0$, where $x = x_0$. If the initial data is $u(x,0) = \phi(x)$, then on the characteristic through $(x_0, 0)$ the solution takes the value $\phi(x_0) = \phi(x - ct)$. We have proved:

Proposition 2.2 (Solution of the transport equation)**.** *The solution of $u_t + cu_x = 0$ with $u(x,0) = \phi(x)$ is*

$$u(x,t) = \phi(x - ct). \tag{2.2}$$

The solution is the initial profile ϕ translated to the right at speed c, rigid and undistorted. This is the prototype of *wave propagation*, and it already shows the hyperbolic personality: a feature in the initial data (a bump, a discontinuity) travels at finite speed and keeps its shape. Notice also that (2.2) makes sense even when ϕ is not differentiable — a transported step function, say — which is a first hint that the notion of "solution" may sometimes be usefully relaxed.

Example 2.3 (A travelling pulse)**.** For $u_t + 2u_x = 0$ with $u(x,0) = e^{-x^2}$, the solution is $u(x,t) = e^{-(x-2t)^2}$: a Gaussian bump sliding right at speed 2, unchanged in shape. At $t = 3$ it is centered at $x = 6$.

2.2 The general first-order linear equation

The same geometry handles the general constant-coefficient first-order linear equation

$$a\,u_x + b\,u_t = 0, \tag{2.3}$$

with a, b constants. This is again a directional-derivative statement: the left side is the derivative of u in the direction of the vector (a, b) in the (x, t) plane. The equation says u is constant in that direction, so the characteristics are the lines parallel to (a, b), namely $bx - at = \text{const}$, and $u(x,t) = \phi(bx - at)$ for an arbitrary function ϕ fixed by the data. Dividing (2.3) by b recovers the transport form with speed $c = a/b$, so nothing new happens — but the vector picture is worth having, because it generalizes: the characteristics are always the curves tangent to the direction in which the equation differentiates u.

When a lower-order term is present, $au_x + bu_t + d\,u = 0$, the solution is no longer constant along characteristics but *decays* (or grows) along them: on the characteristic, $\frac{d}{ds}u = -\frac{d}{b}u$ (parametrizing by t), an ODE giving exponential change. This is the prototype of how zeroth-order terms enter — they make the transported profile shrink or swell as it travels, without changing the characteristics themselves.

2.3 Variable coefficients

The geometric method extends at once to variable speed. Consider

$$u_t + c(x)\,u_x = 0, \qquad u(x,0) = \phi(x). \tag{2.4}$$

Now the characteristics are no longer straight, but they are still the curves along which u is constant. Along a curve $x = X(t)$ we compute

$$\frac{d}{dt}\,u(X(t),t) = u_x\,X'(t) + u_t.$$

Comparing with (2.4), this vanishes provided the curve satisfies the *characteristic ODE*

$$\frac{dX}{dt} = c(X), \qquad X(0) = x_0. \tag{2.5}$$

So the recipe is: solve the ODE (2.5) to find the characteristic through each starting point x_0; the solution u is constant along it and equal to $\phi(x_0)$. The PDE has been reduced to a family of ODEs, one per characteristic — this is the method of characteristics in general.

Example 2.4 (A variable-speed equation)**.** Solve $u_t + x u_x = 0$ with $u(x,0) = \phi(x)$. The characteristic ODE is $X' = X$, with solution $X(t) = x_0 e^t$, so the characteristics are the curves $x = x_0 e^t$, i.e. $x_0 = x e^{-t}$. Since u is constant along each,

$$u(x,t) = \phi(x_0) = \phi(x e^{-t}).$$

One checks directly: $u_t = \phi'(xe^{-t}) \cdot (-xe^{-t})$ and $u_x = \phi'(xe^{-t}) \cdot e^{-t}$, so $u_t + xu_x = \phi' \cdot (-xe^{-t}) + x\phi' \cdot e^{-t} = 0$. The characteristics fan out from the origin; a feature of the initial data is carried along them, stretching as they separate.

Example 2.5 (Time-dependent speed)**.** Solve $u_t + t\, u_x = 0$ with $u(x,0) = \phi(x)$. The characteristic ODE is now $X'(t) = t$ (the speed depends on t, not x), with solution $X(t) = x_0 + \frac{1}{2}t^2$. So $x_0 = x - \frac{1}{2}t^2$ and

$$u(x,t) = \phi\big(x - \tfrac{1}{2}t^2\big).$$

The characteristics are parabolas in the (x,t) plane, all congruent translates of one another; the profile accelerates to the right, having travelled a distance $\frac{1}{2}t^2$ by time t. This shows that variable coefficients can bend the characteristics (here into parabolas) while still leaving u constant along each.

2.4 The method of characteristics, summarized

The pattern is general and worth stating as a method. To solve a first-order equation of the form $u_t + c(x,t)\, u_x = 0$:

1. Write the characteristic ODE $dX/dt = c(X,t)$ and solve it, obtaining a family of curves indexed by their starting point x_0.

2. Recognize that the PDE says u is constant along each characteristic.

3. Determine the constant from the data: trace each characteristic back to where the data is given (usually $t = 0$) and read off the value.

The method turns a PDE into ODEs. It works because the equation specifies the derivative of u in one particular direction at each point — the direction of the characteristic — and once a directional derivative is known, recovering the function along that direction is an ODE.

Equations with a right-hand side. If the equation has a source, $u_t + cu_x = g(x,t)$, the solution is no longer constant along characteristics; instead it *changes* along them at the rate g. Along $x = X(t)$ we have $\frac{d}{dt}u(X(t),t) = g(X(t),t)$, an ODE to be integrated. Thus the method still applies: the characteristics are unchanged, but on each one solves a simple ODE driven by the source rather than a constant. We pursue this in the exercises.

Example 2.6 (A source term along characteristics)**.** Solve $u_t + u_x = x$ with $u(x,0) = 0$. The characteristics are $X(t) = x_0 + t$ (speed 1), so along the characteristic through x_0,

$$\frac{d}{dt}u(X(t),t) = X(t) = x_0 + t,$$

which integrates (with $u = 0$ at $t = 0$) to $u = x_0 t + \frac{1}{2}t^2$. Substituting $x_0 = x - t$ gives $u(x,t) = (x-t)t + \frac{1}{2}t^2 = xt - \frac{1}{2}t^2$. One verifies $u_t = x - t$, $u_x = t$, so $u_t + u_x = x$, and $u(x,0) = 0$. The source feeds the solution as it rides along each characteristic.

2.5 A glimpse of nonlinearity: shock formation

The method of characteristics survives even when the equation is *nonlinear*, with one dramatic new feature. Consider the inviscid *Burgers equation*

$$u_t + u\,u_x = 0, \qquad u(x,0) = \phi(x), \tag{2.6}$$

in which the transport speed is the solution itself: tall parts of the profile move faster than short parts. The characteristic ODE is now

$$\frac{dX}{dt} = u(X(t), t),$$

and since u is constant along characteristics (the equation has no source), the speed dX/dt is constant on each — so the characteristics are again *straight* lines, but with slopes that depend on the initial height: through $(x_0, 0)$ the line is $x = x_0 + \phi(x_0)\,t$, carrying the value $\phi(x_0)$.

The trouble is that these lines need not stay parallel. Where the initial profile decreases ($\phi'(x_0) < 0$), faster-moving characteristics from behind catch up to slower ones ahead. When two characteristics carrying different values cross, the formula assigns u two values at once: the smooth solution breaks down, and a *shock* — a moving discontinuity — forms. One can compute the first crossing time: characteristics first intersect at

$$t^\star = -\frac{1}{\min_{x_0} \phi'(x_0)},$$

the reciprocal of the steepest negative initial slope. Before $t^\star$ the characteristic construction gives a genuine smooth solution; after it, the notion of solution must be enlarged to admit discontinuities, a subject (conservation laws and weak solutions) beyond this course but worth glimpsing here, because it shows both the reach of the method and the genuinely new phenomena that nonlinearity brings.

Remark 2.7 (Why we dwell on first-order equations)**.** First-order equations are not an end in themselves for this course; their purpose is to introduce characteristics in their cleanest setting. When we factor the wave operator in Chapter 4 as $\partial_{tt} - c^2\partial_{xx} = (\partial_t - c\partial_x)(\partial_t + c\partial_x)$, we will see the wave equation as a pair of transport equations, its solution a superposition of a right-moving and a left-moving wave — exactly the travelling profiles of this chapter. The characteristics of the wave equation, the lines $x \pm ct = \text{const}$, are the direct descendants of the characteristics introduced here.

Exercises

Exercise 2.1 (*Constant-coefficient transport*). Solve $u_t + 3u_x = 0$ with $u(x,0) = \dfrac{1}{1+x^2}$. State the characteristics explicitly and give the formula for $u(x,t)$. Where is the peak of the profile at time $t = 4$?

Exercise 2.2 (*Transport to the left*). Solve $u_t - 2u_x = 0$ with $u(x,0) = \cos x$. In which direction and at what speed does the profile move?

Exercise 2.3 (*The general linear equation*). Solve $2u_x + 3u_t = 0$ with $u(x,0) = \phi(x)$ by identifying the characteristic direction $(2,3)$. Write the characteristics and the solution.

Exercise 2.4 (*Verifying by substitution*). For the solution $u(x,t) = \phi(x - ct)$ of the transport equation, verify directly by the chain rule that $u_t + cu_x = 0$ for any differentiable ϕ.

Exercise 2.5 (*Variable speed*). Solve $u_t + 2x\, u_x = 0$ with $u(x, 0) = \phi(x)$ by the method of characteristics. *(Solve $X' = 2X$ for the characteristics, then express the constant data along them.)*

Exercise 2.6 (*Time-dependent speed*). Solve $u_t + t\, u_x = 0$ with $u(x, 0) = \phi(x)$, following Example 2.5. What shape are the characteristics in the (x, t) plane?

Exercise 2.7 (*A linear source term*). Solve $u_t + cu_x = u$ with $u(x, 0) = \phi(x)$. *(Along the characteristic $x = ct + x_0$, show that u satisfies $\frac{d}{dt}u = u$, hence grows like e^t; conclude $u(x, t) = e^t\, \phi(x - ct)$.)*

Exercise 2.8 (*An inhomogeneous transport equation*). Solve $u_t + u_x = x$ with $u(x, 0) = 0$, following Example 2.6.

Exercise 2.9 (*Characteristics that fan out*). For Example 2.4, sketch several characteristics in the (x, t) plane for $t \geq 0$. Do they converge or diverge as t increases? What does this imply about how a bump in the initial data deforms?

Exercise 2.10 (*Shock formation*). For the Burgers equation $u_t + uu_x = 0$ with $u(x, 0) = -x$ on a suitable interval, find the time $t^\star$ at which characteristics first cross. *(Use $t^\star = -1/\min \phi'$.)* What does the crossing of characteristics mean for the solution?

Exercise 2.11 (*Reduction to an ODE*). Explain in one or two sentences why the method of characteristics reduces a first-order PDE to a family of ordinary differential equations, and what role the initial data plays in pinning down the solution.

Chapter 3

Initial Conditions, Boundary Conditions, and Well-Posedness

A PDE by itself rarely has a unique solution; the transport equation, for instance, is solved by $\phi(x - ct)$ for *any* ϕ. To single out one solution we must supply additional data — the state of the system at an initial time, or its behavior at the boundary of the spatial region, or both. This chapter catalogs the standard kinds of auxiliary data, explains which kinds suit which equations, and makes precise the notion of a *well-posed* problem, the standard of a sensible mathematical model.

3.1 Initial conditions

For an equation that evolves in time, an *initial condition* prescribes the state of the unknown at a starting time, usually $t = 0$. How many initial conditions are needed is governed by the order of the time derivative.

The diffusion equation $u_t = ku_{xx}$ is first order in t, so it needs a single initial condition,

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

the initial temperature distribution. The wave equation $u_{tt} = c^2 u_{xx}$ is second order in t, so — like the ODE $m\ddot{x} = F$, which needs both initial position and initial velocity — it requires two:

$$u(x, 0) = \phi(x), \qquad u_t(x, 0) = \psi(x),$$

the initial displacement and the initial velocity. The number of initial conditions matching the order of the time derivative is not a coincidence; it is what makes the forward evolution determinate, as the explicit solutions of later chapters confirm. The same rule holds for ODEs — a first-order ODE needs one initial value, a second-order ODE two — and for the same reason: each integration in time introduces one constant to be fixed.

3.2 Boundary conditions

When the spatial variable ranges over a bounded region — a rod of length l, a drum, a domain Ω in the plane — we must also say how the unknown behaves at the boundary. There are three standard types, plus one more that ties the two ends together. We state them for a rod $0 < x < l$, where the boundary is the two endpoints $x = 0$ and $x = l$; the same types appear for a region Ω with boundary $\partial\Omega$, with the spatial derivative replaced by the outward normal derivative $\partial u/\partial n$.

Definition 3.1 (Boundary conditions)**.** At an endpoint (or on $\partial\Omega$):

- *Dirichlet* (first kind): the value of u is prescribed, $u = g$. For the rod, $u(0,t) = g_1(t)$, $u(l,t) = g_2(t)$. Physically, the ends are held at prescribed temperatures.
- *Neumann* (second kind): the normal derivative is prescribed, $\partial u/\partial n = g$. For the rod, $u_x(0,t)$ and $u_x(l,t)$ are given. Physically, the heat flux through the ends is prescribed; the homogeneous case $u_x = 0$ means insulated ends.
- *Robin* (third kind, or radiation): a combination of value and normal derivative is prescribed, $\partial u/\partial n + \alpha u = g$ with α a constant. Physically, this models heat exchange with a surrounding medium at a rate proportional to the temperature difference (Newton's law of cooling).

A boundary condition is *homogeneous* if its right-hand side is zero.

A fourth type binds the two ends of an interval together.

Definition 3.2 (Periodic boundary conditions)**.** The *periodic* boundary conditions on $0 < x < l$ are

$$u(0,t) = u(l,t), \qquad u_x(0,t) = u_x(l,t),$$

matching both the value and the slope at the two ends. They are natural when the spatial domain is really a circle of circumference l, so that $x = 0$ and $x = l$ are the same physical point.

The choice of boundary condition is part of the modeling, and it materially affects the solution. The same diffusion equation with the same initial data relaxes to different steady states under Dirichlet, Neumann, and Robin conditions — to the prescribed end temperatures, to a uniform temperature (insulated, conserving total heat), or to the temperature of the surrounding medium, respectively. When we solve these problems by separation of variables in Chapter 8, each boundary condition will produce its own family of eigenfunctions, and the differences among Definitions 3.1 and 3.2 become the differences between sine series, cosine series, and full Fourier series.

Example 3.3 (Mixed boundary conditions)**.** A single problem may carry different conditions at its two ends. A rod with one end held in ice and the other insulated obeys $u(0,t) = 0$ (Dirichlet) and $u_x(l,t) = 0$ (Neumann) — a *mixed* boundary-value problem. The separated eigenfunctions are then $\sin\left(\frac{(2n-1)\pi x}{2l}\right)$, the sines that vanish at $x = 0$ and have zero slope at $x = l$ (a quarter-wave fitting in the interval). The lesson is that each end is treated independently; the eigenfunctions adjust to satisfy whatever condition holds at each.

3.3 Boundary conditions for the three types

Which data make a problem determinate depends on the *type* of the equation (Section 1.4), and matching them correctly is the practical payoff of the classification.

Parabolic equations like the diffusion equation evolve forward in time from initial data and require boundary data along the spatial boundary for all time. The natural setting is a space-time region — the rod for $0 < x < l$, $t > 0$ — with one initial condition at $t = 0$ and one boundary condition at each end for $t > 0$. The data is prescribed on the bottom and sides of the space-time strip, but not the top: the solution at later times is what we solve for.

Hyperbolic equations like the wave equation also evolve from initial data — two conditions, matching the second time derivative — and on a bounded interval take one boundary condition at each end. On the whole line no boundary conditions are needed, and the pure initial-value problem is governed by d'Alembert's formula (Chapter 4).

Elliptic equations like Laplace's equation have no time variable; they describe equilibrium, and they take data on the *entire* boundary of the spatial region at once. A typical well-posed problem is the Dirichlet problem: find u harmonic inside Ω with prescribed values $u = g$ on $\partial\Omega$. There is no notion of "initial" data; the boundary surrounds the region and determines the interior.

Remark 3.4 (Mismatched data is dangerous)**.** Pairing the wrong kind of data with an equation can destroy well-posedness. Prescribing initial data for Laplace's equation — treating one spatial variable as "time" and giving u and its normal derivative on a single curve — produces the notorious *Cauchy problem for the Laplacian*, which is ill-posed: arbitrarily small changes in the data can produce arbitrarily large changes in the solution. Hadamard's example, explored in the exercises, makes this concrete. The lesson is that the type of the equation is not a mere label; it dictates the kind of data the equation can sensibly accept.

3.4 Well-posedness

We can now state the standard of a good mathematical model, due to Hadamard.

Definition 3.5 (Well-posed problem)**.** A problem (a PDE together with initial and/or boundary data) is *well-posed* if it has all three of the following properties:

1. **Existence:** a solution exists.
2. **Uniqueness:** the solution is unique.
3. **Stability:** the solution depends continuously on the data — small changes in the initial or boundary data produce small changes in the solution.

A problem that fails any of these is *ill-posed*.

The three conditions answer three practical worries. Existence asks whether the model is consistent — whether we have over-determined it by demanding too much. Uniqueness asks whether the model is complete — whether we have supplied enough data to single out one answer, or left it ambiguous. Stability, the subtlest, asks whether the model is usable: since real data carries measurement error, a solution that swings wildly under tiny data perturbations predicts nothing. These are not abstract niceties; each of the three can fail, and a central activity of the next chapters is *proving* that the standard problems for the wave and diffusion equations are well-posed.

How we will establish each property. The three properties call for different tools. *Existence* we typically establish constructively, by exhibiting a solution — d'Alembert's formula for the wave equation, the heat kernel for diffusion, a convergent Fourier series on an interval. *Uniqueness* we prove by the *energy method* or the *maximum principle*: assume two solutions, show their difference must vanish. *Stability* often follows from the same energy or maximum-principle estimates, which bound the size of the solution by the size of the data. The energy method in particular — multiply the equation by the unknown, integrate, and integrate by parts — is a recurring engine, introduced for the wave equation in Chapter 4 and for diffusion in Chapter 5.

Example 3.6 (Counting data for a well-posed problem)**.** Consider the wave equation on $0 < x < l$ for $t > 0$. A well-posed problem needs: two initial conditions (because the equation is second order in t), namely $u(x,0) = \phi(x)$ and $u_t(x,0) = \psi(x)$; and one boundary condition at each end for all $t > 0$, say $u(0,t) = u(l,t) = 0$. With fewer conditions the solution is not determined (uniqueness fails); with more — for instance, also prescribing u at the final time $t = T$ — the problem is generally over-determined and no solution exists (existence fails). Counting conditions against the order of the derivatives is the first check on whether a problem is sensibly posed.

Example 3.7 (Uniqueness can fail without enough conditions)**.** The bare diffusion equation $u_t = ku_{xx}$ on $0 < x < l$, with an initial condition but *no* boundary conditions, does not have a unique solution: heat can enter or leave through the unspecified ends in infinitely many ways. Supplying one boundary condition at each end removes the ambiguity. This illustrates that uniqueness is a statement about the *problem* — equation plus data — not about the equation alone.

3.5 Looking ahead

With the apparatus of this chapter in place we can pose the problems the rest of the course solves. For the wave equation: the initial-value problem on the line (two initial conditions, no boundary), and the initial-boundary-value problem on an interval (two initial conditions, one boundary condition per end). For the diffusion equation: the initial-boundary-value problem on an interval (one initial condition, one boundary condition per end) and the pure initial-value problem on the line. For Laplace's equation: the boundary-value problem on a region (data on the whole boundary). In every case we will ask Hadamard's three questions, and the techniques of the coming chapters are, in large part, the means of answering them.

Exercises

Exercise 3.1 (*Counting conditions*). State how many initial conditions in time each equation requires, and why: (a) the diffusion equation $u_t = ku_{xx}$; (b) the wave equation $u_{tt} = c^2u_{xx}$; (c) the beam equation $u_{tt} = -u_{xxxx}$.

Exercise 3.2 (*Naming boundary conditions*). Name the type of each boundary condition on a rod $0 < x < l$: (a) $u(0,t) = 5$; (b) $u_x(l,t) = 0$; (c) $u_x(0,t) - 2u(0,t) = 0$; (d) $u(0,t) = u(l,t)$ and $u_x(0,t) = u_x(l,t)$.

Exercise 3.3 (*Physical meaning*). For a heat-conducting rod, give the physical meaning of each: (a) homogeneous Dirichlet $u = 0$ at an end; (b) homogeneous Neumann $u_x = 0$ at an end; (c) Robin $u_x + \alpha u = 0$ at an end with $\alpha > 0$.

Exercise 3.4 (*Mixed boundary conditions*). For a rod with $u(0,t) = 0$ and $u_x(l,t) = 0$ (Example 3.3), find the spatial eigenfunctions X_n satisfying $X'' + \lambda X = 0$ with $X(0) = 0$, $X'(l) = 0$. *(The eigenvalues are $\lambda_n = ((2n-1)\pi/2l)^2$.)*

Exercise 3.5 (*Matching data to type*). For each equation, state what kind of auxiliary data makes a natural well-posed problem (initial, boundary, or both; on what part of the domain): (a) $u_t = ku_{xx}$ on $0 < x < l$; (b) $u_{xx} + u_{yy} = 0$ on a disk; (c) $u_{tt} = c^2u_{xx}$ on $-\infty < x < \infty$.

Exercise 3.6 (*Counting data*). For the wave equation on $0 < x < l$, list a complete set of initial and boundary conditions for a well-posed problem, and explain what goes wrong if you supply one fewer, or one more (such as a condition at a final time).

Exercise 3.7 (*Steady state under different boundary conditions*). A rod obeys $u_t = ku_{xx}$. Describe (without solving) the steady state $u_t = 0$ it approaches under (a) Dirichlet ends held at temperatures

T_1 and T_2; (b) both ends insulated (homogeneous Neumann). *(For (a), a steady state satisfies* $u_{xx} = 0$*, a straight line; for (b), what conserved quantity fixes the final uniform temperature?)*

Exercise 3.8 (*Non-uniqueness without boundary data*). Explain why the diffusion equation on $0 < x < l$ with only an initial condition fails to have a unique solution, and how supplying boundary conditions repairs this.

Exercise 3.9 (*Hadamard's three conditions*). State the three conditions for well-posedness and, for each, describe a practical failure it guards against.

Exercise 3.10 (*Stability matters*). Suppose a problem has a unique solution, but a measurement error of 0.001 in the initial data changes the solution at a later time by 1000. Is the problem well-posed? Why does this make the model useless for prediction, even though a unique solution exists?

Exercise 3.11 (*An ill-posed Cauchy problem**). Consider Laplace's equation $u_{xx} + u_{yy} = 0$ with Cauchy data $u(x, 0) = 0$ and $u_y(x, 0) = \frac{1}{n}\sin(nx)$. Verify that $u(x, y) = \frac{1}{n^2}\sin(nx)\sinh(ny)$ solves it. As $n \to \infty$ the data tends to zero uniformly, but show the solution does not. Which condition of well-posedness fails, and what does this say about posing initial-value problems for elliptic equations?

Part II

Waves and Diffusions

Chapter 4

The Wave Equation

The wave equation $u_{tt} = c^2 u_{xx}$ is the prototype of a hyperbolic equation, and it is the one second-order equation we can solve completely and explicitly on the whole line. This chapter derives that solution — d'Alembert's formula — reads off from it the hallmark features of wave propagation (finite speed, domains of dependence and influence), and then introduces the *energy method*, the technique that proves uniqueness and stability and that will reappear throughout the course.

4.1 The general solution by factoring

The wave operator factors. Writing ∂_t and ∂_x for the partial-derivative operators,

$$\partial_{tt} - c^2 \partial_{xx} = (\partial_t - c\,\partial_x)(\partial_t + c\,\partial_x),$$

as one checks by expanding (the cross terms $\mp c\,\partial_x \partial_t$ cancel because mixed partials commute). So the wave equation is

$$(\partial_t - c\,\partial_x)(\partial_t + c\,\partial_x)\,u = 0.$$

Each factor is a transport operator of the kind solved in Chapter 2. Introduce *characteristic coordinates*

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

A computation with the chain rule (Exercise 4) turns the wave equation into

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

which we solved in Example 1.2: u is a sum of a function of ξ alone and a function of η alone. Returning to x and t:

Proposition 4.1 (General solution of the wave equation)**.** *Every solution of $u_{tt} = c^2 u_{xx}$ on the line has the form*

$$u(x,t) = F(x + ct) + G(x - ct), \tag{4.1}$$

where F and G are arbitrary twice-differentiable functions. The term $G(x - ct)$ is a wave of fixed shape travelling right at speed c; the term $F(x + ct)$ a wave travelling left at speed c.

The solution of the wave equation is thus a superposition of two rigid travelling waves, one moving each way at speed c. The lines $x \pm ct = \text{const}$ along which each travels are the *characteristics* of the wave equation, the direct descendants of the transport characteristics of Chapter 2.

4.2 D'Alembert's formula

To solve the initial-value problem we choose F and G to match the initial data

$$u(x,0) = \phi(x), \qquad u_t(x,0) = \psi(x).$$

From (4.1), at $t = 0$ we have $F(x) + G(x) = \phi(x)$, and differentiating in t gives $u_t = cF'(x+ct) - cG'(x-ct)$, so at $t = 0$, $cF'(x) - cG'(x) = \psi(x)$. Solving these two relations for F and G (the second integrates to $F(x) - G(x) = \frac{1}{c}\int_0^x \psi + \text{const}$) and substituting back yields the celebrated formula.

Theorem 4.2 (d'Alembert, 1747)**.** *The solution of $u_{tt} = c^2 u_{xx}$ on $-\infty < x < \infty$ with $u(x,0) = \phi(x)$ and $u_t(x,0) = \psi(x)$ is*

$$u(x,t) = \tfrac{1}{2}\big[\phi(x+ct) + \phi(x-ct)\big] + \frac{1}{2c}\int_{x-ct}^{x+ct} \psi(s)\, ds. \tag{4.2}$$

The formula rewards close reading. The displacement part $\frac{1}{2}[\phi(x+ct) + \phi(x-ct)]$ splits the initial shape into two half-amplitude copies that march off in opposite directions. The velocity part spreads the initial velocity over the interval $[x-ct, x+ct]$ between the two characteristics through (x,t). Existence is settled by the formula itself: it manifestly solves the problem (one can substitute and check), provided ϕ is twice and ψ once differentiable.

Example 4.3 (A plucked string)**.** Release a string from rest ($\psi = 0$) with initial shape ϕ a narrow triangular bump. By (4.2) the solution is $\frac{1}{2}[\phi(x+ct)+\phi(x-ct)]$: the bump splits into two half-height triangles, one gliding left and one right at speed c, each keeping its shape. This is the mathematical content of a plucked string releasing two travelling pulses.

Example 4.4 (A struck string)**.** Now strike a string at rest, giving it an initial velocity but no initial displacement: $\phi = 0$ and $\psi(x) = v$ for $|x| < a$, zero otherwise (a hammer blow over a small region). D'Alembert's formula gives $u(x,t) = \frac{1}{2c}\int_{x-ct}^{x+ct}\psi(s)\,ds = \frac{v}{2c} \times$ (length of overlap of $[x-ct, x+ct]$ with $(-a,a)$). For a point far from the struck region, u is zero until the disturbance arrives, then rises to a plateau of height $\frac{v}{2c}\cdot 2a = \frac{va}{c}$ as the struck region's influence fully passes, between the outgoing wavefronts. Unlike the plucked string, whose displacement returns to zero behind the pulses, the struck string is left permanently displaced in the wake of the velocity pulse — the integral of the velocity accumulates and does not undo itself. This contrast between ϕ-driven and ψ-driven solutions is one of the most instructive features of d'Alembert's formula.

4.3 Causality: domains of dependence and influence

D'Alembert's formula (4.2) shows that the solution at a point (x_0, t_0) depends on the initial data only through ϕ at the two points $x_0 \pm ct_0$ and ψ on the interval between them. This interval,

$$[\,x_0 - ct_0,\ x_0 + ct_0\,],$$

is the *domain of dependence* of (x_0, t_0): the solution there is completely determined by — and only by — the data on this interval. Data outside it has no effect whatever on $u(x_0, t_0)$.

Turning this around, the data at a single point x_0 at time zero can influence the solution at time t only within the interval $[x_0 - ct, x_0 + ct]$, the *domain of influence*, an expanding "light cone" opening at speed c. Together these express *finite propagation speed*: signals in the wave equation travel no faster than c. A disturbance confined initially to a small region is not felt elsewhere until enough time has passed for a wave to arrive at speed c — in sharp contrast to the diffusion equation, which we will find spreads influence *instantly* (Chapter 6).

Example 4.5 (The parallelogram rule)**.** The characteristic structure yields a pretty identity. Take any four points forming a "characteristic parallelogram" — a parallelogram whose sides lie along the characteristics $x \pm ct =$ const. Label the vertices A (bottom), B (right), C (top), D (left). Because the general solution is $F(x+ct)+G(x-ct)$ and each side holds one of $x \pm ct$ fixed, a short computation shows

$$u(A) + u(C) = u(B) + u(D) :$$

the sums of u over opposite vertices are equal. This *parallelogram rule* is a purely geometric consequence of the travelling-wave structure, it requires no integration, and it provides both a quick way to find the solution at one vertex from the other three and a tool for solving boundary-value problems by propagating values along characteristics. It is the cleanest expression of how the wave equation moves information along its two characteristic families.

Remark 4.6 (The hyperbolic personality)**.** Finite speed, the preservation of shapes by travelling waves, and the central role of characteristics are the defining traits of hyperbolic equations. D'Alembert's formula displays all three at once. When in later chapters we contrast waves with diffusions, this triad — propagation without distortion, finite speed, dependence on data through characteristics — is the wave side of the contrast.

4.4 The energy method

We now prove uniqueness for the wave equation by the energy method, the single most important proof technique in the course. The physical idea is that a vibrating string has a conserved total energy, and that a solution with zero initial energy must stay zero.

For the wave equation on the line (or on an interval with suitable boundary conditions) define the *energy* at time t,

$$E(t) = \tfrac{1}{2} \int \left[\, u_t^2 + c^2 u_x^2 \,\right] dx, \tag{4.3}$$

the integral of kinetic energy ($\propto u_t^2$) plus potential energy ($\propto u_x^2$). We compute its rate of change. Differentiating under the integral sign,

$$\frac{dE}{dt} = \int \left[\, u_t u_{tt} + c^2 u_x u_{xt} \,\right] dx.$$

Integrate the second term by parts in x; on the whole line (with u decaying at infinity) or on an interval with $u_t = 0$ or $u_x = 0$ at the ends, the boundary terms vanish, leaving $\int c^2 u_x u_{xt}\, dx = -\int c^2 u_{xx} u_t\, dx$. Hence

$$\frac{dE}{dt} = \int u_t \big(u_{tt} - c^2 u_{xx}\big)\, dx = 0,$$

because $u_{tt} - c^2 u_{xx} = 0$ by the equation. The energy is *conserved*: $E(t) = E(0)$ for all time. This is a theorem about the wave equation, proved in three lines, and it is the model for every energy argument to come.

Theorem 4.7 (Uniqueness for the wave equation)**.** *The initial-value problem for $u_{tt} = c^2 u_{xx}$ with given ϕ and ψ (on the line, or on an interval with Dirichlet or Neumann boundary conditions) has at most one solution.*

Proof. Suppose u_1 and u_2 both solve the problem with the same data. Their difference $w = u_1 - u_2$ is a linear combination of solutions, hence solves the wave equation (Proposition 1.5), and has *zero* initial data: $w(x, 0) = 0$ and $w_t(x, 0) = 0$. Its energy at $t = 0$ is therefore $E(0) = \frac{1}{2} \int [w_t^2 + c^2 w_x^2]\, dx = 0$, since $w_t = 0$ and also $w_x = 0$ (as $w(\cdot, 0) \equiv 0$ has zero spatial derivative). By conservation, $E(t) = 0$ for all t. But $E(t)$ is an integral of squares, so $E(t) = 0$ forces $w_t \equiv 0$ and $w_x \equiv 0$ everywhere. A function with vanishing partial derivatives is constant, and since $w = 0$ initially, $w \equiv 0$. Thus $u_1 = u_2$. □

The structure of this proof recurs verbatim for the diffusion equation (Chapter 5) and elsewhere: form the difference of two solutions, note it satisfies the equation with zero data, write down an energy that the equation forces to be non-increasing (here, constant), and conclude from $E \leq E(0) = 0$ and $E \geq 0$ that the difference vanishes. The "vanishing theorem" — that a continuous nonnegative function with zero integral is itself zero — is the small lemma that finishes such arguments.

Remark 4.8 (Energy and stability)**.** Conservation of energy gives more than uniqueness; it gives stability. Since $E(t) = E(0)$, the size of the solution (measured by the energy) at any later time is controlled by the size of the initial data. Small data means small energy for all time, so the solution depends continuously on the data in the energy norm — the third pillar of well-posedness (Definition 3.5). The single energy identity thus delivers two of Hadamard's three conditions at once.

4.5 The inhomogeneous wave equation

A forced string — driven by an external transverse force — obeys the *inhomogeneous* wave equation

$$u_{tt} - c^2 u_{xx} = f(x, t),$$

with f the applied force per unit mass. Duhamel's principle (Chapter 7) solves it by superposing the responses to the forcing at each instant. The result, for zero initial data, is

$$u(x, t) = \frac{1}{2c} \int_0^t \int_{x-c(t-s)}^{x+c(t-s)} f(y, s)\, dy\, ds,$$

an integral of the source over the backward characteristic triangle with apex at (x, t) — exactly the region of the past that can influence the point, by finite speed. The structure mirrors the velocity term of d'Alembert's formula: where ψ was spread over the interval $[x - ct, x + ct]$, the forcing $f(\cdot, s)$ at each earlier time s is spread over the smaller interval $[x - c(t - s), x + c(t - s)]$ and the contributions accumulated. We return to the principle behind this formula in Chapter 7; here it illustrates that the travelling-wave and finite-speed structure persists when the equation is forced.

4.6 The wave equation on an interval, briefly

On a finite string $0 < x < l$ with fixed ends $u(0, t) = u(l, t) = 0$, the travelling-wave picture acquires reflections: a pulse reaching an end inverts and returns, as we work out in Chapter 7. The energy method above still applies — the boundary terms vanish because u, hence u_t, is zero at the fixed ends — so uniqueness and stability hold on the interval too. The explicit solution, however, is more naturally obtained by separation of variables (Chapter 8), which expands the vibration in the normal modes $\sin(n\pi x/l)$ of the string. The two viewpoints — travelling waves and standing-wave modes — are complementary.

Exercises

Exercise 4.1 (*Characteristic coordinates*). Using $\xi = x + ct$, $\eta = x - ct$, show by the chain rule that $u_{tt} - c^2 u_{xx} = -4c^2 u_{\xi\eta}$, so that the wave equation becomes $u_{\xi\eta} = 0$. Conclude the general solution (4.1).

Exercise 4.2 (*Applying d'Alembert*). Solve $u_{tt} = u_{xx}$ (so $c = 1$) with $u(x, 0) = \sin x$ and $u_t(x, 0) = 0$. Then solve it with $u(x, 0) = 0$ and $u_t(x, 0) = \cos x$.

Exercise 4.3 (*Both data nonzero*). Solve $u_{tt} = u_{xx}$ with $u(x, 0) = e^{-x^2}$ and $u_t(x, 0) = 1$. Write the solution as the sum of its displacement and velocity contributions, and identify the (linearly growing) effect of the constant initial velocity.

Exercise 4.4 (*The velocity term*). Solve $u_{tt} = 4u_{xx}$ with $u(x, 0) = 0$ and $u_t(x, 0) = \psi(x)$ where $\psi(x) = 1$ for $|x| < 1$ and 0 otherwise. Describe the solution's shape for large t. *(The displacement spreads the initial velocity over* $[x - 2t, x + 2t]$*.)*

Exercise 4.5 (*Plucked versus struck*). Contrast the long-time displacement of a plucked string (Example 4.3) and a struck string (Example 4.4). Why does the struck string remain displaced behind the pulse while the plucked string returns to zero?

Exercise 4.6 (*Domain of dependence*). For $u_{tt} = c^2 u_{xx}$, the solution at (x_0, t_0) depends on the initial data over what set? If the initial data is changed only on the interval $|x| < 1$, for which points (x_0, t_0) is the solution affected?

Exercise 4.7 (*The parallelogram rule*). State and verify the parallelogram rule $u(A) + u(C) = u(B) + u(D)$ of Example 4.5 for the four vertices of a characteristic parallelogram, using the general solution $F(x + ct) + G(x - ct)$.

Exercise 4.8 (*Finite speed*). Explain, using d'Alembert's formula, why a disturbance initially confined to $|x| < a$ cannot affect the solution at a point x_0 with $|x_0| > a$ until time $t = (|x_0| - a)/c$. Contrast this with what you expect for the diffusion equation.

Exercise 4.9 (*Energy is conserved*). Carry out the energy computation in detail for the wave equation on the whole line, justifying the vanishing of the boundary terms. Where is the assumption that u solves the wave equation used?

Exercise 4.10 (*Uniqueness on an interval*). State and prove uniqueness for the wave equation on $0 < x < l$ with fixed ends $u(0, t) = u(l, t) = 0$. *(Check that the boundary terms in the energy computation vanish because* $u_t = 0$ *at the ends.)*

Exercise 4.11 (*Energy with a damping term**). For the damped wave equation $u_{tt} + au_t = c^2 u_{xx}$ with $a > 0$, compute dE/dt for the energy (4.3) and show E is non-increasing. What does this say about the long-time behavior of the vibration, and about uniqueness?

Chapter 5

The Diffusion Equation

The diffusion equation $u_t = ku_{xx}$ is the prototype of a parabolic equation, and its character could hardly be more different from the wave equation's. Where waves propagate at finite speed and preserve shapes, diffusion spreads instantly, smooths everything, and decays toward equilibrium. This chapter derives the equation carefully, proves the *maximum principle* that governs its qualitative behavior, and gives two proofs of uniqueness — one by the maximum principle, one by the energy method of Chapter 4 — including the energy argument with Neumann and Robin boundary conditions that figures in the course's examinations.

5.1 Derivation from conservation and Fourier's law

Let $u(x,t)$ be the temperature in a thin uniform rod, and consider the heat contained in a segment $a \le x \le b$. Two principles govern it.

Conservation of heat. The heat stored in the segment is proportional to $\int_a^b u\,dx$, and it changes only through flux across the two ends. If $q(x,t)$ denotes the heat flux (rate of heat flow in the $+x$ direction), then the rate of change of stored heat equals the flux in at $x = a$ minus the flux out at $x = b$:

$$c_p\rho\,\frac{d}{dt}\int_a^b u\,dx = q(a,t) - q(b,t) = -\int_a^b q_x\,dx,$$

where $c_p\rho$ bundles the specific heat and density and the last step is the fundamental theorem of calculus.

Fourier's law. Heat flows from hot to cold, at a rate proportional to the temperature gradient: $q = -Ku_x$, with $K > 0$ the conductivity. Substituting,

$$c_p\rho\int_a^b u_t\,dx = -\int_a^b (-Ku_x)_x\,dx = \int_a^b Ku_{xx}\,dx.$$

Since this holds for *every* segment $[a,b]$, the integrands must agree, giving $c_p\rho\,u_t = Ku_{xx}$, that is,

$$u_t = ku_{xx}, \qquad k = \frac{K}{c_p\rho} > 0, \tag{5.1}$$

the diffusion equation, with *diffusivity* k. The identical derivation, with Fourier's law replaced by Fick's law (matter flows down concentration gradients), governs the diffusion of a dissolved substance — hence the two names for (5.1). In two or three space dimensions the same reasoning, using the divergence theorem in place of the fundamental theorem of calculus, gives $u_t = k\,\Delta u$.

5.2 The maximum principle

The qualitative behavior of diffusion is governed by a single powerful statement: heat cannot spontaneously concentrate. Made precise, this is the *maximum principle.*

Theorem 5.1 (Maximum principle)**.** *Let $u(x,t)$ solve the diffusion equation $u_t = ku_{xx}$ in the space-time rectangle $0 \le x \le l$, $0 \le t \le T$. Then the maximum of u over the closed rectangle is attained on the* parabolic boundary *— the bottom $t = 0$ or the two sides $x = 0$, $x = l$. The same holds for the minimum.*

In words: the hottest the rod ever gets, over the whole time interval, occurs either initially or at one of the ends; the temperature in the interior at a later time cannot exceed what was supplied by the initial and boundary data. The intuition is that an interior hot spot would have to be a local maximum in x, where $u_{xx} \le 0$, forcing $u_t = ku_{xx} \le 0$ so the spot cools — heat flows away from a maximum, never toward it.

Idea of the proof. The intuition above is almost a proof, but the case $u_{xx} = 0$ at the maximum needs care. The standard device is to perturb: set $v = u + \varepsilon x^2$ for small $\varepsilon > 0$. Then $v_t - kv_{xx} = u_t - ku_{xx} - 2k\varepsilon = -2k\varepsilon < 0$, so v *strictly* satisfies the inequality. At an interior maximum of v one would need $v_t \ge 0$ (time could only have increased to reach it) and $v_{xx} \le 0$, giving $v_t - kv_{xx} \ge 0$, a contradiction. So v attains its maximum on the parabolic boundary; letting $\varepsilon \to 0$ transfers the conclusion to u. (The minimum follows by applying the result to $-u$.) □

The maximum principle is the parabolic counterpart of conservation of energy for waves: a structural fact, provable in a few lines, from which uniqueness and stability flow.

Corollary 5.2 (Uniqueness by the maximum principle)**.** *The diffusion equation on $0 < x < l$ with given initial data $u(x,0) = \phi(x)$ and Dirichlet boundary data $u(0,t)$, $u(l,t)$ has at most one solution.*

Proof. If u_1, u_2 solve the problem, $w = u_1 - u_2$ solves the diffusion equation with zero initial and zero boundary data. By the maximum principle the maximum *and* minimum of w are attained on the parabolic boundary, where $w = 0$; hence $w \equiv 0$, so $u_1 = u_2$. □

Corollary 5.3 (Stability)**.** *Under the same conditions, if the data of two problems differ by at most δ (in initial values and boundary values), then their solutions differ by at most δ everywhere. Diffusion is stable: small changes in the data produce small changes in the solution.*

Proof. Apply the maximum principle to the difference of the two solutions, whose parabolic-boundary values are bounded in absolute value by δ, so the difference itself is bounded by δ throughout. □

Example 5.4 (A maximum-principle estimate)**.** Consider $u_t = u_{xx}$ on $0 < x < 1$, $t > 0$, with ends held at $u(0,t) = u(1,t) = 0$ and initial data $0 \le \phi(x) \le 10$. What can we say about u at later times without solving? By the maximum principle, the maximum over any rectangle $0 \le t \le T$ is attained on the parabolic boundary, where $u \le 10$ (initially) or $u = 0$ (at the ends); hence $u(x,t) \le 10$ everywhere. The minimum is likewise attained on the parabolic boundary, where $u \ge 0$; hence $u(x,t) \ge 0$ everywhere. So $0 \le u(x,t) \le 10$ for all x and t — the temperature stays within the range set by the data, and (as we will see from the series solution) decays toward zero. This kind of *a priori* bound, obtained before and without solving, is one of the chief uses of the maximum principle.

Remark 5.5 (The comparison principle)**.** The maximum principle has a useful restatement as a *comparison principle*: if u and v both solve the diffusion equation and $u \le v$ on the parabolic boundary (initially and at the ends), then $u \le v$ everywhere and for all time. (Apply the maximum principle to $w = u - v$, which is ≤ 0 on the parabolic boundary.) In words, ordering of the data is preserved by the evolution: a rod that starts cooler and is held cooler at the ends stays cooler. This monotonicity is a powerful tool — it lets one bound an unknown solution above and below by known ones — and it is special to parabolic (and elliptic) equations; the wave equation, which can overshoot, has no comparison principle.

5.3 Uniqueness by the energy method

The energy method of Chapter 4 gives a second, independent proof of uniqueness, and it is the one that adapts most readily to Neumann and Robin boundary conditions. For the diffusion equation the natural energy is simpler than for the wave equation — just the integral of u^2.

Let w solve the diffusion equation $w_t = kw_{xx}$ on $0 < x < l$, and define

$$E(t) = \tfrac{1}{2} \int_0^l w(x,t)^2 \, dx. \tag{5.2}$$

Differentiating and using the equation,

$$\frac{dE}{dt} = \int_0^l w \, w_t \, dx = \int_0^l w \cdot k w_{xx} \, dx = k \Big[w w_x \Big]_0^l - k \int_0^l w_x^2 \, dx,$$

integrating by parts. The boundary term is $k[w(l)w_x(l) - w(0)w_x(0)]$. For *homogeneous Dirichlet* data ($w = 0$ at both ends) or *homogeneous Neumann* data ($w_x = 0$ at both ends) this term vanishes, leaving

$$\frac{dE}{dt} = -k \int_0^l w_x^2 \, dx \;\le\; 0. \tag{5.3}$$

The energy is *non-increasing.* This is the diffusion analogue of energy conservation for waves — but here the inequality is strict decay, reflecting that diffusion dissipates rather than conserves.

Theorem 5.6 (Uniqueness by energy, Neumann case)**.** *The diffusion equation on $0 < x < l$ with given initial data and Neumann boundary data $u_x(0,t) = g_1(t)$, $u_x(l,t) = g_2(t)$ has at most one solution.*

Proof. Let $w = u_1 - u_2$ be the difference of two solutions; it solves the diffusion equation with zero initial data and *homogeneous* Neumann data $w_x(0,t) = w_x(l,t) = 0$ (the inhomogeneous parts cancel). The boundary term in dE/dt vanishes, so by (5.3) E is non-increasing; and $E(0) = \frac{1}{2}\int w(x,0)^2 \, dx = 0$ since $w(\cdot,0) \equiv 0$. As $E \ge 0$ always and $E(0) = 0$ with E non-increasing, $E(t) = 0$ for all t. An integral of a square that vanishes forces $w \equiv 0$. Hence $u_1 = u_2$. □

Remark 5.7 (Robin boundary conditions)**.** The energy method handles Robin conditions with one extra observation, which is exactly the step tested in the course's examinations. Suppose homogeneous Robin data, $w_x(0,t) = \alpha \, w(0,t)$ and $w_x(l,t) = -\beta \, w(l,t)$ with $\alpha, \beta \ge 0$ (signs chosen so heat is lost, not gained, at the ends). Then the boundary term becomes

$$k\big[w(l)w_x(l) - w(0)w_x(0)\big] = k\big[-\beta \, w(l)^2 - \alpha \, w(0)^2\big] \;\le\; 0,$$

so that $dE/dt = -k\int_0^l w_x^2 \, dx - k\beta w(l)^2 - k\alpha w(0)^2 \le 0$, and the uniqueness argument goes through unchanged. The sign conditions $\alpha, \beta \ge 0$ are essential: physically they say the ends radiate heat away, and without them uniqueness can genuinely fail. A reaction term is handled similarly: for $w_t = kw_{xx} - c(x,t)w$ with $c \geq 0$, the energy acquires an extra $-\int c\, w^2\, dx \leq 0$, again preserving the decay.

5.4 Two equations, two personalities

We now contrast the two evolution equations we have solved. The wave equation conserves energy; the diffusion equation dissipates it. The wave equation propagates signals at finite speed c; diffusion, as we will see explicitly in Chapter 6, spreads influence instantly. The wave equation is reversible — run time backward and it is still the wave equation — while diffusion is irreversible: the backward diffusion equation $u_t = -ku_{xx}$ is famously ill-posed, since it would require unsmoothing, amplifying every wrinkle. The wave equation preserves the roughness of its data; diffusion smooths instantly, and a solution at any positive time is infinitely differentiable no matter how rough the initial data. These contrasts are the concrete meaning of the words *hyperbolic* and *parabolic*, and holding the two equations side by side is the best way to understand either.

The table below summarizes the contrast, the entries of which we have now established or will establish in the next chapter.

	Wave equation	*Diffusion equation*
type	hyperbolic	parabolic
energy	conserved	dissipated
speed of influence	finite (c)	infinite
effect on rough data	preserved	smoothed instantly
time-reversal	reversible	irreversible
key structural tool	energy conservation	maximum principle

Exercises

Exercise 5.1 (*Deriving the equation*). Repeat the derivation of the diffusion equation, stating clearly where conservation of heat is used and where Fourier's law is used. Which physical law would change if we modeled diffusion of a chemical instead of heat?

Exercise 5.2 (*The maximum principle in words*). Explain physically why an interior local maximum of temperature must be cooling, using the sign of u_{xx} there and the equation $u_t = ku_{xx}$.

Exercise 5.3 (*Using the maximum principle*). A rod $0 < x < 1$ has initial temperature $u(x,0) = \sin(\pi x) \leq 1$ and ends held at $u(0,t) = u(1,t) = 0$. Without solving, what is the largest value u can take for any x and any $t > 0$? Justify by the maximum principle.

Exercise 5.4 (*A maximum-principle estimate*). For $u_t = u_{xx}$ on $0 < x < 1$ with ends held at 0 and initial data $0 \leq \phi \leq 10$, show $0 \leq u(x,t) \leq 10$ for all $t > 0$ (Example 5.4). Where on the parabolic boundary are the bounds attained?

Exercise 5.5 (*The comparison principle*). State the comparison principle for the diffusion equation and deduce it from the maximum principle. If one rod starts everywhere cooler than another and is held cooler at the ends, what can you conclude about their temperatures at later times?

Exercise 5.6 (*Energy decay*). For the diffusion equation on $0 < x < l$ with homogeneous Dirichlet data, show $E(t) = \frac{1}{2}\int_0^l u^2\, dx$ is non-increasing, and explain why this expresses the dissipative nature of diffusion. Can E ever increase?

Exercise 5.7 (*Uniqueness, Dirichlet case*). Prove uniqueness for the diffusion equation on $0 < x < l$ with Dirichlet boundary data by the energy method, and separately by the maximum principle. Which proof was shorter?

Exercise 5.8 (*Robin boundary conditions*). Prove uniqueness for the diffusion equation on $0 < x < l$ with homogeneous Robin conditions $u_x(0,t) = \alpha\, u(0,t)$ and $u_x(l,t) = -\beta\, u(l,t)$, $\alpha, \beta \geq 0$. Identify precisely where the sign conditions on α, β are used.

Exercise 5.9 (*A reaction term*). Prove uniqueness for $u_t = ku_{xx} - c(x,t)\, u$ with $c(x,t) \geq 0$ and Neumann boundary data, by the energy method. Show that without the assumption $c \geq 0$ the energy argument breaks down.

Exercise 5.10 (*Backward diffusion is ill-posed**). Consider the backward diffusion equation $u_t = -ku_{xx}$. Show that $u(x,t) = e^{kn^2 t} \sin(nx)$ is a solution with initial data $\sin(nx)$. As n grows, the initial data stays bounded but the solution at any fixed $t > 0$ blows up. Which condition of well-posedness fails, and how does this reflect the irreversibility of diffusion?

Chapter 6

The Diffusion Equation on the Whole Line

On a finite interval the diffusion equation is best solved by separation of variables (Chapter 8). On the whole line $-\infty < x < \infty$ there are no boundaries, and a different and very beautiful solution is available: an explicit formula built from a single special solution, the *heat kernel*. This chapter derives the heat kernel, solves the initial-value problem by convolution, and reads off the striking qualitative properties — instantaneous smoothing and infinite propagation speed — that set diffusion apart from waves. The transform machinery that produces the kernel systematically is developed in Chapter 11; here we obtain it by a more elementary route, exploiting a symmetry of the equation.

6.1 A special solution from scaling

The diffusion equation $u_t = ku_{xx}$ has a remarkable symmetry: it looks the same under the rescaling $x \mapsto \lambda x$, $t \mapsto \lambda^2 t$. That is, if $u(x,t)$ is a solution, so is $u(\lambda x, \lambda^2 t)$ for any $\lambda > 0$ (check by the chain rule). The pairing of one power of x with *two* of $\sqrt{t}$ — equivalently, the combination $x/\sqrt{t}$ is unchanged by the scaling — is the signature of diffusion, and it suggests looking for a solution that depends only on the scale-invariant variable $x/\sqrt{t}$.

Pursuing this idea (and fixing the constant by requiring the total heat $\int u\,dx$ to be conserved and equal to one) leads to the *fundamental solution* of the diffusion equation, also called the *heat kernel*:

$$S(x,t) = \frac{1}{\sqrt{4\pi kt}} \exp\left(-\frac{x^2}{4kt}\right), \qquad t > 0. \tag{6.1}$$

One verifies directly (Exercise 6) that S solves $u_t = ku_{xx}$ for $t > 0$. As a function of x at fixed t, it is a Gaussian bump centered at the origin, with total area $\int_{-\infty}^{\infty} S(x,t)\,dx = 1$ for every t, and width growing like $\sqrt{t}$. As $t \to 0^+$ the bump becomes infinitely tall and narrow while keeping unit area: it represents a unit of heat initially concentrated at a single point, spreading out as time advances. This is the diffusion equation's answer to "what happens to a hot spot?"

Remark 6.1 (Why $\sqrt{t}$, and what it means)**.** The width of the heat kernel grows like $\sqrt{t}$, not like t. This is the mathematical fingerprint of diffusion and the reason diffusive spreading is *slow*: to spread twice as far takes four times as long. It contrasts sharply with wave propagation, where a signal travels a distance proportional to t (at constant speed c). The diffusion length scale $\sqrt{kt}$

recurs throughout the subject — it is, for instance, how deep a daily or seasonal temperature swing penetrates into the ground.

6.2 Solving the initial-value problem by convolution

The heat kernel solves the problem with a point source of heat. By superposition (Proposition 1.5) we can build the solution for *any* initial distribution by regarding it as a continuous superposition of point sources. If the initial temperature is $u(x,0) = \phi(x)$, think of ϕ as a weighted collection of point sources, the source at location y carrying "strength" $\phi(y)\,dy$. Each evolves into a spreading Gaussian $S(x-y,t)$, and summing (integrating) them gives the *convolution* of ϕ with the kernel.

Theorem 6.2 (Solution of the diffusion equation on the line)**.** *For suitable initial data ϕ, the solution of $u_t = ku_{xx}$ on $-\infty < x < \infty$ with $u(x,0) = \phi(x)$ is*

$$u(x,t) = \int_{-\infty}^{\infty} S(x-y,t)\,\phi(y)\,dy = \frac{1}{\sqrt{4\pi kt}} \int_{-\infty}^{\infty} e^{-(x-y)^2/4kt}\,\phi(y)\,dy. \tag{6.2}$$

The formula expresses the temperature at (x,t) as a weighted average of the initial data, with weight $S(x-y,t)$ that is largest for y near x and falls off like a Gaussian of width $\sqrt{4kt}$. As t increases, the averaging window widens, and the temperature at any point reflects an ever broader swath of the initial data. That $u(x,t) \to \phi(x)$ as $t \to 0^+$, recovering the initial data, is the statement that the kernel becomes a point source — made rigorous in Chapter 11.

Example 6.3 (Spreading of a temperature step)**.** Let the initial temperature be a step, $\phi(x) = 1$ for $x > 0$ and 0 for $x < 0$. Formula (6.2) gives, after the substitution $z = (y-x)/\sqrt{4kt}$,

$$u(x,t) = \frac{1}{\sqrt{\pi}} \int_{-x/\sqrt{4kt}}^{\infty} e^{-z^2}\,dz = \tfrac{1}{2}\Big(1 + \operatorname{erf}\big(\tfrac{x}{\sqrt{4kt}}\big)\Big),$$

where erf is the error function. The sharp step immediately becomes a smooth transition, of width $\sim \sqrt{kt}$, that broadens with time — the rough initial data is smoothed at once.

6.3 The error function

The error function that appeared in Example 6.3 is worth recording, since it is to diffusion what the trigonometric functions are to waves — the special function in terms of which solutions are expressed. It is defined by

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

the (normalized) area under a Gaussian out to z. Its basic properties follow from this definition: it is odd, $\operatorname{erf}(-z) = -\operatorname{erf}(z)$; it starts flat at the origin with $\operatorname{erf}(0) = 0$; and it saturates, $\operatorname{erf}(z) \to 1$ as $z \to \infty$ (using the Gaussian integral $\int_0^\infty e^{-s^2} ds = \sqrt{\pi}/2$) and $\to -1$ as $z \to -\infty$. The normalization $2/\sqrt{\pi}$ is chosen exactly so that the limits are ± 1. The step-data solution $\frac{1}{2}(1 + \operatorname{erf}(x/\sqrt{4kt}))$ thus runs smoothly from 0 (deep on the cold side, $x \to -\infty$) to 1 (deep on the hot side, $x \to +\infty$), with the transition centered at $x = 0$ and of width $\sim \sqrt{kt}$ — the boundary between hot and cold blurring as $\sqrt{t}$, precisely the diffusion length scale.

6.4 Smoothing and infinite propagation speed

Two qualitative features of (6.2) deserve emphasis, both sharp contrasts with the wave equation.

Instantaneous smoothing. For any $t > 0$, the solution $u(\cdot, t)$ is infinitely differentiable in x, no matter how rough ϕ is — even if ϕ is discontinuous, as in Example 6.3. The reason is visible in (6.2): u is an integral of the smooth kernel S against ϕ, and derivatives in x fall on S, which is infinitely differentiable. Diffusion regularizes instantly; this is the precise opposite of the wave equation, which preserves the roughness of its data and propagates discontinuities undimmed along characteristics.

Infinite propagation speed. Because the Gaussian kernel $S(x - y, t)$ is strictly positive for *all* x and y (however far apart) at every $t > 0$, the value $u(x, t)$ depends on the initial data at every point y, no matter how distant. Concretely, if $\phi \geq 0$ is positive only on a tiny interval, then $u(x, t) > 0$ *everywhere* for every $t > 0$: the heat is felt instantly at arbitrarily large distances. This *infinite propagation speed* is physically an idealization — the diffusion equation is a continuum model that ignores the finite speed of molecules — but mathematically it is a defining trait of parabolic equations, and the exact opposite of the wave equation's finite speed c (Section 4.3).

Remark 6.4 (Diffusion and the random walk)**.** The Gaussian heat kernel is no coincidence; it is the same Gaussian that governs the sum of many small random steps, by the central limit theorem. Indeed the diffusion equation is the continuum limit of a *random walk*: a particle taking tiny random steps left and right has, after time t, a probability distribution of position that is exactly the spreading Gaussian $S(x, t)$, with the diffusivity k set by the step size and rate. This is the probabilistic meaning of diffusion — heat (or a dissolved dye, or a pollen grain in water) spreads because its carriers wander randomly — and it explains both the Gaussian shape and the $\sqrt{t}$ spreading: the standard deviation of a sum of n independent steps grows like $\sqrt{n}$, and $n \propto t$. The connection runs deep, and it is the bridge between this course and the theory of stochastic processes.

Remark 6.5 (The contrast, sharpened)**.** Place the two evolution equations side by side once more. The wave equation: finite speed, no smoothing, reversible, dependence on data through two characteristics. The diffusion equation: infinite speed, instant smoothing, irreversible, dependence on *all* the data through a positive kernel. The heat kernel (6.1) and d'Alembert's formula (4.2) are the two explicit solutions that make these opposite personalities concrete, and a student who understands both formulas understands the difference between hyperbolic and parabolic equations.

6.5 Looking ahead to the Fourier transform

We obtained the heat kernel here by guessing its form from the scaling symmetry. There is a systematic method that produces it — and solves the diffusion equation on the line — without guessing: the *Fourier transform*, developed in Chapter 11. The transform turns the diffusion equation, a PDE in x, into a simple ODE in t for each frequency, solves it, and inverts; the Gaussian heat kernel emerges because the Fourier transform of a Gaussian is again a Gaussian. The convolution structure of (6.2) is likewise no accident: it reflects the convolution theorem, that the transform turns convolution into ordinary multiplication. We will see all of this in due course; for now, the heat kernel stands as a complete and explicit solution, however obtained.

Example 6.6 (A box of heat)**.** Let the initial temperature be a unit box, $\phi(x) = 1$ for $|x| < a$ and 0 otherwise — a slab of warm material between two cold half-lines. The convolution (6.2) integrates the kernel over the box,

$$u(x,t) = \frac{1}{\sqrt{4\pi kt}} \int_{-a}^{a} e^{-(x-y)^2/4kt}\, dy = \tfrac{1}{2}\Big[\operatorname{erf}\Big(\tfrac{a-x}{\sqrt{4kt}}\Big) + \operatorname{erf}\Big(\tfrac{a+x}{\sqrt{4kt}}\Big)\Big],$$

a difference of two error-function fronts, one spreading from each edge of the box. For small t the profile is still nearly the box, with the two edges blurred over a width $\sim \sqrt{kt}$; for large t (when $\sqrt{kt} \gg a$) the box has spread so far that u is approximately the point-source kernel of total heat $2a$, namely $u \approx \frac{2a}{\sqrt{4\pi kt}} e^{-x^2/4kt}$ — the slab, viewed from far away, looks like a point source. The example shows both edges of the box smoothing independently and the eventual loss of memory of the initial shape, retaining only its total heat.

Exercises

Exercise 6.1 (*Verifying the heat kernel*). Verify by direct differentiation that $S(x,t) = (4\pi kt)^{-1/2} e^{-x^2/4kt}$ satisfies $u_t = ku_{xx}$ for $t > 0$. *(Compute S_t and S_{xx} separately and compare.)*

Exercise 6.2 (*Unit mass*). Show that $\int_{-\infty}^{\infty} S(x,t)\,dx = 1$ for every $t > 0$, using the Gaussian integral $\int_{-\infty}^{\infty} e^{-ax^2} dx = \sqrt{\pi/a}$. Why is this the statement that total heat is conserved?

Exercise 6.3 (*The scaling symmetry*). Show that if $u(x,t)$ solves the diffusion equation, so does $u_\lambda(x,t) = u(\lambda x, \lambda^2 t)$ for any $\lambda > 0$. Which combination of x and t is left invariant, and how does this motivate the form of the heat kernel?

Exercise 6.4 (*Spreading of a step*). Carry out the substitution in Example 6.3 to obtain the error-function solution for step initial data. What is the width of the transition region as a function of t?

Exercise 6.5 (*Properties of erf*). Using the definition $\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-s^2} ds$, show that erf is odd, that $\operatorname{erf}(0) = 0$, and that $\operatorname{erf}(z) \to 1$ as $z \to \infty$. Why is the constant $2/\sqrt{\pi}$ the right normalization?

Exercise 6.6 (*A Gaussian stays Gaussian*). Take initial data $\phi(x) = e^{-x^2}$ and use (6.2) to show the solution remains Gaussian for all t, with increasing width. *(The convolution of two Gaussians is a Gaussian; you may use this.)*

Exercise 6.7 (*Instantaneous smoothing*). Explain, from the convolution formula (6.2), why the solution is infinitely differentiable in x for every $t > 0$ even when ϕ is discontinuous. Where do the x-derivatives act?

Exercise 6.8 (*Infinite propagation speed*). Suppose $\phi \geq 0$ is positive only on $|x| < 1$ and zero elsewhere. Show that $u(x,t) > 0$ for *every* x and every $t > 0$. Contrast this with the domain of influence of the wave equation.

Exercise 6.9 (*The random walk***). Explain, qualitatively, why the position of a particle performing a random walk of many small steps is approximately Gaussian after time t, and how this relates the diffusion equation to the central limit theorem. Why does the spread grow like $\sqrt{t}$?

Exercise 6.10 (*The diffusion length*). A daily temperature oscillation at the ground surface penetrates to a depth of order $\sqrt{kt}$ with t one day. Using this scaling, explain qualitatively why a seasonal oscillation (period one year) penetrates roughly $\sqrt{365} \approx 19$ times deeper than a daily one.

Chapter 7

Reflections and Sources

So far we have solved the wave and diffusion equations on the whole line, where there are no boundaries. Two natural extensions remain before we turn to finite intervals: problems on a *half-line*, where a single boundary reflects waves and constrains diffusion, and problems with a *source* term, where the equation is inhomogeneous. Both are handled by clever applications of the whole-line solutions we already have — reflections by the method of odd and even extensions, sources by Duhamel's principle. The techniques are characteristic of the subject's economy: rather than solve a new problem from scratch, we reduce it to one already solved.

7.1 The half-line: reflection

Consider the diffusion equation on the half-line $0 < x < \infty$, with a boundary condition at $x = 0$ and initial data $\phi(x)$ for $x > 0$. The idea is to *extend* the data to the whole line in a way that automatically enforces the boundary condition, solve the resulting whole-line problem by the heat kernel (Chapter 6), and restrict the answer to $x > 0$.

Dirichlet condition: odd reflection. Suppose the boundary is held at zero, $u(0, t) = 0$. Extend ϕ to the whole line as an *odd* function: define $\phi_{\text{odd}}(x) = \phi(x)$ for $x > 0$ and $\phi_{\text{odd}}(-x) = -\phi(x)$. Solve the whole-line problem with data ϕ_{odd}. Because the heat kernel is even and the data is odd, the solution $u(x, t)$ is odd in x for all time, and an odd function automatically vanishes at $x = 0$ — so the Dirichlet condition holds for free. Restricting to $x > 0$ gives the solution of the half-line problem. Explicitly,

$$u(x, t) = \int_0^\infty \big[\, S(x - y, t) - S(x + y, t) \,\big]\, \phi(y)\, dy,$$

the whole-line formula with the odd extension folded back onto $x > 0$; the second term is the "reflected" contribution.

Neumann condition: even reflection. Suppose instead the boundary is insulated, $u_x(0, t) = 0$. Now extend ϕ as an *even* function, $\phi_{\text{even}}(-x) = \phi(x)$. The solution is then even in x, and an even function has zero slope at $x = 0$ — so the Neumann condition holds automatically. The formula is the same as above but with a *plus* sign: $S(x - y, t) + S(x + y, t)$.

The same device works for the wave equation. A wave on a half-line with a fixed end ($u(0, t) = 0$) is solved by odd reflection: a pulse travelling toward the boundary is met by an inverted mirror-image pulse coming the other way, and their superposition keeps $u(0, t) = 0$ at all times. The physical picture — a pulse reflecting off a fixed end and returning *inverted* — is exactly the odd reflection made visible. A free end ($u_x(0,t) = 0$) reflects without inversion, the even case.

Remark 7.1 (Why reflection works)**.** The method of reflection is an instance of a general principle: build the symmetry of the boundary condition into the data, so that the symmetry of the solution enforces the condition. Odd data gives an odd solution, which vanishes at the origin (Dirichlet); even data gives an even solution, which has zero slope at the origin (Neumann). The same idea, applied on a *bounded* interval, requires reflecting repeatedly to produce a periodic extension — and that periodic extension is precisely the Fourier series of Chapter 9. Reflection on the half-line is the simplest case of the extension idea that organizes all of Part III.

7.2 The method of images

The reflection idea has a vivid physical name: the *method of images.* To enforce $u = 0$ at a wall, one imagines an *image source* of opposite sign placed at the mirror position behind the wall; the real and image sources together produce a field that vanishes on the wall by symmetry. For the half-line, a unit of heat released at $x = a > 0$ with a cold wall at $x = 0$ is represented by the real source at a together with an image *sink* (a negative source) at $-a$, giving

$$u(x,t) = S(x-a,t) - S(x+a,t),$$

which vanishes at $x = 0$ for all t because the two terms are equal there. For an insulated wall the image is a source of the *same* sign, giving a sum and hence zero slope at the wall. The method extends to two walls — a rod of finite length, or a point charge between two grounded planes — by placing an infinite train of images, reflecting repeatedly across both walls. That infinite train of images is, once again, the periodic extension whose Fourier series solves the bounded-interval problem; the method of images and separation of variables are two routes to the same answer, the former more geometric and the latter more computational.

7.3 Inhomogeneous equations and Duhamel's principle

Now consider a *source* term. The inhomogeneous diffusion equation

$$u_t - ku_{xx} = f(x,t), \qquad u(x,0) = 0, \tag{7.1}$$

models a rod heated by a distributed source f (with zero initial temperature, the homogeneous initial data being recoverable separately by superposition). *Duhamel's principle* solves it by treating the source as a continuous succession of instantaneous initial-value problems.

The idea: the heat deposited by the source during a short time interval $[s, s+ds]$ acts like a small burst of initial data, $f(x,s)\,ds$, introduced at time s. After that instant it evolves freely according to the *homogeneous* equation. Let $U(x,t;s)$ denote the solution of the homogeneous diffusion equation that starts, at time s, from initial data $f(x,s)$:

$$U_t = kU_{xx} \quad (t > s), \qquad U(x,s;s) = f(x,s).$$

Then the solution of the inhomogeneous problem is the accumulation of all these bursts:

$$u(x,t) = \int_0^t U(x,t;s)\,ds. \tag{7.2}$$

Proposition 7.2 (Duhamel's principle)**.** *The function* (7.2) *solves the inhomogeneous problem* (7.1)*.*

Proof. Differentiate (7.2) using the Leibniz rule for differentiating an integral whose upper limit is the variable t:

$$u_t = U(x,t;t) + \int_0^t U_t(x,t;s)\,ds = f(x,t) + \int_0^t k\,U_{xx}(x,t;s)\,ds,$$

using the initial condition $U(x,t;t) = f(x,t)$ for the boundary term and the homogeneous equation $U_t = kU_{xx}$ inside the integral. The remaining integral is ku_{xx} (differentiating (7.2) twice in x), so $u_t = f + ku_{xx}$, which is (7.1). And $u(x,0) = 0$ since the integral from 0 to 0 vanishes. □

Duhamel's principle is a recurring tool, and it has the same flavor as the convolution solution of Chapter 6: decompose the forcing into elementary pieces, evolve each by the homogeneous solution we already know, and superpose. For the diffusion equation on the line, combining Duhamel with the heat kernel gives a closed double-integral formula for the solution with both a source and initial data. The same principle applies to the wave equation, where it reproduces the velocity term of d'Alembert's formula and extends it to forced vibrations.

Duhamel for the wave equation. The principle is general; only the "elementary solution" changes. For the inhomogeneous wave equation $u_{tt} - c^2 u_{xx} = f(x,t)$ with zero initial data, the burst of forcing $f(\cdot,s)\,ds$ at time s is treated as an initial *velocity* (since the wave equation is second order in time, a forcing enters at the level of acceleration, hence contributes to velocity). The solution starting at time s from zero displacement and velocity $f(\cdot,s)$ is, by d'Alembert's velocity term, $\frac{1}{2c}\int_{x-c(t-s)}^{x+c(t-s)} f(y,s)\,dy$. Accumulating over s recovers the characteristic-triangle formula of Section 4.5:

$$u(x,t) = \frac{1}{2c}\int_0^t \int_{x-c(t-s)}^{x+c(t-s)} f(y,s)\,dy\,ds.$$

The same recipe — evolve each instant's forcing by the known homogeneous solution and integrate — yields the forced response of both equations.

Example 7.3 (A steadily heated rod⋆)**.** If the source is constant in time, $f(x,t) = f(x)$, Duhamel's formula and the heat kernel combine to describe a rod approaching a steady state. As $t \to \infty$, $u(x,t)$ tends to the solution of the steady equation $-ku_{xx} = f(x)$ (with the given boundary conditions) — the balance in which the source's heating is exactly carried away by conduction. The approach to this steady state, and the steady state itself as a solution of an ordinary differential equation in x, tie the time-dependent and equilibrium problems together.

Example 7.4 (Cooling of a half-line with a hot patch)**.** A half-line rod $x > 0$ has its end held at zero, $u(0,t) = 0$, and starts with a hot patch $\phi(x) = 1$ on $1 < x < 2$ and zero elsewhere. By odd reflection (Section 7.1) the solution is

$$u(x,t) = \int_1^2 \big[S(x-y,t) - S(x+y,t)\big]\,dy,$$

which evaluates, exactly as in the box example of Chapter 6, to a combination of four error-function terms — two from the real patch at $(1,2)$ and two from its mirror image (a cold patch) at $(-2,-1)$. The image patch is what drives heat *out* through the cold end: near $x = 0$ the real and image contributions cancel, pinning $u(0,t) = 0$, while the temperature gradient there carries heat steadily out of the rod. As $t \to \infty$ all the heat eventually escapes through the end and $u \to 0$ everywhere — in contrast to the insulated (even-reflection) case, where the image patch is warm, no heat escapes, and the rod would instead relax toward a nonzero spread-out profile. The sign of the image, set by the boundary condition, thus decides the rod's ultimate fate.

Exercises

Exercise 7.1 (*Odd reflection*). Solve the diffusion equation on $0 < x < \infty$ with $u(0,t) = 0$ and initial data $\phi(x)$, by odd reflection. Write the solution as an integral over $y > 0$ of the kernel and its reflection, and explain why the boundary condition holds automatically.

Exercise 7.2 (*Even reflection*). Repeat the previous exercise for the insulated boundary $u_x(0,t) = 0$, using even reflection. How does the sign of the reflected term differ from the Dirichlet case?

Exercise 7.3 (*Method of images*). A unit of heat is released at $x = a > 0$ on the half-line with a cold wall $u(0,t) = 0$. Write the image representation of the solution and verify it vanishes at $x = 0$. What image would enforce an insulated wall instead?

Exercise 7.4 (*A reflected wave*). A wave on $0 < x < \infty$ with a fixed end $u(0,t) = 0$ has initial pulse ϕ supported in $1 < x < 2$, released from rest. Using odd reflection, describe what happens when the left-moving part of the pulse reaches the boundary at $x = 0$. Does it return inverted or upright?

Exercise 7.5 (*Free-end reflection*). Repeat the previous exercise for a free end $u_x(0,t) = 0$ (even reflection). Is the reflected pulse inverted or upright now?

Exercise 7.6 (*Duhamel by hand*). Use Duhamel's principle to solve the ODE analogue $u'(t) = au(t) + f(t)$, $u(0) = 0$, and compare with the integrating-factor solution $u(t) = \int_0^t e^{a(t-s)} f(s)\, ds$. *(Here the "homogeneous solution starting at s" is $e^{a(t-s)} f(s)$.)*

Exercise 7.7 (*Verifying Duhamel*). Carry out the differentiation in the proof of Proposition 7.2 in full detail, stating where the Leibniz rule, the initial condition $U(x,t;t) = f(x,t)$, and the homogeneous equation are each used.

Exercise 7.8 (*Duhamel for the wave equation*). Using the wave-equation form of Duhamel's principle, derive the characteristic-triangle formula for $u_{tt} - c^2 u_{xx} = f$, $u(x,0) = u_t(x,0) = 0$. Why does the forcing enter as an initial velocity rather than an initial displacement?

Exercise 7.9 (*Source plus initial data*). Explain how to solve the diffusion equation with *both* a source $f(x,t)$ and nonzero initial data $\phi(x)$, by splitting into two problems and superposing. Which principle licenses the splitting?

Exercise 7.10 (*Approach to steady state**). For a rod $0 < x < l$ with ends held at 0 and a constant source $f(x)$, write the steady-state equation that $u(x,t)$ approaches as $t \to \infty$, and explain physically what balance it represents.

Part III

Fourier Series and Separation of Variables

Chapter 8

Separation of Variables

We come to the central computational technique of the course. On a finite interval, the wave and diffusion equations are solved not by travelling waves or kernels but by *separation of variables*: one seeks solutions that are products of a function of x alone and a function of t alone, finds infinitely many such product solutions, and superposes them to match the initial data. The method works because the boundary conditions pick out a special set of spatial shapes — the *eigenfunctions* — in which any reasonable initial profile can be expanded. That expansion is a Fourier series, developed in the next chapter; here we see how separation of variables produces it and why the boundary conditions are decisive.

8.1 The method on the diffusion equation

Consider the diffusion equation on a rod with both ends held at zero:

$$u_t = ku_{xx}, \quad 0 < x < l,\ t > 0; \qquad u(0,t) = u(l,t) = 0; \qquad u(x,0) = \phi(x). \tag{8.1}$$

Seek a *product solution* $u(x,t) = X(x)\,T(t)$. Substituting into the equation, $XT' = kX''T$, and dividing by kXT,

$$\frac{T'}{kT} = \frac{X''}{X}.$$

The left side depends only on t, the right only on x; for them to be equal for all x and t, both must equal the same constant, which we write $-\lambda$ (the sign and notation chosen for later convenience):

$$\frac{X''}{X} = -\lambda = \frac{T'}{kT}.$$

This *separates* the PDE into two ODEs,

$$X'' + \lambda X = 0, \qquad T' + \lambda k\, T = 0. \tag{8.2}$$

The boundary conditions $u(0,t) = u(l,t) = 0$ become $X(0) = X(l) = 0$ (since $T(t)$ is not identically zero). The spatial problem is thus

$$X'' + \lambda X = 0, \qquad X(0) = X(l) = 0, \tag{8.3}$$

a *boundary-value problem* for X. This is not an ordinary initial-value ODE: we seek values of λ for which a *nonzero* solution X exists.

8.2 The eigenvalue problem

Problem (8.3) is an *eigenvalue problem*: the values of λ admitting a nonzero solution are the *eigenvalues*, and the corresponding solutions X are the *eigenfunctions*. Let us find them, examining the sign of λ.

If $\lambda < 0$, write $\lambda = -\mu^2$; the general solution of $X'' = \mu^2 X$ is $X = Ae^{\mu x} + Be^{-\mu x}$, and the conditions $X(0) = X(l) = 0$ force $A = B = 0$. No nonzero solution; negative λ gives nothing. If $\lambda = 0$, then $X'' = 0$ so $X = A + Bx$, and again $X(0) = X(l) = 0$ forces $X \equiv 0$. So we need $\lambda > 0$. Write $\lambda = \beta^2$; then $X = A\cos\beta x + B\sin\beta x$. The condition $X(0) = 0$ gives $A = 0$, leaving $X = B\sin\beta x$. The condition $X(l) = 0$ requires $\sin\beta l = 0$, hence $\beta l = n\pi$ for a positive integer n. We have found the spectrum.

Proposition 8.1 (Dirichlet eigenvalues and eigenfunctions)**.** *The eigenvalue problem* (8.3) *has eigenvalues and eigenfunctions*

$$\lambda_n = \left(\frac{n\pi}{l}\right)^2, \qquad X_n(x) = \sin\frac{n\pi x}{l}, \qquad n = 1, 2, 3, \ldots \tag{8.4}$$

For each eigenvalue the time equation $T' + \lambda_n k\,T = 0$ has solution $T_n(t) = e^{-\lambda_n kt}$, decaying exponentially, faster for larger n. Each product

$$u_n(x,t) = e^{-(n\pi/l)^2 kt}\,\sin\frac{n\pi x}{l}$$

solves the diffusion equation and the boundary conditions. These are the *normal modes* of the rod: spatial shapes $\sin(n\pi x/l)$ that decay in place without changing shape, the higher modes (more wiggles) decaying faster.

8.3 Superposition and the Fourier series

A single mode rarely matches the initial data. But by superposition (Proposition 1.5) any combination

$$u(x,t) = \sum_{n=1}^{\infty} b_n\, e^{-(n\pi/l)^2 kt}\,\sin\frac{n\pi x}{l} \tag{8.5}$$

also solves the equation and the boundary conditions, for any coefficients b_n (granting convergence). It remains to choose the b_n so that the initial condition holds. Setting $t = 0$ in (8.5),

$$\phi(x) = \sum_{n=1}^{\infty} b_n \sin\frac{n\pi x}{l}. \tag{8.6}$$

This is the demand that the initial data be expressible as a sum of the eigenfunctions — a *Fourier sine series*. The question of whether an arbitrary ϕ admits such an expansion, and how to compute the coefficients b_n, is the subject of Chapter 9. The answer, in brief: yes, for any reasonable ϕ, and the coefficients are given by an integral.

In (8.5), as t grows, every term decays, the higher modes vanishing fastest, so after a short time only the first mode $b_1 e^{-(\pi/l)^2 kt}\sin(\pi x/l)$ remains appreciable: *the solution relaxes to the slowest-decaying mode and then dies away*, regardless of how complicated the initial data was. This is the diffusion equation's smoothing and decay, seen now through the modes rather than the heat kernel.

8.4 A fully worked example

It is worth carrying one problem all the way through, coefficients and all, to see the method in operation. Solve

$$u_t = u_{xx}, \quad 0 < x < \pi,\ t > 0; \qquad u(0,t) = u(\pi,t) = 0; \qquad u(x,0) = x(\pi - x).$$

Here $k = 1$, $l = \pi$, so the eigenfunctions are $\sin(nx)$ with eigenvalues n^2, and the solution is $u(x,t) = \sum_{n=1}^{\infty} b_n e^{-n^2 t} \sin(nx)$ with the b_n the sine coefficients of $\phi(x) = x(\pi - x)$. By the coefficient formula (anticipating Chapter 9),

$$b_n = \frac{2}{\pi} \int_0^{\pi} x(\pi - x) \sin(nx)\, dx.$$

Two integrations by parts give $\int_0^{\pi} x(\pi - x)\sin(nx)\, dx = \frac{2}{n^3}\big(1 - (-1)^n\big)$, which is $4/n^3$ for odd n and 0 for even n. Hence $b_n = \frac{8}{\pi n^3}$ for odd n, and

$$u(x,t) = \frac{8}{\pi} \sum_{n \text{ odd}} \frac{1}{n^3}\, e^{-n^2 t} \sin(nx).$$

Several features are now concrete. The coefficients decay like $1/n^3$, so the series converges quickly even at $t = 0$ (the initial data is smooth and vanishes at both ends, matching the eigenfunctions). For $t > 0$ the factor $e^{-n^2 t}$ accelerates the decay enormously: by $t = 1$ the $n = 3$ term is already down by $e^{-9} \approx 10^{-4}$ relative to its $t = 0$ size, so the solution is essentially $\frac{8}{\pi} e^{-t} \sin x$ — the fundamental mode alone. Rapid smoothing followed by slow decay of the lowest mode is visible in this one formula.

8.5 Other boundary conditions, other series

The boundary conditions determine the eigenfunctions, and changing them changes the series. The method is identical; only the spatial eigenvalue problem changes.

Neumann (insulated ends). For $u_x(0,t) = u_x(l,t) = 0$ the spatial conditions are $X'(0) = X'(l) = 0$. The same analysis gives eigenfunctions $X_n(x) = \cos(n\pi x/l)$ for $n = 0, 1, 2, \ldots$ — note that $n = 0$ now contributes a nonzero constant eigenfunction $X_0 = 1$ with $\lambda_0 = 0$. The expansion of the initial data is a *Fourier cosine series*, and the constant mode $X_0 = 1$ does not decay, reflecting that insulated ends conserve total heat: the rod relaxes to its average initial temperature rather than to zero.

Periodic (a ring). For periodic conditions $u(0,t) = u(l,t)$, $u_x(0,t) = u_x(l,t)$, both $\cos(2n\pi x/l)$ and $\sin(2n\pi x/l)$ are eigenfunctions, and the expansion is a *full Fourier series*.

Robin (radiating ends). For Robin conditions the eigenvalue condition becomes a transcendental equation (such as $\tan \beta l =$ a function of β) with no closed-form roots, though the eigenfunctions still exist and the method proceeds. We treat the general pattern under Sturm–Liouville theory in Chapter 10.

The three classical boundary conditions thus generate the three classical Fourier series — sine, cosine, and full — and the choice among them is dictated entirely by the physics at the ends. This is the deep reason the next chapter studies Fourier series: they are not an arbitrary tool but the precise language in which separated solutions are assembled.

8.6 Inhomogeneous boundary conditions

Separation of variables needs *homogeneous* boundary conditions — the step $X(0) = 0$ used the boundary value being zero. When the ends are held at nonzero temperatures, say $u(0,t) = A$ and $u(l,t) = B$ (constants), we first subtract a function carrying those values. The natural choice is the *steady state* $v(x) = A + (B - A)x/l$, the straight line solving $v'' = 0$ with $v(0) = A$, $v(l) = B$. Then $w = u - v$ satisfies the heat equation (since v is time-independent and $v_{xx} = 0$), now with *homogeneous* Dirichlet conditions $w(0,t) = w(l,t) = 0$ and initial data $w(x,0) = \phi(x) - v(x)$. We solve for w by the sine series as above, and recover $u = v + w$. The physical reading is clean: v is the final equilibrium the rod settles into, and w is the decaying transient by which it approaches it. The same subtract-the-steady-state device handles steady sources and is the bridge from the homogeneous theory to applied problems.

8.7 Separation of variables for the wave equation

The same method solves the wave equation on a string with fixed ends:

$$u_{tt} = c^2 u_{xx}, \quad 0 < x < l; \qquad u(0,t) = u(l,t) = 0.$$

Separation $u = X(x)T(t)$ gives the same spatial problem (8.3), hence the same eigenfunctions $\sin(n\pi x/l)$, but now the time equation is $T'' + \lambda_n c^2 T = 0$, whose solutions *oscillate* — $\cos(n\pi ct/l)$ and $\sin(n\pi ct/l)$ — rather than decay. Each mode

$$u_n(x,t) = \left(a_n \cos \tfrac{n\pi ct}{l} + b_n \sin \tfrac{n\pi ct}{l}\right) \sin \frac{n\pi x}{l}$$

is a *standing wave*: a fixed spatial shape $\sin(n\pi x/l)$ vibrating in time at frequency $n\pi c/l$. These are the harmonics of a musical string — the fundamental ($n = 1$) and its overtones — and their superposition, matching the initial displacement and velocity, is the general vibration. The contrast with the heat equation is instructive: identical spatial modes, but decaying for diffusion and oscillating for waves, the difference residing entirely in the time equation and hence in the order of the time derivative.

Example 8.2 (A plucked string by separation)**.** Release the string $0 < x < l$ from rest ($u_t(x,0) = 0$) with initial shape $\phi(x)$. Zero initial velocity kills the $\sin(n\pi ct/l)$ terms (their time derivative at $t = 0$ would be nonzero), so $b_n = 0$ and $u(x,t) = \sum_n a_n \cos(n\pi ct/l)\sin(n\pi x/l)$, with a_n the sine coefficients of ϕ. Each mode simply oscillates in place at its own frequency; the string's motion is the superposition, and because the frequencies $n\pi c/l$ are integer multiples of the fundamental, the motion is periodic in time with period $2l/c$. This is why a plucked string sounds a definite musical pitch — its overtones are harmonically related.

Remark 8.3 (Standing waves and travelling waves)**.** We now have two descriptions of the vibrating string: the travelling waves of d'Alembert (Chapter 4) and the standing-wave modes here. They are equivalent — a standing wave is the superposition of two travelling waves moving in opposite directions, as the identity $2\sin(n\pi x/l)\cos(n\pi ct/l) = \sin\frac{n\pi}{l}(x + ct) + \sin\frac{n\pi}{l}(x - ct)$ shows. Which description is more useful depends on the question: travelling waves for propagation and reflection on the line, standing modes for vibration on a bounded interval.

Exercises

Exercise 8.1 (*Separating the equation*). Carry out separation of variables for $u_t = ku_{xx}$, deriving the two ODEs (8.2), and explain why the separation constant must be the same for both.

Exercise 8.2 (*Dirichlet eigenvalues*). Solve the eigenvalue problem $X'' + \lambda X = 0$, $X(0) = X(l) = 0$ in full, treating the cases $\lambda < 0$, $\lambda = 0$, $\lambda > 0$ separately, and confirm (8.4).

Exercise 8.3 (*Neumann eigenvalues*). Solve $X'' + \lambda X = 0$ with $X'(0) = X'(l) = 0$. Show the eigenfunctions are $\cos(n\pi x/l)$ for $n = 0, 1, 2, \ldots,$ and explain why $n = 0$ now gives a nonzero (constant) eigenfunction.

Exercise 8.4 (*A single mode*). Solve (8.1) on $0 < x < \pi$ (so $l = \pi$) with $k = 1$ and initial data $\phi(x) = 5\sin(3x)$. *(Only one term of the series is needed; identify it.)* What is the solution at time $t = 1$?

Exercise 8.5 (*A sum of modes*). Solve the same problem with $\phi(x) = 2\sin x - \sin(4x)$. Write the full time-dependent solution, and state which mode dominates for large t.

Exercise 8.6 (*A fully worked series*). Solve $u_t = u_{xx}$ on $0 < x < \pi$ with fixed ends and $\phi(x) = x(\pi - x)$, following Section 8.4. Compute the coefficients b_n and write the solution; estimate how many terms matter at $t = 1$.

Exercise 8.7 (*Insulated rod*). Solve $u_t = u_{xx}$ on $0 < x < \pi$ with insulated ends $u_x(0, t) = u_x(\pi, t) = 0$ and $\phi(x) = 3 + \cos(2x)$. To what value does the temperature relax as $t \to \infty$, and why does it not go to zero?

Exercise 8.8 (*Inhomogeneous boundary data*). Solve $u_t = u_{xx}$ on $0 < x < 1$ with $u(0, t) = 2$, $u(1, t) = 5$, and $u(x, 0) = 2 + 3x$. *(Subtract the steady state $v(x) = 2 + 3x$; what does the transient w satisfy here?)*

Exercise 8.9 (*The vibrating string*). Solve $u_{tt} = c^2u_{xx}$ on $0 < x < l$ with fixed ends, initial shape $\phi(x) = \sin(\pi x/l)$ and zero initial velocity. Identify the frequency of the resulting vibration.

Exercise 8.10 (*Standing equals travelling*). Verify the identity $2\sin(n\pi x/l)\cos(n\pi ct/l) = \sin\frac{n\pi}{l}(x + ct) + \sin\frac{n\pi}{l}(x - ct)$, and explain how it expresses a standing wave as a superposition of two travelling waves.

Exercise 8.11 (*Decay rates*). In the series solution (8.5), the nth mode decays like $e^{-(n\pi/l)^2kt}$. Explain why higher modes decay faster, and why this means the diffusion equation smooths rough initial data: which features of ϕ are carried by the high modes?

Chapter 9

Fourier Series

Separation of variables left us with a debt: it produced solutions as infinite series of eigenfunctions — sines, cosines — and asked us to believe that any reasonable initial profile could be written as such a series. This chapter pays the debt. We compute the coefficients of a Fourier series, using the crucial *orthogonality* of the eigenfunctions, and then survey what is known about when, and in what sense, the series converges. Fourier series are of independent importance throughout mathematics and physics; for us they are above all the language in which the solutions of Chapter 8 are expressed and made rigorous.

9.1 The Fourier coefficients via orthogonality

Suppose we wish to expand a function ϕ on $0 < x < l$ in a sine series,

$$\phi(x) = \sum_{n=1}^{\infty} b_n \sin\frac{n\pi x}{l}, \tag{9.1}$$

as the Dirichlet problem demands. How do we find the coefficients b_n? The key is a computation: the sine functions are *orthogonal* over the interval. For positive integers m, n,

$$\int_0^l \sin\frac{m\pi x}{l}\,\sin\frac{n\pi x}{l}\,dx = \begin{cases} 0, & m \neq n, \\ \dfrac{l}{2}, & m = n. \end{cases} \tag{9.2}$$

(The off-diagonal vanishing follows from the product-to-sum identity; the diagonal value from $\int_0^l \sin^2(n\pi x/l)\,dx = l/2$.) Orthogonality is the discrete analogue of perpendicular coordinate axes: each eigenfunction points in its own independent "direction," and we extract a coefficient by projecting onto it.

To find b_m, multiply (9.1) by $\sin(m\pi x/l)$ and integrate term by term. By orthogonality every term on the right vanishes except $n = m$, which contributes $b_m \cdot l/2$:

$$\int_0^l \phi(x)\sin\frac{m\pi x}{l}\,dx = b_m \cdot \frac{l}{2}.$$

Solving gives the coefficient formula.

Proposition 9.1 (Fourier sine coefficients)**.** *The coefficients in the sine series* (9.1) *are*

$$b_n = \frac{2}{l}\int_0^l \phi(x)\,\sin\frac{n\pi x}{l}\,dx. \tag{9.3}$$

The same recipe — multiply by an eigenfunction, integrate, use orthogonality — gives the coefficients of any Fourier series. For the *cosine series* $\phi(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(n\pi x/l)$ arising from Neumann conditions, the cosines satisfy the analogous orthogonality relation, and

$$a_n = \frac{2}{l}\int_0^l \phi(x)\cos\frac{n\pi x}{l}\,dx, \quad n \geq 0,$$

the constant term $\frac{a_0}{2}$ being the average of ϕ over the interval. For the *full Fourier series* on $-l < x < l$, both sines and cosines appear, with coefficients given by integrals over the full interval.

Example 9.2 (Sine series of a constant)**.** Expand $\phi(x) = 1$ on $0 < x < l$ in a sine series. By (9.3),

$$b_n = \frac{2}{l}\int_0^l \sin\frac{n\pi x}{l}\,dx = \frac{2}{l}\cdot\frac{l}{n\pi}\big(1 - \cos n\pi\big) = \frac{2}{n\pi}\big(1 - (-1)^n\big),$$

which is $4/(n\pi)$ for odd n and 0 for even n. So $1 = \frac{4}{\pi}\sum_{n\text{ odd}} \frac{1}{n}\sin(n\pi x/l)$ on the open interval. Notice the series represents the constant 1 inside $(0, l)$ but, being a sum of sines, vanishes at the endpoints — a first sign that the series may behave specially at points where the expanded function and the eigenfunctions disagree.

Example 9.3 (A sawtooth and the value of a classical sum)**.** Expand $\phi(x) = x$ on $0 < x < l$ in a sine series. Integrating (9.3) by parts,

$$b_n = \frac{2}{l}\int_0^l x\sin\frac{n\pi x}{l}\,dx = \frac{2l}{n\pi}(-1)^{n+1},$$

so $x = \frac{2l}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}\sin\frac{n\pi x}{l}$. The coefficients decay only like $1/n$, because the odd periodic extension of x has jumps at $x = \pm l$ (it leaps from l down to $-l$). This slow decay is the price of the discontinuity, and it foreshadows the Gibbs phenomenon at those jumps.

9.2 Even, odd, and periodic extensions

Why does a function on $0 < x < l$ have both a sine series and a cosine series? Because a series of sines, evaluated outside $(0, l)$, defines the *odd*, $2l$-periodic extension of ϕ, while a series of cosines defines the *even*, $2l$-periodic extension. The Fourier series on the interval is really the full Fourier series of one of these periodic extensions, restricted back to $(0, l)$. This is the same extension idea that powered the method of reflection in Chapter 7: odd extension enforces a Dirichlet (vanishing) condition at the ends, even extension a Neumann (zero-slope) condition. The choice of sine versus cosine series is thus the choice of how to extend ϕ past the boundary, which is in turn dictated by the boundary condition of the PDE.

9.3 The complex form

For many purposes — especially the transition to the Fourier transform in Chapter 11 — the most economical way to write a full Fourier series is in *complex exponential* form. Using Euler's formula $e^{i\theta} = \cos\theta + i\sin\theta$, the sines and cosines combine into a single family $e^{in\pi x/l}$, and the full Fourier series on $-l < x < l$ becomes

$$\phi(x) = \sum_{n=-\infty}^{\infty} c_n\, e^{in\pi x/l}, \qquad c_n = \frac{1}{2l}\int_{-l}^{l} \phi(x)\, e^{-in\pi x/l}\,dx,$$

the sum now running over all integers, positive and negative. The complex coefficients c_n package the sine and cosine coefficients together (c_n and c_{-n} encode a_n and b_n), and the orthogonality relation $\int_{-l}^{l} e^{im\pi x/l}\overline{e^{in\pi x/l}}\,dx = 2l\,\delta_{mn}$ that produces the coefficient formula is cleaner than for sines and cosines separately. As $l \to \infty$, this complex series passes formally into the Fourier transform integral, the discrete index n becoming a continuous frequency — a limit we make precise in Chapter 11.

9.4 Convergence: in what sense, and how fast

A Fourier series is an infinite sum, and we must ask whether it converges and to what. The answer is subtle, and three distinct notions of convergence are worth distinguishing.

Pointwise convergence. At a point x where ϕ is continuous and reasonably smooth, the Fourier series converges to the value $\phi(x)$. At a jump discontinuity, it converges to the *average* of the left and right limits, $\frac{1}{2}[\phi(x^-)+\phi(x^+)]$ — the series "splits the difference" at a jump. This is the content of *Dirichlet's theorem*, the basic pointwise convergence result, valid for piecewise-smooth functions.

Uniform convergence. If ϕ is continuous, periodic, and piecewise smooth, the Fourier series converges *uniformly* — the approximation is good everywhere at once. Uniform convergence fails, however, when ϕ has a jump, and the manner of its failure is the Gibbs phenomenon below.

Mean-square (L^2) convergence. The most robust notion: for any square-integrable ϕ (any ϕ with $\int_0^l \phi^2\,dx < \infty$), the partial sums converge to ϕ in the mean-square sense, meaning the integral of the squared error tends to zero. This holds with no smoothness assumption at all, and it is the natural notion for the energy methods of the course, since the energy is a mean-square quantity. Mean-square convergence is equivalent to the *completeness* of the eigenfunctions, discussed next.

Remark 9.4 (Smoothness and decay of coefficients)**.** How fast the coefficients b_n decay reflects how smooth ϕ is, and governs how fast the series converges. A jump discontinuity forces the coefficients to decay only like $1/n$ (as in Examples 9.2 and 9.3); a continuous function with a corner gives decay like $1/n^2$; an infinitely smooth periodic function gives faster than any power. This is a useful diagnostic in both directions: rough functions have slowly-decaying, hard-to-sum Fourier series, while the rapid decay of a smooth function's coefficients is exactly why the high modes in the heat series (8.5) contribute so little after a short time.

9.5 Differentiating and integrating Fourier series

A practical question: may one differentiate a Fourier series term by term? Not always — and the reason connects to the decay of coefficients. Differentiating $\sin(n\pi x/l)$ brings down a factor $n\pi/l$, so term-by-term differentiation of $\sum b_n \sin(n\pi x/l)$ produces $\sum b_n(n\pi/l)\cos(n\pi x/l)$, whose coefficients are *larger* by a factor n. If the original coefficients decayed only like $1/n$, the differentiated series has coefficients that do not decay at all and the series diverges. Differentiation is legitimate only when the original series is smooth enough — roughly, when ϕ is continuous (no jumps in the periodic extension) so the coefficients decay fast enough to survive the extra factor of n. *Integration*, by contrast, is always safe: it divides coefficients by n, improving convergence. The asymmetry is a useful rule of thumb — integrating a Fourier series smooths it and is harmless, differentiating it roughens it and must be justified — and it is the series counterpart of the fact that differentiation amplifies high frequencies while integration suppresses them.

9.6 Completeness

The statement that *every* square-integrable function on the interval can be expanded in the eigenfunctions — that no function is "missed" — is the *completeness* of the eigenfunction system. Completeness is what guarantees that the separation-of-variables series can match *any* reasonable initial data, and hence that the method of Chapter 8 actually solves the initial-boundary-value problem rather than only special cases. It is expressed by *Parseval's identity*,

$$\int_0^l \phi(x)^2\,dx = \frac{l}{2}\sum_{n=1}^{\infty} b_n^2 \qquad \text{(sine case)},$$

an "infinite Pythagorean theorem": the squared length of ϕ equals the sum of the squared lengths of its components along the orthogonal eigenfunctions. Parseval's identity is also the precise sense in which the Fourier series captures all of ϕ's energy, and it furnishes a charming way to evaluate certain infinite sums in closed form, as the exercises show. We will see in Chapter 10 that completeness holds not just for sines and cosines but for the eigenfunctions of any Sturm–Liouville problem, which is why separation of variables works for the whole range of boundary conditions.

9.7 The Gibbs phenomenon

Near a jump discontinuity, the partial sums of a Fourier series overshoot the jump by a fixed percentage — about 9% of the jump's height (more precisely, the overshoot tends to about 8.95%, governed by the integral $\frac{2}{\pi}\int_0^{\pi} \frac{\sin s}{s}\,ds \approx 1.179$) — no matter how many terms are taken. The overshoot does not shrink as more terms are added; it merely narrows, crowding closer to the discontinuity. This is the *Gibbs phenomenon*. It is not a numerical error but a genuine feature of how a series of smooth functions approximates a discontinuous one: the partial sums converge to ϕ pointwise away from the jump and to the midpoint at the jump, but they cannot converge *uniformly* across it, and the persistent overshoot is the visible signature of that failure. The phenomenon matters in practice — it limits the fidelity of Fourier-based approximation of functions with sharp features, from signal processing to the numerical solution of PDEs with discontinuous data — and it is a vivid reminder that the three notions of convergence in Section 9.4 are genuinely different.

9.8 Inhomogeneous boundary conditions

A practical remark closes the chapter. Separation of variables requires *homogeneous* boundary conditions, because the argument that $X(0) = 0$ relied on the boundary value being zero. When the boundary data is nonzero — ends held at fixed nonzero temperatures, say — one first subtracts off a particular function (often the steady state) that carries the inhomogeneous boundary values, reducing the problem to one with homogeneous conditions for the remainder, which separation of variables then solves. For ends held at constant temperatures T_1 and T_2, the steady state is the straight line interpolating them, and the transient remainder is expanded in the sine series. This subtract-the-steady-state device recurs whenever boundary data or sources are inhomogeneous, and it is the bridge between the homogeneous theory of this chapter and the inhomogeneous problems of applications.

Exercises

Exercise 9.1 (*Orthogonality*). Prove the orthogonality relation (9.2) using the product-to-sum identity $\sin A \sin B = \frac{1}{2}[\cos(A-B) - \cos(A+B)]$. Compute the diagonal value $\int_0^l \sin^2(n\pi x/l)\,dx$ separately.

Exercise 9.2 (*A sine series*). Find the Fourier sine series of $\phi(x) = x$ on $0 < x < l$ (Example 9.3). Do the coefficients decay like $1/n$ or $1/n^2$, and what does that say about the smoothness of the odd periodic extension of x?

Exercise 9.3 (*A cosine series*). Find the Fourier cosine series of $\phi(x) = x$ on $0 < x < l$, including the constant term. Why does the cosine series of x converge faster at the endpoints than the sine series?

Exercise 9.4 (*Sine series of a constant*). Verify the expansion of Example 9.2, and evaluate the resulting series at $x = l/2$ to obtain a classical formula for $\pi/4$. *(This is the Leibniz series.)*

Exercise 9.5 (*Complex form*). Write the full Fourier series of $\phi(x) = x$ on $-l < x < l$ in complex exponential form, computing the coefficients $c_n = \frac{1}{2l}\int_{-l}^{l} x\, e^{-in\pi x/l}\,dx$. Relate c_n to the sine coefficients found earlier.

Exercise 9.6 (*Convergence at a jump*). The sine series of the constant 1 on $(0, l)$ vanishes at $x = 0$ and $x = l$. Reconcile this with Dirichlet's theorem on pointwise convergence at the endpoints, treating $x = 0$ as a jump of the odd periodic extension.

Exercise 9.7 (*Differentiating a series*). Explain why the sine series of $\phi(x) = x$ on $(0, l)$ cannot be differentiated term by term to give a convergent series. Which property of the coefficients fails after differentiation?

Exercise 9.8 (*Three kinds of convergence*). State the difference between pointwise, uniform, and mean-square convergence of a Fourier series, and give the smoothness hypothesis under which each is guaranteed. Which notion requires the least of ϕ?

Exercise 9.9 (*Parseval*). For the sine series of $\phi(x) = 1$ on $(0, l)$ (Example 9.2), write Parseval's identity explicitly and use it to evaluate $\sum_{n \text{ odd}} 1/n^2$.

Exercise 9.10 (*Gibbs phenomenon*). Explain why the Gibbs overshoot near a jump does not disappear as more Fourier terms are added, and what does happen to it. Which notion of convergence fails across the jump, and which still hold?

Exercise 9.11 (*Inhomogeneous boundary data*). To solve $u_t = k u_{xx}$ on $0 < x < l$ with ends held at $u(0,t) = T_1$ and $u(l,t) = T_2$, one subtracts a function $v(x)$ carrying the boundary values. Find the appropriate $v(x)$ (the steady state) and state the homogeneous problem the remainder $u - v$ satisfies.

Chapter 10

Sturm–Liouville Theory

The eigenvalue problems of Chapter 8 — $X''+\lambda X = 0$ with various boundary conditions — all shared a remarkable list of properties: real eigenvalues, orthogonal eigenfunctions, and a completeness that let any reasonable function be expanded in them. These properties are not peculiar to the sine and cosine functions. They hold for a whole class of eigenvalue problems, the *Sturm–Liouville problems*, and the general theory explains *why* separation of variables works across the full range of boundary conditions and even for equations with variable coefficients. This chapter, which may be treated as optional ($\star$) on a first pass, develops that theory at the level of statement and example rather than full proof.

10.1 The Sturm–Liouville problem

Separation of variables on more general equations — a rod of varying material properties, problems in cylindrical or spherical geometry — produces a spatial eigenvalue problem of the form

$$-\big(p(x)\,X'\big)' + q(x)\,X = \lambda\, w(x)\,X, \qquad a < x < b, \tag{10.1}$$

together with boundary conditions at $x = a$ and $x = b$. Here $p(x) > 0$, $w(x) > 0$ (the *weight*), and $q(x)$ are given functions determined by the equation and geometry. This is the *Sturm–Liouville* (S–L) form. The constant-coefficient problem $X'' + \lambda X = 0$ is the special case $p = w = 1$, $q = 0$, so everything in Chapter 8 is a Sturm–Liouville problem in disguise.

The boundary conditions are required to be of a type — Dirichlet, Neumann, Robin, or periodic — that makes the problem *self-adjoint*, a symmetry property of the differential operator analogous to a symmetric matrix. A problem in S–L form with such boundary conditions is called *regular* if $p, w > 0$ on the closed interval and the interval is finite. Regularity is the hypothesis under which the strong conclusions below hold; the singular case (where p or w vanishes at an endpoint, as in Bessel's equation) is richer and is where the special functions of mathematical physics arise.

10.2 The main theorem

The reason Sturm–Liouville theory matters is the following list of guarantees, which generalize every property we observed for sines and cosines.

Theorem 10.1 (Sturm–Liouville, regular case)**.** *For a regular Sturm–Liouville problem* (10.1) *with self-adjoint boundary conditions:*

1. *The eigenvalues are* real *and form an increasing sequence* $\lambda_1 < \lambda_2 < \cdots \to \infty$.

2. *To each eigenvalue corresponds (up to a constant multiple) a single eigenfunction X_n; the nth eigenfunction has exactly $n-1$ zeros in the open interval.*

3. *Eigenfunctions for different eigenvalues are* orthogonal *with respect to the weight w:*

$$\int_a^b X_m(x)\,X_n(x)\,w(x)\,dx = 0, \qquad m \neq n.$$

4. *The eigenfunctions are* complete*: every square-integrable function on (a, b) can be expanded in a (mean-square convergent) series $\sum_n c_n X_n$.*

Each conclusion has a familiar shadow. The reality and discreteness of the eigenvalues (1) is why separation of variables yields a sequence of decaying or oscillating modes indexed by n. The zero-counting (2) generalizes the fact that $\sin(n\pi x/l)$ has $n-1$ interior zeros — higher modes wiggle more. Orthogonality (3) is exactly what let us compute Fourier coefficients by projection (Section 9.1), now with a weight w when the equation demands it. Completeness (4) is the guarantee, promised in Chapter 9, that the expansion can represent any reasonable initial data — and it holds not just for sines and cosines but for the eigenfunctions of *every* regular S–L problem.

10.3 Why orthogonality holds

The orthogonality in Theorem 10.1 is not a lucky accident of the sine functions; it follows from the self-adjoint structure, by an argument worth seeing because it is short and explains the role of the boundary conditions. Write the S–L operator as $\mathcal{L}X = -(pX')' + qX$, so the problem is $\mathcal{L}X = \lambda w X$. For two eigenfunctions X_m, X_n with eigenvalues λ_m, λ_n, a double integration by parts gives *Lagrange's identity*,

$$\int_a^b \big(X_m \mathcal{L} X_n - X_n \mathcal{L} X_m\big)\,dx = \Big[\,p\,(X_n X_m' - X_m X_n')\,\Big]_a^b,$$

and the self-adjoint boundary conditions are precisely those making the right-hand boundary term vanish. The left side is then $(\lambda_n - \lambda_m)\int_a^b X_m X_n w\,dx$ (using the eigenvalue equation), so when $\lambda_m \neq \lambda_n$ the weighted integral must vanish — orthogonality. The same computation, with $X_n = \overline{X_m}$, shows the eigenvalues are real. Thus reality and orthogonality both flow from self-adjointness, and the boundary conditions earn their "self-adjoint" name by making the boundary term disappear — exactly as the homogeneous Dirichlet, Neumann, and Robin conditions did in the energy arguments of Chapter 5.

10.4 The Rayleigh quotient

Sturm–Liouville theory has a variational side that both estimates eigenvalues and reveals their sign. Multiplying the eigenvalue equation $\mathcal{L}X = \lambda w X$ by X, integrating, and integrating by parts once gives, for an eigenfunction X with eigenvalue λ,

$$\lambda = \frac{\big[-pXX'\big]_a^b + \displaystyle\int_a^b \big(p\,(X')^2 + qX^2\big)\,dx}{\displaystyle\int_a^b X^2\,w\,dx}.$$

This expression is the *Rayleigh quotient.* When the boundary term vanishes (Dirichlet or Neumann) and $q \geq 0$, the numerator is non-negative, so *every eigenvalue is non-negative* — which is why the separation constants in the heat and wave equations came out positive, giving decaying or oscillating (not growing) modes. Beyond fixing the sign, the Rayleigh quotient gives a powerful estimate: the *smallest* eigenvalue λ_1 is the minimum of the quotient over all admissible functions, so plugging in any trial function that satisfies the boundary conditions yields an upper bound for λ_1. This variational characterization is the engine behind many practical eigenvalue estimates and is the finite-dimensional shadow of the fact that, for a symmetric matrix, the least eigenvalue minimizes the Rayleigh quotient $x^\top Ax/x^\top x$.

10.5 Eigenfunction expansions and generalized Fourier series

Completeness licenses the *generalized Fourier series*: to expand a function ϕ in the eigenfunctions of a Sturm–Liouville problem, write

$$\phi(x) = \sum_{n=1}^{\infty} c_n X_n(x), \qquad c_n = \frac{\displaystyle\int_a^b \phi\, X_n\, w\, dx}{\displaystyle\int_a^b X_n^2\, w\, dx},$$

the coefficient formula obtained, exactly as for sines, by multiplying by $X_n w$, integrating, and using orthogonality to kill all but one term. Solving a separated PDE then proceeds as in Chapter 8: expand the initial data in the eigenfunctions, attach to each the appropriate time factor (decaying for diffusion, oscillating for waves), and superpose. The whole apparatus of separation of variables — Fourier coefficients, mode-by-mode evolution, long-time dominance of the lowest mode — carries over verbatim, with $\sin(n\pi x/l)$ replaced by the S–L eigenfunctions X_n and ordinary integration replaced by integration against the weight.

Example 10.2 (The Robin problem revisited)**.** The Robin eigenvalue problem $X'' + \lambda X = 0$ with $X'(0) = 0$ and $X'(l) + \beta X(l) = 0$ ($\beta > 0$) is a regular Sturm–Liouville problem ($p = w = 1$, $q = 0$). Its eigenvalues solve a transcendental equation — $\beta = -\sqrt{\lambda}\tan(\sqrt{\lambda}\, l)$ or similar — with no closed form, so the eigenfunctions are sines and cosines with *non-integer* frequencies determined numerically. Yet Theorem 10.1 guarantees, without our solving the transcendental equation, that the eigenvalues are real and increasing, the eigenfunctions orthogonal, and the system complete — so the Robin problem can be solved by eigenfunction expansion exactly like the Dirichlet and Neumann problems, even though we cannot write its eigenvalues in closed form. This is the practical power of the abstract theorem: it certifies that the method works before any explicit computation.

10.6 Singular problems and the special functions

The richest applications of Sturm–Liouville theory are *singular*: the coefficient p or the weight w vanishes at an endpoint, or the interval is infinite. These arise the moment one separates variables in a curved geometry.

Example 10.3 (Bessel's equation from the disk*)**.** Separating the heat or wave equation in a disk (polar coordinates r, θ) produces, for the radial factor, *Bessel's equation*

$$\big(rR'\big)' + \Big(\lambda r - \frac{n^2}{r}\Big)R = 0, \qquad 0 < r < a,$$

which is in Sturm–Liouville form (10.1) with $p(r) = r$, weight $w(r) = r$, and $q(r) = n^2/r$. It is *singular* because $p(r) = r$ vanishes at the center $r = 0$. Its bounded solutions are the *Bessel functions* $J_n(\sqrt{\lambda}\, r)$, and the boundary condition at $r = a$ (say $R(a) = 0$) quantizes λ through the zeros of J_n. Despite the singularity, the essential Sturm–Liouville conclusions survive: the Bessel functions for different eigenvalues are orthogonal *with weight* r, $\int_0^a J_n(\sqrt{\lambda_j}\, r) J_n(\sqrt{\lambda_k}\, r)\, r\, dr = 0$ for $j \neq k$, and they are complete, so a radial profile can be expanded in a *Fourier–Bessel series* exactly as an interval profile is expanded in sines. The vibrations of a circular drumhead are governed by precisely this expansion.

Remark 10.4 (Where this leads)**.** Sturm–Liouville theory is the gateway to the special functions of mathematical physics. Separation of variables for the wave or heat equation in a disk produces Bessel's equation, as above; in a sphere it produces *Legendre's equation* and the Legendre polynomials. In each case the same four conclusions — real eigenvalues, orthogonality (with the appropriate weight), completeness — justify expanding data in these exotic eigenfunctions exactly as we expand in sines. We touch this circle of ideas again when separating Laplace's equation in a disk (Chapter 13); the deeper theory belongs to a later course, but its foundation is the Sturm–Liouville theorem stated here.

Example 10.5 (Locating Robin eigenvalues)**.** Consider $X'' + \lambda X = 0$ on $0 < x < 1$ with $X(0) = 0$ and $X'(1) + X(1) = 0$ (a Dirichlet end and a Robin end with $\beta = 1$). The condition $X(0) = 0$ gives $X = \sin(\sqrt{\lambda}\, x)$, and imposing the Robin condition at $x = 1$ yields the transcendental equation

$$\sqrt{\lambda}\, \cos\sqrt{\lambda} + \sin\sqrt{\lambda} = 0, \quad \text{i.e.} \quad \tan\sqrt{\lambda} = -\sqrt{\lambda}.$$

This has no closed-form solution, but its roots $\mu_n = \sqrt{\lambda_n}$ are the intersections of $y = \tan\mu$ with the line $y = -\mu$, one in each interval $(\frac{(2n-1)\pi}{2}, n\pi)$. Numerically the first few are $\mu_1 \approx 2.029$, $\mu_2 \approx 4.913$, $\mu_3 \approx 7.979$, so $\lambda_1 \approx 4.12$, $\lambda_2 \approx 24.1$, $\lambda_3 \approx 63.7$. For large n the roots approach $\mu_n \approx (n - \frac{1}{2})\pi$, the Dirichlet–Neumann values, since the Robin term βX becomes negligible compared to X' at high frequency. The Sturm–Liouville theorem (Theorem 10.1) guarantees in advance that these eigenvalues are real, simple, and increase to infinity, and that the corresponding $\sin(\mu_n x)$ are orthogonal and complete — so a function on $(0, 1)$ may be expanded in them exactly as in a sine series, despite the absence of a closed form for the μ_n.

Exercises

Exercise 10.1 (*Recognizing S–L form*). Write the constant-coefficient problem $X'' + \lambda X = 0$ in Sturm–Liouville form (10.1), identifying p, q, and w. Is it regular on a finite interval?

Exercise 10.2 (*Putting an equation in S–L form*). Show that the equation $X'' + 2X' + \lambda X = 0$ can be put in Sturm–Liouville form by multiplying through by an integrating factor e^{2x}, and identify the resulting p, q, w.

Exercise 10.3 (*Counting zeros*). For the Dirichlet problem on $(0, l)$, verify that the nth eigenfunction $\sin(n\pi x/l)$ has exactly $n - 1$ zeros in the open interval, confirming conclusion (2) of Theorem 10.1.

Exercise 10.4 (*Orthogonality with a weight*). Suppose a Sturm–Liouville problem has weight $w(x) = x$. Write the orthogonality relation its eigenfunctions satisfy, and the coefficient formula for expanding a function ϕ in them.

Exercise 10.5 (*Lagrange's identity*). Carry out the double integration by parts to establish Lagrange's identity $\int_a^b (X_m \mathcal{L} X_n - X_n \mathcal{L} X_m)\, dx = [p(X_n X_m' - X_m X_n')]_a^b$ for $\mathcal{L}X = -(pX')' + qX$.

Exercise 10.6 (*Boundary terms vanish*). Show that homogeneous Dirichlet conditions ($X = 0$ at both ends) make the boundary term in Lagrange's identity vanish. Repeat for homogeneous Neumann conditions ($X' = 0$ at both ends).

Exercise 10.7 (*The Rayleigh quotient*). Write the Rayleigh quotient for the Dirichlet problem $X'' + \lambda X = 0$ on $(0, l)$, and use the trial function $X(x) = x(l - x)$ to obtain an upper bound for the smallest eigenvalue $\lambda_1 = (\pi/l)^2$. How close is the bound?

Exercise 10.8 (*Reality of eigenvalues*). Using Lagrange's identity with X_n and its complex conjugate, sketch why the eigenvalues of a self-adjoint S–L problem must be real.

Exercise 10.9 (*Robin eigenvalues*). For the Robin problem of Example 10.2, write the transcendental equation the eigenvalues satisfy. Explain why Theorem 10.1 guarantees an eigenfunction expansion exists even though the eigenvalues have no closed form.

Exercise 10.10 (*Bessel orthogonality**). For Bessel's equation (Example 10.3), state the weight with respect to which the radial eigenfunctions are orthogonal, and write the orthogonality relation. Why does the weight r appear rather than the constant weight of the interval problem?

Part IV

Transform Methods

Chapter 11

The Fourier Transform

On a finite interval, a function is built from a discrete set of frequencies — the Fourier series. On the whole line $-\infty < x < \infty$ there is no fundamental period, and the discrete sum gives way to a continuous integral over all frequencies: the *Fourier transform*. This is the natural tool for PDEs on unbounded spatial domains, and it will, among other things, produce the heat kernel of Chapter 6 systematically rather than by inspired guessing. We define the transform, establish the properties that make it useful for differential equations, and apply it to solve the diffusion and wave equations on the line.

11.1 From series to transform

A Fourier series represents a function on $(-l, l)$ as a sum over the discrete frequencies $n\pi/l$. As the interval grows, $l \to \infty$, these frequencies crowd together and the sum becomes an integral over a continuous frequency variable. The limiting object represents a function on all of $\mathbb{R}$ as a continuous superposition of oscillations $e^{i\xi x}$, weighted by a function $\hat{f}(\xi)$ that records "how much" of each frequency ξ is present. We take the resulting formulas as definitions.

Definition 11.1 (Fourier transform)**.** For a function f on $\mathbb{R}$ (decaying suitably at infinity), the *Fourier transform* is

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x)\, e^{-i\xi x}\, dx, \tag{11.1}$$

and f is recovered by the *inverse Fourier transform*

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\xi)\, e^{i\xi x}\, d\xi. \tag{11.2}$$

We use the frequency variable ξ (not k, which is reserved for the diffusion constant), matching the notation of the course's examinations. The asymmetry of the constants — no factor in the forward transform, $1/2\pi$ in the inverse — is one of several conventions in use; what matters is that the forward and inverse together reproduce f, and we fix (11.1) and (11.2) once and for all.

11.2 The two properties that matter for PDEs

Two properties of the Fourier transform make it a tool for differential equations. Both are consequences of (11.1) and a little calculus.

Differentiation becomes multiplication. Transforming a derivative multiplies by $i\xi$:

$$\widehat{f'}(\xi) = i\xi\,\hat{f}(\xi), \qquad \widehat{f''}(\xi) = (i\xi)^2\hat{f}(\xi) = -\xi^2\,\hat{f}(\xi). \tag{11.3}$$

This is the property on which everything turns. It follows by integrating (11.1) by parts: the derivative on f transfers to the exponential $e^{-i\xi x}$, which differentiates to $-i\xi\,e^{-i\xi x}$, and the boundary terms vanish because f decays. The transform thus converts differentiation in x into multiplication by a number — turning a differential equation in x into an *algebraic* one in ξ.

Convolution becomes multiplication. The Fourier transform of a convolution

$$(f * g)(x) = \int_{-\infty}^{\infty} f(x-y)\,g(y)\,dy$$

is the *product* of the transforms:

$$\widehat{f * g}(\xi) = \hat{f}(\xi)\,\hat{g}(\xi). \tag{11.4}$$

This *convolution theorem* explains, after the fact, why the solution of the diffusion equation on the line came out as a convolution (Chapter 6): in frequency space the solution is a simple product, and a product of transforms is the transform of a convolution.

11.3 Shifting, scaling, and other rules

A handful of further rules make the transform genuinely usable. Each follows from (11.1) by an elementary change of variable, and together they form a small toolkit.

$$\begin{aligned}
&\text{(shift in } x\text{)} && \widehat{f(x-a)}(\xi) = e^{-ia\xi}\hat{f}(\xi),\\
&\text{(modulation)} && \widehat{e^{ibx}f(x)}(\xi) = \hat{f}(\xi-b),\\
&\text{(scaling)} && \widehat{f(ax)}(\xi) = \tfrac{1}{|a|}\hat{f}(\xi/a),\\
&\text{(symmetry)} && f \text{ real and even} \implies \hat{f} \text{ real and even.}
\end{aligned}$$

The shift rule says that translating f multiplies its transform by a pure phase $e^{-ia\xi}$ — the frequencies are unchanged, only their phases shift, which is just what one expects of a rigid translation. The scaling rule expresses the reciprocal relationship between the widths of f and $\hat{f}$: compressing f (large a) stretches $\hat{f}$, the uncertainty principle in algebraic form. These rules let one transform complicated functions by decomposing them into shifts, scalings, and modulations of known ones.

11.4 The Gaussian and its transform

One computation is indispensable: the Fourier transform of a Gaussian is again a Gaussian. Precisely, for $a > 0$,

$$f(x) = e^{-ax^2} \implies \hat{f}(\xi) = \sqrt{\frac{\pi}{a}}\,e^{-\xi^2/4a}. \tag{11.5}$$

The transform of a narrow Gaussian (large a) is a wide one (the exponent $\xi^2/4a$ is small), and vice versa — a manifestation of the uncertainty principle, that a function and its transform cannot both be sharply localized. The Gaussian's self-similarity under the transform is exactly what makes the heat kernel Gaussian, as we now see.

11.5 Plancherel's theorem

The Fourier transform preserves energy, in the following precise sense, which is the whole-line analogue of Parseval's identity for Fourier series (Section 9.6).

Theorem 11.2 (Plancherel)**.** *For suitable f,*

$$\int_{-\infty}^{\infty} |f(x)|^2 \, dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\hat{f}(\xi)|^2 \, d\xi.$$

The total energy of a signal, computed in physical space, equals (up to the 2π) its total energy computed in frequency space: the transform merely redistributes the energy among frequencies without creating or destroying it. This is why the Fourier transform is the natural language for energy methods on the line, and it is indispensable in signal processing, where it says that the power of a signal can be read off equally from its waveform or its spectrum. For the PDEs of this course, Plancherel's theorem underlies the stability estimates: if the transform of the solution is controlled, so is the solution itself.

11.6 Solving the diffusion equation on the line

We can now solve $u_t = ku_{xx}$ on $-\infty < x < \infty$ with $u(x, 0) = \phi(x)$ *systematically.* Take the Fourier transform of the equation in the spatial variable x, writing $\hat{u}(\xi, t)$ for the transform of $u(\cdot, t)$. The time derivative passes through the transform, and the second x-derivative becomes multiplication by $-\xi^2$ via (11.3):

$$\hat{u}_t(\xi, t) = k\,(-\xi^2)\,\hat{u}(\xi, t) = -k\xi^2\,\hat{u}(\xi, t).$$

For each fixed ξ this is an *ordinary* differential equation in t — the PDE has become a family of ODEs indexed by frequency, just as separation of variables produced one ODE per mode. Its solution is

$$\hat{u}(\xi, t) = \hat{\phi}(\xi)\, e^{-k\xi^2 t},$$

with the initial value $\hat{u}(\xi, 0) = \hat{\phi}(\xi)$ the transform of the initial data. The transform of the solution is thus the transform of the data times the Gaussian factor $e^{-k\xi^2 t}$. By the convolution theorem (11.4), multiplication in frequency is convolution in space, and the inverse transform of the Gaussian $e^{-k\xi^2 t}$ is — by (11.5) — precisely the heat kernel. We recover

$$u(x, t) = (S(\cdot, t) * \phi)(x) = \frac{1}{\sqrt{4\pi kt}} \int_{-\infty}^{\infty} e^{-(x-y)^2/4kt} \phi(y)\, dy, \tag{11.6}$$

the formula of Chapter 6, now *derived* rather than guessed. The heat kernel is Gaussian because the time factor $e^{-k\xi^2 t}$ is Gaussian in ξ, and the Gaussian transforms to a Gaussian.

Remark 11.3 (Why the transform method is illuminating)**.** The transform solution explains features that the kernel alone merely exhibits. Each frequency ξ decays in time at rate $k\xi^2$ — so high frequencies (large ξ, sharp features) decay fastest, which *is* the instantaneous smoothing of Chapter 6, now quantified mode by mode. The parallel with the Fourier-series solution on an interval (Chapter 8), where the nth mode decayed at rate $k(n\pi/l)^2$, is exact: the transform is the whole-line version of separation of variables, with a continuum of frequencies in place of the discrete modes.

11.7 Solving the wave equation on the line

The same method handles the wave equation $u_{tt} = c^2 u_{xx}$ on the line, and it reproduces d'Alembert's formula from the frequency side. Transforming in x turns u_{xx} into $-\xi^2 \hat{u}$, giving for each ξ the ODE

$$\hat{u}_{tt}(\xi, t) = -c^2 \xi^2 \, \hat{u}(\xi, t),$$

a harmonic oscillator in t with frequency $c|\xi|$. Its solution is $\hat{u}(\xi, t) = \hat{\phi}(\xi) \cos(c\xi t) + \hat{\psi}(\xi) \frac{\sin(c\xi t)}{c\xi}$, matching the transformed initial data $\hat{\phi}$ and $\hat{\psi}$. Now $\cos(c\xi t) = \frac{1}{2}(e^{ic\xi t} + e^{-ic\xi t})$, and by the shift rule (Section 11.3) multiplication of a transform by $e^{\pm ic\xi t}$ corresponds to translation by $\mp ct$. Inverting therefore turns the $\hat{\phi}$ term into $\frac{1}{2}[\phi(x + ct) + \phi(x - ct)]$ — the two half-amplitude travelling copies of d'Alembert's formula — and the $\hat{\psi}$ term into the velocity integral. Where diffusion gave a single decaying exponential per frequency, the wave equation gives an undamped oscillation, $\cos(c\xi t)$, whose persistence in time is the frequency-space face of energy conservation and of propagation without decay.

11.8 The transform method in general

The pattern generalizes to any constant-coefficient linear PDE on the line. Transform in the spatial variable; derivatives in x become powers of $i\xi$; the PDE becomes an ODE (or algebraic equation) in the remaining variable for each ξ; solve it; invert the transform. Its limitation is the requirement of an unbounded domain and constant coefficients — on a finite interval the discrete Fourier *series* is the right tool, and variable coefficients break the clean conversion of differentiation into multiplication. Within its domain, however, the Fourier transform is the most systematic solution method in the subject.

Exercises

Exercise 11.1 (*A basic transform*). Compute the Fourier transform of $f(x) = e^{-a|x|}$ for $a > 0$. *(Split the integral at $x = 0$.)* Is the result a Gaussian?

Exercise 11.2 (*Differentiation rule*). Prove the differentiation rule $\widehat{f'}(\xi) = i\xi \, \hat{f}(\xi)$ by integrating (11.1) by parts, stating where the decay of f at infinity is used.

Exercise 11.3 (*Shift and scale*). Prove the shift rule $\widehat{f(x-a)} = e^{-ia\xi} \hat{f}(\xi)$ and the scaling rule $\widehat{f(ax)} = \frac{1}{|a|} \hat{f}(\xi/a)$ by changing variables in (11.1).

Exercise 11.4 (*The Gaussian transform*). Verify (11.5) for $f(x) = e^{-ax^2}$. *(One route: show $\hat{f}$ satisfies the ODE $\hat{f}'(\xi) = -\frac{\xi}{2a} \hat{f}(\xi)$ by differentiating under the integral and integrating by parts, then solve it with $\hat{f}(0) = \sqrt{\pi/a}$.)*

Exercise 11.5 (*Convolution theorem*). Starting from the definition of convolution, prove $\widehat{f * g} = \hat{f} \, \hat{g}$. *(Interchange the order of integration.)*

Exercise 11.6 (*Plancherel*). State Plancherel's theorem and explain what it says about the energy of a signal in physical versus frequency space. How is it the whole-line analogue of Parseval's identity for Fourier series?

Exercise 11.7 (*Solving the heat equation*). Carry out the Fourier-transform solution of $u_t = ku_{xx}$ on the line in full: transform, solve the ODE in t for each ξ, and invert to recover the heat-kernel convolution. Identify where the Gaussian transform (11.5) is used.

Exercise 11.8 (*Solving the wave equation*). Solve $u_{tt} = c^2 u_{xx}$ on the line by the Fourier transform, obtaining $\hat{u} = \hat{\phi}\cos(c\xi t) + \hat{\psi}\frac{\sin(c\xi t)}{c\xi}$, and invert the $\hat{\phi}$ term using the shift rule to recover the travelling-wave part of d'Alembert's formula.

Exercise 11.9 (*Frequency-by-frequency decay*). Explain, from $\hat{u}(\xi, t) = \hat{\phi}(\xi)e^{-k\xi^2 t}$, why high spatial frequencies decay faster than low ones, and how this realizes the smoothing property of the diffusion equation.

Exercise 11.10 (*The transport equation by transform*). Solve $u_t + cu_x = 0$ on the line by the Fourier transform: show $\hat{u}(\xi, t) = \hat{\phi}(\xi)e^{-ic\xi t}$, and invert to recover $u(x, t) = \phi(x - ct)$. *(Use the shift rule.)*

Exercise 11.11 (*Uncertainty*). Using the Gaussian pair (11.5), explain the sense in which a function and its Fourier transform cannot both be arbitrarily narrow. What happens to the width of $\hat{f}$ as f is made narrower (larger a)?

Chapter 12

The Laplace Transform

The Fourier transform suits problems on the whole spatial line. The *Laplace transform* suits problems on the half-line $t \geq 0$ in the *time* variable — problems with initial conditions, where we know the state at $t = 0$ and want the evolution for $t > 0$. It converts initial-value problems for ODEs into algebra and initial-boundary-value problems for PDEs into ODEs in the spatial variable. This chapter defines the transform, develops the derivative rule that makes it useful, builds a working table, and applies it to differential equations, including the second-order problems that appear on the course's examinations.

12.1 Definition and basic transforms

Definition 12.1 (Laplace transform)**.** For a function $f(t)$ defined for $t \geq 0$, the *Laplace transform* is

$$F(s) = \mathcal{L}\{f\}(s) = \int_0^\infty f(t)\, e^{-st}\, dt, \tag{12.1}$$

defined for those values of s (real, or complex with sufficiently large real part) for which the integral converges.

The transform replaces a function of t by a function of a new variable s. The exponential weight e^{-st} damps the integrand, so the transform exists for any f that does not grow too fast. A few basic transforms, computed directly from (12.1), recur constantly:

$$\mathcal{L}\{1\} = \frac{1}{s}, \qquad \mathcal{L}\{e^{at}\} = \frac{1}{s-a}, \qquad \mathcal{L}\{t^n\} = \frac{n!}{s^{n+1}},$$
$$\mathcal{L}\{\cos\omega t\} = \frac{s}{s^2+\omega^2}, \qquad \mathcal{L}\{\sin\omega t\} = \frac{\omega}{s^2+\omega^2}.$$

Each follows from an elementary integral; for instance $\mathcal{L}\{e^{at}\} = \int_0^\infty e^{(a-s)t}\, dt = 1/(s-a)$ for $s > a$.

12.2 Linearity and the derivative rule

The Laplace transform is *linear*: $\mathcal{L}\{\alpha f + \beta g\} = \alpha\mathcal{L}\{f\} + \beta\mathcal{L}\{g\}$, immediate from the linearity of the integral.

The property that makes it solve differential equations is its action on derivatives, which — unlike the Fourier rule — explicitly involves the *initial values*. Integrating (12.1) by parts,

$$\mathcal{L}\{f'\}(s) = sF(s) - f(0), \tag{12.2}$$

and applying this twice,

$$\mathcal{L}\{f''\}(s) = s^2 F(s) - s f(0) - f'(0). \tag{12.3}$$

Differentiation in t becomes multiplication by s, with the initial data $f(0)$ and $f'(0)$ entering as additive corrections. This is exactly what one wants for an initial-value problem: the transform builds the initial conditions into the algebra automatically, rather than leaving them as separate side-conditions to impose afterward.

12.3 The shifting theorems and the Heaviside function

Two shift rules greatly extend the table, and they are the Laplace counterparts of the Fourier shift and modulation rules.

Shift in s (multiplication by e^{at}). Multiplying f by an exponential shifts its transform:

$$\mathcal{L}\{e^{at} f(t)\}(s) = F(s-a).$$

Thus $\mathcal{L}\{e^{at}\cos\omega t\} = \frac{s-a}{(s-a)^2+\omega^2}$, obtained from the cosine transform by replacing s with $s-a$. This rule is what produces the transforms of damped oscillations, and hence what lets the transform solve damped-oscillator ODEs.

Shift in t (the Heaviside step). Delaying a function in time multiplies its transform by an exponential. Let $H(t)$ be the *Heaviside step function*, $H(t) = 0$ for $t < 0$ and 1 for $t \geq 0$. Then for a delay $a > 0$,

$$\mathcal{L}\{H(t-a) f(t-a)\}(s) = e^{-as} F(s).$$

The Heaviside function and this rule are the tools for problems with forcing that switches on at a later time — a source applied at $t = a$, a voltage suddenly connected — which are common in applications and which the step function expresses cleanly.

12.4 The Dirac delta and impulse response

Closely related is the *Dirac delta* $\delta(t)$, the idealized "impulse" — an infinitely brief, infinitely strong unit kick, the derivative of the Heaviside step. It is not a function in the ordinary sense but is handled by the rule

$$\int_{-\infty}^{\infty} \delta(t-a)\, g(t)\, dt = g(a),$$

that integrating against $\delta(t-a)$ samples g at $t = a$. Its Laplace transform is especially simple:

$$\mathcal{L}\{\delta(t)\}(s) = 1.$$

The delta represents an instantaneous unit input, and the response of a system to it — the *impulse response* — is fundamental: because a general forcing can be written as a superposition of impulses, $f(t) = \int f(s)\delta(t-s)\, ds$, the response to any forcing is the convolution of f with the impulse response. This is Duhamel's principle (Chapter 7) in transform language, and it connects to the convolution theorem below.

12.5 The convolution theorem

Like the Fourier transform, the Laplace transform turns convolution into multiplication. For functions on $t \geq 0$, the convolution is

$$(f * g)(t) = \int_0^t f(t-\tau)\, g(\tau)\, d\tau,$$

and the convolution theorem states

$$\mathcal{L}\{f * g\}(s) = F(s)\, G(s).$$

This is the key to inverting products: if $H(s) = F(s)G(s)$ is a product of two known transforms, then $h(t) = (f * g)(t)$ is the convolution of their inverses. It also gives the solution of a forced linear ODE as a convolution of the forcing with the impulse response, making precise the impulse-superposition idea of the previous section.

12.6 Solving initial-value problems for ODEs

The method for an ODE is a three-step recipe: transform the equation (turning it into algebra), solve for $F(s)$, and invert to recover $f(t)$. We illustrate with the two orders that appear in practice.

Example 12.2 (A first-order problem)**.** Solve $u'(t) + 3u(t) = e^{-t}$ with $u(0) = 2$. Transform both sides, using (12.2):

$$\big(sU(s) - 2\big) + 3U(s) = \frac{1}{s+1}.$$

Solve for U: $(s+3)U = 2 + \frac{1}{s+1}$, so

$$U(s) = \frac{2}{s+3} + \frac{1}{(s+1)(s+3)}.$$

A partial-fraction decomposition $\frac{1}{(s+1)(s+3)} = \frac{1/2}{s+1} - \frac{1/2}{s+3}$ and the table give

$$u(t) = 2e^{-3t} + \tfrac{1}{2}e^{-t} - \tfrac{1}{2}e^{-3t} = \tfrac{1}{2}e^{-t} + \tfrac{3}{2}e^{-3t}.$$

One checks $u(0) = \frac{1}{2} + \frac{3}{2} = 2$, as required.

Example 12.3 (A second-order problem)**.** Solve $u''(t) + u(t) = 0$ with $u(0) = 1$, $u'(0) = 0$ — the equation of simple harmonic motion. Transform using (12.3):

$$\big(s^2U - s \cdot 1 - 0\big) + U = 0 \quad \Longrightarrow \quad (s^2+1)U = s \quad \Longrightarrow \quad U(s) = \frac{s}{s^2+1}.$$

The table identifies $U(s) = \mathcal{L}\{\cos t\}$, so $u(t) = \cos t$. The second-order derivative rule (12.3), using *both* initial values, is what makes the second-order case go through; this is the structure tested in the course's second-order Laplace problems.

Example 12.4 (A damped oscillator via the shift rule)**.** Solve $u'' + 2u' + u = 0$ with $u(0) = 1$, $u'(0) = 0$. Transforming, $(s^2U - s) + 2(sU - 1) + U = 0$, so $(s^2 + 2s + 1)U = s + 2$ and

$$U(s) = \frac{s+2}{(s+1)^2} = \frac{1}{s+1} + \frac{1}{(s+1)^2}.$$

The shift rule (Section 12.3) inverts each term: $\frac{1}{s+1} \to e^{-t}$ and $\frac{1}{(s+1)^2} \to te^{-t}$ (shifting $\mathcal{L}\{t\} = 1/s^2$). Hence $u(t) = e^{-t} + te^{-t}$, the critically damped response. The repeated root $s = -1$ produces the factor t, the hallmark of critical damping — visible immediately in the transform.

12.7 Inversion

Solving for $F(s)$ is only half the work; one must recover $f(t)$ from $F(s)$. In a first course, inversion is done by recognizing $F(s)$ as a known transform, usually after *partial fractions* reduces a rational $F(s)$ to a sum of simple pieces, each matched to the table. The examples above are typical: the algebra produces a rational function of s, partial fractions break it into terms like $1/(s-a)$ and $s/(s^2+\omega^2)$, and the table reads off the inverse term by term. A general inversion formula exists — a contour integral in the complex s-plane, the Bromwich integral — but it requires complex analysis beyond this course; for our purposes the table, partial fractions, and the shift rules suffice.

12.8 Application to PDEs

For a PDE in x and t, transforming in t removes the time derivatives — converting them to algebra in s via (12.2)–(12.3) and absorbing the initial conditions — and leaves an *ordinary* differential equation in x for the transformed unknown $U(x,s)$, with s a parameter. One solves this ODE (applying the boundary conditions in x, which transform straightforwardly) and then inverts in s to recover $u(x,t)$.

Example 12.5 (A suddenly heated semi-infinite rod⋆)**.** A rod occupying $x>0$, initially at zero temperature, has its end suddenly raised to temperature 1 at $t=0$: solve $u_t = ku_{xx}$ for $x>0$, $t>0$, with $u(x,0)=0$, $u(0,t)=1$, and $u\to 0$ as $x\to\infty$. Transforming in t (so $u_t \to sU$, with zero initial data), the equation becomes the ODE $sU = kU_{xx}$, i.e. $U_{xx} = (s/k)U$, whose decaying solution is $U(x,s) = U(0,s)\,e^{-\sqrt{s/k}\,x}$. The boundary condition $u(0,t)=1$ transforms to $U(0,s)=1/s$, so $U(x,s) = \frac{1}{s}e^{-\sqrt{s/k}\,x}$. Inverting this transform (a standard table entry) gives

$$u(x,t) = \operatorname{erfc}\Big(\frac{x}{\sqrt{4kt}}\Big) = 1 - \operatorname{erf}\Big(\frac{x}{\sqrt{4kt}}\Big),$$

where $\operatorname{erfc} = 1 - \operatorname{erf}$ is the complementary error function. The heat penetrates from the boundary to a depth $\sim \sqrt{kt}$ — the diffusion length again — and the error function reappears, now generated by the Laplace method rather than the similarity argument of Chapter 6. The two routes to the error function, transform and similarity, illuminate the problem from different sides.

Remark 12.6 (Fourier or Laplace?)**.** The two transforms are complementary, and choosing between them is a matter of matching the transform to the domain. Use the *Fourier* transform for the spatial variable on an unbounded domain $-\infty < x < \infty$, where the natural data is decay at infinity. Use the *Laplace* transform for the time variable on $t \geq 0$, where the natural data is the initial state, or for a semi-infinite spatial domain with a boundary condition at one end. The derivative rules reflect the difference: the Fourier rule (11.3) involves no boundary data (the function decays), while the Laplace rule (12.2) explicitly carries the initial value, which is exactly the information an initial-value problem supplies.

Exercises

Exercise 12.1 (*Basic transforms*). Compute directly from the definition: (a) $\mathcal{L}\{1\}$; (b) $\mathcal{L}\{e^{at}\}$; (c) $\mathcal{L}\{t\}$. State the range of s for which each integral converges.

Exercise 12.2 (*Linearity and the table*). Find $\mathcal{L}\{3 + 2e^{-t} - \sin 2t\}$ using linearity and the table.

Exercise 12.3 (*The shift rules*). Use the s-shift rule to find $\mathcal{L}\{e^{-t}\cos 3t\}$, and the t-shift rule to find $\mathcal{L}\{H(t-2)\}$ where H is the Heaviside step.

Exercise 12.4 (*First-order IVP*). Solve $u'(t) - 2u(t) = 0$ with $u(0) = 5$ by the Laplace transform, and check against the elementary solution.

Exercise 12.5 (*First-order with forcing*). Solve $u'(t) + u(t) = t$ with $u(0) = 0$ by the Laplace transform. *(You will need $\mathcal{L}\{t\} = 1/s^2$ and partial fractions.)*

Exercise 12.6 (*Second-order IVP*). Solve $u''(t) + 4u(t) = 0$ with $u(0) = 0$, $u'(0) = 6$, using the second-derivative rule (12.3). Identify the resulting motion.

Exercise 12.7 (*A damped oscillator*). Solve $u''(t) + 2u'(t) + u(t) = 0$ with $u(0) = 1$, $u'(0) = 0$ (Example 12.4), and explain how the repeated root produces the te^{-t} term.

Exercise 12.8 (*The convolution theorem*). State the Laplace convolution theorem and use it to invert $F(s) = \frac{1}{s(s+1)}$ as a convolution. Check against the partial-fraction inverse.

Exercise 12.9 (*The derivative rule*). Derive the first-order rule $\mathcal{L}\{f'\} = sF(s) - f(0)$ by integrating $\int_0^\infty f'e^{-st}\,dt$ by parts, and explain where the initial value $f(0)$ enters. Then derive the second-order rule.

Exercise 12.10 (*Partial fractions*). Invert $F(s) = \dfrac{s+3}{(s+1)(s+2)}$ by partial fractions and the table.

Exercise 12.11 (*Fourier versus Laplace*). For each problem, state which transform is the natural tool and why: (a) the diffusion equation on $-\infty < x < \infty$; (b) an initial-value problem $u'' + u = f(t)$, $t \geq 0$; (c) the heat equation on a semi-infinite rod $0 < x < \infty$ with the end suddenly heated.

Part V

Laplace's Equation and Computation

Chapter 13

Laplace's Equation and Harmonic Functions

Laplace's equation $\Delta u = 0$ is the prototype of an elliptic equation and the one that describes *equilibrium* — a steady temperature distribution, an electrostatic potential, a stretched membrane at rest. It has no time variable; its solutions, the *harmonic functions*, are determined by data on the boundary of a region. This chapter develops the two structural properties that govern harmonic functions — the mean value property and the maximum principle — and then solves the equation explicitly by separation of variables in the two geometries that admit clean answers: the rectangle and the disk.

13.1 Harmonic functions and where they arise

A function u is *harmonic* in a region Ω if it satisfies Laplace's equation there,

$$\Delta u = u_{xx} + u_{yy} = 0 \quad \text{in } \Omega$$

(in two dimensions; in three, $\Delta u = u_{xx} + u_{yy} + u_{zz}$). We met the equation in Chapter 1 as the steady state of the diffusion equation: when $u_t = 0$, the heat equation $u_t = k\Delta u$ reduces to $\Delta u = 0$. Harmonic functions thus describe equilibrium temperature distributions — configurations that have stopped changing because heat flows in and out in balance at every point. The same equation governs electrostatic potentials in charge-free regions, gravitational potentials in empty space, the velocity potential of an incompressible irrotational flow, and the displacement of a soap film spanning a wire boundary. The recurrence of Laplace's equation across so many fields is one of the unifying facts of mathematical physics, and it makes the harmonic functions worth understanding in their own right.

The natural problem is the *Dirichlet problem*: given boundary values $u = g$ on $\partial\Omega$, find the harmonic function inside taking those values. Physically, fix the temperature around the edge of a plate and ask for the steady temperature inside. Unlike the evolution equations, there is no initial condition — the boundary data all around the region determines the interior completely, as the type-elliptic classification of Chapter 3 led us to expect.

13.2 The mean value property

Harmonic functions satisfy a striking averaging property that captures their essence.

Theorem 13.1 (Mean value property)**.** *If u is harmonic in a region containing a disk, then the value of u at the center of the disk equals the average of u over the bounding circle:*

$$u(center) = \frac{1}{2\pi}\int_0^{2\pi} u(center + re^{i\theta})\, d\theta,$$

and likewise equals the average over the solid disk. The same holds for spheres in three dimensions.

The value at a point is the average of the values around it: a harmonic function has no isolated bumps or dips. This is the equilibrium condition made visible — at a steady state, the temperature at a point is the average of its surroundings, which is exactly the condition for no net heat flow. The mean value property is the elliptic counterpart of conservation of energy (waves) and the maximum principle (diffusion): a structural fact from which much follows.

13.3 The maximum principle for harmonic functions

From the mean value property follows a maximum principle even cleaner than the parabolic one.

Theorem 13.2 (Maximum principle)**.** *A non-constant harmonic function on a bounded region Ω attains its maximum and minimum only on the boundary $\partial\Omega$, never at an interior point.*

Idea. If u attained an interior maximum at a point P, the mean value property would force u on every small circle around P to average to $u(P)$; but if $u(P)$ is the maximum, no value on the circle can exceed it, so u must equal $u(P)$ all around — and, propagating outward, throughout Ω. Hence a non-constant harmonic function has no interior maximum. The minimum follows by applying this to $-u$. □

The maximum principle delivers uniqueness and stability for the Dirichlet problem, exactly as it did for diffusion.

Corollary 13.3 (Uniqueness for the Dirichlet problem)**.** *The Dirichlet problem — $\Delta u = 0$ in Ω, $u = g$ on $\partial\Omega$ — has at most one solution.*

Proof. If u_1, u_2 both solve it, $w = u_1 - u_2$ is harmonic with $w = 0$ on $\partial\Omega$. By the maximum principle its maximum and minimum, both attained on the boundary, are zero; hence $w \equiv 0$ and $u_1 = u_2$. □

Remark 13.4 (The Neumann problem and a compatibility condition)**.** The Neumann problem — prescribe the normal derivative $\partial u/\partial n = g$ on $\partial\Omega$ rather than the value — behaves differently in two instructive ways. First, the solution is determined only *up to an additive constant*: if u is harmonic with given normal derivative, so is $u + C$, since a constant has zero derivative. (Physically, prescribing the heat flux through the boundary fixes the temperature distribution's shape but not its overall level.) Second, the data cannot be arbitrary: integrating $\Delta u = 0$ over Ω and applying the divergence theorem gives $\oint_{\partial\Omega} \partial u/\partial n\, ds = 0$, so the prescribed flux g must satisfy the *compatibility condition* $\oint g\, ds = 0$. Physically obvious — at a steady state the net heat flux through the boundary must vanish, or the body would keep heating or cooling — it is a first example of a solvability condition, a constraint the data must meet for a solution to exist at all.

13.4 Separation of variables in a rectangle

On a rectangle $0 < x < a$, $0 < y < b$, Laplace's equation separates just as the heat and wave equations did. Seeking $u = X(x)Y(y)$ and substituting into $u_{xx} + u_{yy} = 0$ gives $X''/X = -Y''/Y = -\lambda$, hence

$$X'' + \lambda X = 0, \qquad Y'' - \lambda Y = 0.$$

The crucial difference from the evolution equations: *both* variables are spatial, so both ODEs are of boundary-value type, and the signs differ — one equation oscillates while the other grows and decays. With homogeneous Dirichlet data on the two sides $x = 0$ and $x = a$, the X equation is the familiar eigenvalue problem with $X_n = \sin(n\pi x/a)$, while the Y equation $Y'' = \lambda_n Y$ has *hyperbolic* solutions $\sinh(n\pi y/a)$ and $\cosh(n\pi y/a)$. Superposing,

$$u(x,y) = \sum_{n=1}^{\infty} \sin\frac{n\pi x}{a}\Big(A_n \cosh\tfrac{n\pi y}{a} + B_n \sinh\tfrac{n\pi y}{a}\Big),$$

and the coefficients are fixed by the boundary data on the remaining two sides $y = 0$, $y = b$ through Fourier sine series, exactly as in Chapter 9. The method is identical to separation of variables for the heat equation; only the time factor (decaying exponential) is replaced by the transverse spatial factor (growing/decaying hyperbolic function), reflecting that y here plays a spatial, not temporal, role.

Example 13.5 (A rectangle with one hot side)**.** Solve $\Delta u = 0$ on the square $0 < x, y < \pi$ with $u = 0$ on the three sides $x = 0$, $x = \pi$, $y = 0$, and $u(x, \pi) = \sin x$ on the top. The homogeneous conditions on the vertical sides give $X_n = \sin(nx)$; the condition $u(x, 0) = 0$ on the bottom selects the Y-solution vanishing there, namely $\sinh(ny)$. So $u(x,y) = \sum_n B_n \sinh(ny)\sin(nx)$, and the top condition $\sum_n B_n \sinh(n\pi)\sin(nx) = \sin x$ matches only the $n = 1$ term: $B_1 \sinh\pi = 1$. Hence

$$u(x,y) = \frac{\sinh y}{\sinh\pi}\,\sin x.$$

The solution is largest near the hot top edge and decays (through the factor $\sinh y/\sinh\pi$, which falls from 1 at $y = \pi$ toward 0 at $y = 0$) as one moves into the interior — the steady temperature interpolating smoothly between the hot top and the cold sides.

13.5 Laplace's equation in a disk and Poisson's formula

For a disk the natural coordinates are polar, (r, θ), in which the Laplacian becomes

$$\Delta u = u_{rr} + \frac{1}{r}u_r + \frac{1}{r^2}u_{\theta\theta} = 0.$$

Separation $u = R(r)\Theta(\theta)$ gives an angular equation $\Theta'' + \lambda\Theta = 0$ whose solutions must be *periodic* in θ with period 2π (since θ and $\theta + 2\pi$ are the same point), forcing $\lambda = n^2$ and $\Theta_n = \cos n\theta, \sin n\theta$. The radial equation is then an Euler equation with power solutions r^n and r^{-n}; discarding r^{-n} (singular at the center) leaves $R_n = r^n$. Superposing gives the general harmonic function in the disk,

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

with the coefficients determined by the boundary data $u(R, \theta) = g(\theta)$ on the circle of radius R — again a Fourier series, now of the boundary data in the angle θ. Summing this series in closed form yields the following.

Theorem 13.6 (Poisson's formula)**.** *The solution of the Dirichlet problem in the disk of radius R, with boundary data $g(\theta)$, is*

$$u(r,\theta) = \frac{R^2 - r^2}{2\pi} \int_0^{2\pi} \frac{g(\phi)}{R^2 - 2Rr\cos(\theta - \phi) + r^2}\, d\phi. \tag{13.1}$$

Poisson's formula expresses the value at any interior point as a weighted average of the boundary data, the *Poisson kernel* $\frac{R^2-r^2}{R^2-2Rr\cos(\theta-\phi)+r^2}$ weighting boundary points by their proximity. Setting $r = 0$ recovers the mean value property: at the center, the kernel is constant and u is the plain average of g. The formula is the elliptic analogue of d'Alembert's formula and the heat-kernel convolution — an explicit solution of the boundary-value problem in terms of the data — and it makes manifest both the averaging character of harmonic functions and their smoothness, since the kernel is infinitely differentiable inside the disk.

Example 13.7 (A harmonic function in the disk)**.** Find the harmonic function in the unit disk ($R = 1$) with boundary value $g(\theta) = 2\cos\theta$. Rather than integrate Poisson's formula, match to the series: $g(\theta) = 2\cos\theta$ is already a single Fourier mode ($a_1 = 2$, all others zero), so $u(r,\theta) = 2r\cos\theta$. In Cartesian terms this is $u = 2x$ — a simple linear function, manifestly harmonic, taking the value $2\cos\theta = 2x$ on the unit circle. The example shows that when the boundary data is a finite combination of $\cos n\theta$ and $\sin n\theta$, the solution is read off instantly from the series, no integration required.

13.6 Green's functions, in brief

Poisson's formula is one instance of a general representation: the solution of a boundary-value problem as an integral of the data against a kernel built from the geometry. That kernel is the *Green's function.* For the Dirichlet problem on a region Ω, the Green's function $G(x, y)$ is the response at x to a unit point source at y that vanishes on the boundary — physically, the potential of a point charge inside a grounded conductor shaped like $\partial\Omega$. Once G is known, the solution with boundary data g is an integral of g against the normal derivative of G over the boundary, and the solution with an interior source f is an integral of f against G itself. The Poisson kernel of (13.1) is precisely (the boundary form of) the Green's function of the disk, computable there because the disk's symmetry allows the method of images — the image of an interior point in the circle. Green's functions unify the explicit-solution formulas of the course under one idea, and they are a central topic of the next course in the subject; we note them here to place Poisson's formula in its proper context.

Remark 13.8 (The three explicit solutions)**.** We now have, for each of the three model equations, one geometry in which the solution is completely explicit: d'Alembert's formula for the wave equation on the line, the heat-kernel convolution for diffusion on the line, and Poisson's formula for Laplace's equation in the disk. Each writes the solution as the data acted on by a kernel — two travelling copies, a spreading Gaussian, a boundary average — and the character of the kernel encodes the personality of the equation: finite-speed propagation, diffusive smoothing, equilibrium averaging. These three formulas are the central explicit solutions of the classical theory, and all three are Green's functions for their respective problems.

Example 13.9 (Steady temperature in an annulus)**.** Consider the ring $1 < r < 2$ between two concentric circles, with the inner circle held at $u = 0$ and the outer at $u = 10$, the data being independent of the angle θ. By symmetry the solution depends on r alone, so Laplace's equation reduces to the radial part $u_{rr} + \frac{1}{r}u_r = 0$, i.e. $(ru_r)' = 0$. Integrating twice gives the general radially-symmetric harmonic function $u(r) = A + B\ln r$ — note the logarithm, the radial harmonic function in two dimensions, which we discard inside a full disk (where it blows up at the center) but *keep* in an annulus, which excludes the center. Imposing $u(1) = 0$ and $u(2) = 10$ gives $A + B\ln 1 = 0$, so $A = 0$, and $B\ln 2 = 10$, so $B = 10/\ln 2$. Hence

$$u(r) = \frac{10\ \ln r}{\ln 2}.$$

The temperature rises logarithmically from the cold inner wall to the hot outer one — not linearly, as it would in a slab, because the geometry spreads the heat over circles of growing circumference. This is the simplest problem in which the logarithmic harmonic function plays its proper role, and it shows how the admissible radial solutions depend on whether the domain includes the origin.

Exercises

Exercise 13.1 (*Verifying harmonic functions*). Verify that each is harmonic in the plane: (a) $u = x^2 - y^2$; (b) $u = xy$; (c) $u = e^x \cos y$; (d) $u = \ln(x^2 + y^2)$ away from the origin.

Exercise 13.2 (*Mean value property*). State the mean value property and explain why it expresses the equilibrium (steady-state) condition for heat: why must the temperature at a point equal the average of its surroundings at a steady state?

Exercise 13.3 (*Maximum principle*). A harmonic function on the unit disk equals $3 + \sin\theta$ on the boundary circle. Without solving, what can you say about its maximum and minimum values inside? *(Bound them by the boundary values.)*

Exercise 13.4 (*Uniqueness*). Prove that the Dirichlet problem for Laplace's equation has at most one solution, using the maximum principle. Where is harmonicity of the difference used?

Exercise 13.5 (*Neumann compatibility*). Show that the Neumann problem $\Delta u = 0$ in Ω, $\partial u/\partial n = g$ on $\partial\Omega$, can have a solution only if $\oint_{\partial\Omega} g\, ds = 0$. *(Integrate $\Delta u = 0$ over Ω and use the divergence theorem.)* Why is the solution determined only up to a constant?

Exercise 13.6 (*Rectangle*). Solve $\Delta u = 0$ on the square $0 < x, y < \pi$ with $u = 0$ on the sides $x = 0$, $x = \pi$, $y = 0$, and $u(x, \pi) = \sin x$ on the top (Example 13.5). Where in the square is u largest?

Exercise 13.7 (*Separation in polar coordinates*). Carry out the separation $u = R(r)\Theta(\theta)$ for Laplace's equation in the disk, obtaining the angular equation $\Theta'' + n^2\Theta = 0$ and the radial solutions r^n. Explain why periodicity in θ forces integer n, and why r^{-n} is discarded.

Exercise 13.8 (*Poisson at the center*). Set $r = 0$ in Poisson's formula (13.1) and show it reduces to the mean value property, $u(0) = \frac{1}{2\pi}\int_0^{2\pi} g(\phi)\, d\phi$.

Exercise 13.9 (*Harmonic in the disk*). Find the harmonic function in the unit disk with boundary value $g(\theta) = 2\cos\theta$ (Example 13.7), and identify it as a simple function of x and y.

Exercise 13.10 (*Boundary data with several modes*). Find the harmonic function in the unit disk with boundary value $g(\theta) = 1 + \cos\theta + \sin 2\theta$. *(Match each term to the series; no integration needed.)*

Chapter 14

Numerical Methods

The explicit formulas of the preceding chapters — d'Alembert, the heat kernel, Fourier series, Poisson — solve a handful of equations on a handful of domains. Most PDEs that arise in practice have no such formula: the geometry is irregular, the coefficients vary, the equation is nonlinear. For these we turn to *numerical methods*, which replace the continuous equation by a discrete one that a computer can solve. This chapter introduces the most basic approach, the *finite-difference method*, analyzes its accuracy through truncation error and Big-O notation, and confronts the central subtlety of time-dependent schemes: *stability*, the requirement that errors not explode as the computation marches forward.

14.1 Finite-difference approximations

The idea is to approximate derivatives by difference quotients on a grid. Place grid points $x_j = j\,h$ spaced a distance h apart. Taylor's theorem gives the standard approximations. The *forward difference*

$$u'(x) \approx \frac{u(x+h) - u(x)}{h}$$

has error of order h (we make this precise below). The *centered difference*

$$u'(x) \approx \frac{u(x+h) - u(x-h)}{2h}$$

is more accurate, with error of order h^2. For the second derivative, the *centered second difference*

$$u''(x) \approx \frac{u(x+h) - 2u(x) + u(x-h)}{h^2} \tag{14.1}$$

has error of order h^2 and is the workhorse for the diffusion, wave, and Laplace equations, all of which involve u_{xx}. Replacing the derivatives in a PDE by such differences turns the equation into a system of algebraic relations among the grid values — a system a computer solves.

14.2 Truncation error and Big-O

To quantify accuracy we use *Big-O notation*: we write $E(h) = O(h^p)$ to mean that, for small h, the error $E(h)$ is bounded by a constant times h^p. The exponent p is the *order of accuracy*: a higher p means the error shrinks faster as the grid is refined. Halving h reduces an $O(h)$ error by a factor of

2, an $O(h^2)$ error by a factor of 4 — so second-order methods are dramatically more accurate than first-order ones for fine grids.

The order of a difference formula is found by Taylor expansion. For the forward difference, expand $u(x+h) = u(x) + hu'(x) + \frac{h^2}{2}u''(x) + \cdots$, so

$$\frac{u(x+h) - u(x)}{h} = u'(x) + \frac{h}{2}u''(x) + \cdots = u'(x) + O(h),$$

exhibiting the $O(h)$ *truncation error* — the error made by truncating the Taylor series, i.e. by replacing the derivative with the difference. For the centered second difference (14.1), expanding both $u(x \pm h)$ and adding, the odd-order terms cancel:

$$\frac{u(x+h) - 2u(x) + u(x-h)}{h^2} = u''(x) + \frac{h^2}{12}u''''(x) + \cdots = u''(x) + O(h^2),$$

the cancellation of the odd terms being exactly why centered differences are second-order while one-sided differences are only first-order. The leading neglected term — here $\frac{h^2}{12}u''''$ — both sets the order and warns when the method will struggle: where the higher derivative u'''' is large (a sharp feature), the error is large even at fixed h.

Example 14.1 (Order in practice)**.** Approximate $u''(x)$ for $u(x) = e^x$ at $x = 0$ (true value 1) by the centered second difference (14.1) with $h = 0.1$ and $h = 0.05$. One finds errors of about 8.3×10^{-4} and 2.1×10^{-4} respectively — a roughly fourfold reduction when h is halved, confirming the $O(h^2)$ order. (For a first-order method the reduction would be only twofold.)

Consistency, stability, convergence. It is worth naming the framework these notions belong to. A scheme is *consistent* if its truncation error tends to zero as the grid is refined — the discrete equation genuinely approximates the differential one. It is *stable* if errors are not amplified as the computation proceeds (the subject of Section 14.4). And it is *convergent* if the numerical solution approaches the true solution as the grid is refined. The fundamental theorem of the subject, the *Lax equivalence theorem*, states that for a consistent scheme, *stability is equivalent to convergence*: a consistent scheme converges if and only if it is stable. This is why stability, to which we now turn, is not a technicality but the crux of whether a method works at all.

14.3 An explicit scheme for the diffusion equation

Apply finite differences to the diffusion equation $u_t = ku_{xx}$ on a grid in both space ($x_j = jh$) and time ($t^n = n\,\Delta t$), writing u_j^n for the approximation to $u(x_j, t^n)$. Use a forward difference in time and the centered second difference in space:

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = k\,\frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{h^2}.$$

Solving for the new value u_j^{n+1} gives an *explicit* scheme — the temperature at each grid point at the next time step is computed directly from known values at the current step:

$$u_j^{n+1} = u_j^n + r\big(u_{j+1}^n - 2u_j^n + u_{j-1}^n\big), \qquad r = \frac{k\,\Delta t}{h^2}. \tag{14.2}$$

Starting from the initial data and applying (14.2) repeatedly, we march the solution forward in time. The dimensionless number $r = k\Delta t/h^2$ — the ratio of the time step to the square of the space step, scaled by the diffusivity — turns out to control everything about whether the scheme works.

14.4 Stability and the CFL condition

Here is the subtlety that makes numerical PDEs more than bookkeeping. The explicit scheme (14.2) is *conditionally stable*: it produces sensible results only if the step ratio satisfies

$$r = \frac{k\,\Delta t}{h^2} \le \frac{1}{2}. \tag{14.3}$$

If $r > \frac{1}{2}$, small errors — rounding, or the discretization itself — are *amplified* at each step, growing exponentially until the computed "solution" is swamped by violent, unphysical oscillation. This is numerical *instability*, and it is catastrophic: the scheme does not merely lose accuracy, it produces garbage.

Von Neumann stability analysis. The condition (14.3) is not a guess; it is derived by a beautiful Fourier-based argument due to von Neumann. Test the scheme on a single spatial Fourier mode by substituting $u_j^n = G^n e^{i\beta jh}$ — a wave of spatial frequency β whose amplitude is multiplied by a *growth factor* G at each time step. Inserting into (14.2) and dividing by $e^{i\beta jh}$,

$$G = 1 + r\big(e^{i\beta h} - 2 + e^{-i\beta h}\big) = 1 + 2r(\cos\beta h - 1) = 1 - 4r\sin^2\tfrac{\beta h}{2}.$$

The mode neither grows nor blows up provided $|G| \le 1$ for every frequency β. The most dangerous frequency is the one making $\sin^2(\beta h/2) = 1$, giving $G = 1 - 4r$; the requirement $|1 - 4r| \le 1$ then forces $0 \le r \le \frac{1}{2}$. This is exactly (14.3), now *derived*: stability holds iff every Fourier mode's growth factor stays within the unit disk, and the highest-frequency mode (the sawtooth, which the grid can barely resolve) is the first to go unstable.

The condition has a sharp practical consequence. To halve the spatial grid spacing h for better spatial accuracy, one must *quarter* the time step Δt to keep $r \le \frac{1}{2}$ — so refining the grid is expensive, the number of time steps growing like $1/h^2$. The constraint reflects a genuine feature of diffusion: information at a grid point should not, in one time step, be allowed to depend on data further away than the physics permits. Conditions of this kind, tying the time step to the space step for stability, are called *CFL conditions* (after Courant, Friedrichs, and Lewy); for the wave equation the analogous condition is $c\,\Delta t/h \le 1$, requiring the numerical domain of dependence to contain the true one — a direct echo of the finite propagation speed of Chapter 4.

14.5 Implicit schemes and Crank–Nicolson

The stability restriction can be removed by an *implicit* scheme, which evaluates the spatial difference at the *new* time level rather than the old:

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = k\,\frac{u_{j+1}^{n+1} - 2u_j^{n+1} + u_{j-1}^{n+1}}{h^2}.$$

The update now couples all the new grid values together — each equation involves three unknowns at the new level — so advancing one step requires solving a linear system (a tridiagonal one, efficiently solvable). The reward is decisive: a von Neumann analysis gives growth factor $G = 1/(1 + 4r\sin^2(\beta h/2))$, which satisfies $|G| \le 1$ for *every* $r > 0$. The implicit scheme is therefore *unconditionally stable*: any time step gives a bounded result, and Δt may be chosen for accuracy alone, free of the CFL straitjacket.

A particularly good compromise is the *Crank–Nicolson* scheme, which averages the explicit and implicit spatial differences (evaluating half at the old level and half at the new). It is unconditionally stable *and* second order in time — $O(\Delta t^2)$ rather than the $O(\Delta t)$ of the plain explicit and implicit schemes — so it attains good accuracy without a tiny time step. The trade-off it embodies is characteristic of numerical analysis: explicit methods are cheap per step but constrained in step size; implicit methods cost a linear solve per step but lift the constraint; and the best practical schemes, like Crank–Nicolson, balance the two.

14.6 A scheme for the wave equation

The wave equation, second order in time, is discretized with a centered second difference in time as well as space:

$$\frac{u_j^{n+1} - 2u_j^n + u_j^{n-1}}{\Delta t^2} = c^2\,\frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{h^2},$$

an explicit scheme that computes the new level u_j^{n+1} from the two previous levels u^n and u^{n-1} (two are needed because the equation is second order in time, mirroring the two initial conditions). A von Neumann analysis yields the stability condition

$$\frac{c\,\Delta t}{h} \le 1,$$

the *Courant number* not exceeding one. Its meaning is geometric and ties back to Chapter 4: in one time step a signal must not be required to travel more than one grid spacing, so the numerical domain of dependence (the grid points feeding u_j^{n+1}) must contain the true domain of dependence (the characteristics through the point). When it does not, the scheme cannot "see" the data that physically determines the solution, and it goes unstable — a direct numerical manifestation of finite propagation speed.

Remark 14.2 (Three sources of error)**.** A numerical solution differs from the true solution for three distinct reasons, worth separating. *Truncation error* comes from replacing derivatives by differences, and shrinks as $O(h^p)$ when the grid is refined. *Stability* concerns whether errors already present are amplified or controlled as the computation proceeds — a refined grid that violates the CFL condition is worse, not better. *Round-off error* comes from finite-precision arithmetic and sets a floor below which refinement cannot improve accuracy. A reliable computation requires attention to all three: a consistent discretization (small truncation error) that is stable (errors not amplified), run in adequate precision. This triad — consistency, stability, convergence — is the organizing principle of the numerical analysis of PDEs.

Exercises

Exercise 14.1 (*Difference formulas*). Write the forward, backward, and centered difference approximations to $u'(x)$, and the centered approximation to $u''(x)$. State the order of accuracy of each.

Exercise 14.2 (*Deriving the order*). Use Taylor expansion to show that the forward difference $(u(x+h) - u(x))/h$ approximates $u'(x)$ with error $O(h)$, and that the centered difference $(u(x+h) - u(x-h))/(2h)$ has error $O(h^2)$. Which terms cancel in the centered case?

Exercise 14.3 (*Second difference*). Show by Taylor expansion that the centered second difference (14.1) approximates $u''(x)$ with error $O(h^2)$, and identify the leading error term.

Exercise 14.4 (*Order in practice*). For $u(x) = \sin x$ at $x = 1$, where $u''(1) = -\sin 1$, compute the centered second-difference approximation for $h = 0.1$ and $h = 0.05$, and verify the error drops by about a factor of 4.

Exercise 14.5 (*The explicit scheme*). Write out the explicit finite-difference scheme (14.2) for $u_t = u_{xx}$ (so $k = 1$) and identify the step ratio r. If $h = 0.1$, what is the largest stable time step Δt?

Exercise 14.6 (*Von Neumann analysis*). Carry out the von Neumann stability analysis of the explicit heat scheme: substitute $u_j^n = G^n e^{i\beta jh}$, find the growth factor $G = 1 - 4r\sin^2(\beta h/2)$, and deduce the condition $r \leq \frac{1}{2}$ from $|G| \leq 1$. Which frequency is the most unstable?

Exercise 14.7 (*Stability restriction*). Explain why halving h forces Δt to be quartered to keep $r \leq \frac{1}{2}$, and what this implies about computational cost as the grid is refined.

Exercise 14.8 (*Implicit stability*). Show that the implicit heat scheme has growth factor $G = 1/(1 + 4r\sin^2(\beta h/2))$, and conclude it is unconditionally stable. What is the cost, per time step, of this stability?

Exercise 14.9 (*The CFL condition for waves*). State the CFL condition $c\Delta t/h \leq 1$ for the explicit wave scheme, and explain its meaning in terms of the domain of dependence (Chapter 4). Why must the numerical domain of dependence contain the true one?

Exercise 14.10 (*Explicit, implicit, Crank–Nicolson*). Contrast the explicit, implicit, and Crank–Nicolson schemes for the diffusion equation in terms of cost per step, stability, and order of accuracy in time. When is Crank–Nicolson preferable?

Exercise 14.11 (*Three errors*). Name the three distinct sources of error in a numerical solution of a PDE, and state how each responds (or fails to respond) to refining the grid spacing h.

Appendix A

Reference Tables and Formulas

This appendix collects the formulas used repeatedly in the text, for convenience during problem-solving. Nothing here is new; it is gathered in one place so the reader need not hunt through chapters for a transform pair or an orthogonality relation.

A.1 Trigonometric and orthogonality identities

The product-to-sum identities, used to establish orthogonality:

$$\begin{aligned}
\sin A \sin B &= \tfrac{1}{2}[\cos(A-B) - \cos(A+B)],\\
\cos A \cos B &= \tfrac{1}{2}[\cos(A-B) + \cos(A+B)],\\
\sin A \cos B &= \tfrac{1}{2}[\sin(A+B) + \sin(A-B)].
\end{aligned}$$

The orthogonality relations on $0 < x < l$ (with m, n positive integers):

$$\begin{aligned}
&\int_0^l \sin\tfrac{m\pi x}{l} \sin\tfrac{n\pi x}{l}\, dx = \tfrac{l}{2}\,\delta_{mn},\\
&\int_0^l \cos\tfrac{m\pi x}{l} \cos\tfrac{n\pi x}{l}\, dx = \tfrac{l}{2}\,\delta_{mn} \quad (m, n \geq 1),\\
&\int_0^l \sin\tfrac{m\pi x}{l} \cos\tfrac{n\pi x}{l}\, dx = 0 \;\; \text{in general not zero; see text for the full-interval case.}
\end{aligned}$$

On the symmetric interval $-l < x < l$ the sines and cosines together are orthogonal, which is what makes the full Fourier series work. The Gaussian integral, used throughout the diffusion chapters:

$$\int_{-\infty}^{\infty} e^{-ax^2}\, dx = \sqrt{\tfrac{\pi}{a}} \quad (a > 0).$$

A.2 Fourier series

For a function ϕ on $0 < x < l$, the three classical series and their coefficients:

$$\begin{array}{lll} \text{sine series} & \phi = \sum_{n=1}^{\infty} b_n \sin \frac{n\pi x}{l}, & b_n = \frac{2}{l}\int_0^l \phi \sin \frac{n\pi x}{l}\, dx; \\ \text{cosine series} & \phi = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos \frac{n\pi x}{l}, & a_n = \frac{2}{l}\int_0^l \phi \cos \frac{n\pi x}{l}\, dx; \\ \text{complex (on } -l < x < l) & \phi = \sum_{n=-\infty}^{\infty} c_n e^{in\pi x/l}, & c_n = \frac{1}{2l}\int_{-l}^l \phi\, e^{-in\pi x/l}\, dx. \end{array}$$

Parseval's identity (sine case): $\int_0^l \phi^2\, dx = \frac{l}{2}\sum_{n\geq 1} b_n^2$.

A.3 Fourier transform

Definition and inversion (frequency variable ξ):

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x)e^{-i\xi x}\, dx, \qquad f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}(\xi)e^{i\xi x}\, d\xi.$$

Properties:

$f(x)$	$\hat{f}(\xi)$
$f'(x)$	$i\xi\, \hat{f}(\xi)$
$f''(x)$	$-\xi^2\, \hat{f}(\xi)$
$f(x-a)$	$e^{-ia\xi}\hat{f}(\xi)$
$e^{ibx}f(x)$	$\hat{f}(\xi - b)$
$f(ax)$	$\frac{1}{\lvert a\rvert}\hat{f}(\xi/a)$
$(f * g)(x)$	$\hat{f}(\xi)\, \hat{g}(\xi)$
e^{-ax^2}	$\sqrt{\pi/a}\; e^{-\xi^2/4a}$

Plancherel: $\int \lvert f\rvert^2\, dx = \frac{1}{2\pi}\int \lvert \hat{f}\rvert^2\, d\xi$.
Heat kernel: $S(x,t) = (4\pi kt)^{-1/2}e^{-x^2/4kt}$, with $\hat{S}(\xi,t) = e^{-k\xi^2 t}$.

A.4 Laplace transform

Definition: $F(s) = \mathcal{L}\{f\}(s) = \displaystyle\int_0^{\infty} f(t)e^{-st}\, dt$. Pairs and properties:

$f(t)$	$F(s)$
1	$1/s$
t^n	$n!/s^{n+1}$
e^{at}	$1/(s-a)$
$\cos\omega t$	$s/(s^2+\omega^2)$
$\sin\omega t$	$\omega/(s^2+\omega^2)$
$\delta(t)$	1
$f'(t)$	$sF(s)-f(0)$
$f''(t)$	$s^2F(s)-sf(0)-f'(0)$
$e^{at}f(t)$	$F(s-a)$
$H(t-a)f(t-a)$	$e^{-as}F(s)$
$(f*g)(t)$	$F(s)\,G(s)$

Here H is the Heaviside step and δ the Dirac delta. A useful PDE entry: the inverse transform of $\frac{1}{s}e^{-\sqrt{s/k}\,x}$ is $\operatorname{erfc}(x/\sqrt{4kt})$, the suddenly-heated-rod solution.

A.5 The three model equations at a glance

	Wave	Diffusion	Laplace
equation	$u_{tt}=c^2u_{xx}$	$u_t=ku_{xx}$	$u_{xx}+u_{yy}=0$
type	hyperbolic	parabolic	elliptic
data	u, u_t at $t=0$	u at $t=0$	u on $\partial\Omega$
explicit solution	d'Alembert	heat kernel	Poisson
key tool	energy	max. principle	max. principle
speed	finite (c)	infinite	— (steady)
smoothing	none	instant	(smooth)

Appendix B

A Guide to Choosing a Method

A recurring difficulty for students is not solving a given problem but knowing *which* method to reach for. This short appendix is a decision guide, collecting in one place the logic scattered through the text. It is meant to be read after the methods themselves are understood; it organizes them, but does not replace them.

B.1 First, classify

The single most useful first step with any new problem is to identify the *type* of the equation (Section 1.4), because type dictates both the behavior and the appropriate data.

- *First order* (e.g. $u_t + cu_x = 0$): method of characteristics (Chapter 2). Reduce to ODEs along characteristic curves.
- *Hyperbolic* ($u_{tt} = c^2 u_{xx}$): on the line, d'Alembert's formula (Chapter 4); on an interval, separation of variables into standing modes (Chapter 8).
- *Parabolic* ($u_t = ku_{xx}$): on the line, the heat kernel (Chapter 6); on an interval, separation of variables into decaying modes (Chapter 8).
- *Elliptic* ($\Delta u = 0$): separation of variables in a rectangle or disk, or Poisson's formula in a disk (Chapter 13).

B.2 Then, match the method to the domain

Within a type, the *domain* selects the technique. The governing question is whether the spatial domain is bounded.

- *Bounded interval* $0 < x < l$, homogeneous boundary conditions: **separation of variables**, expanding the data in the eigenfunctions the boundary conditions select (sine for Dirichlet, cosine for Neumann, full series for periodic). This is the default for finite domains.
- *Whole line* $-\infty < x < \infty$: the **Fourier transform** (Chapter 11), or equivalently the fundamental-solution convolution (heat kernel, d'Alembert). Natural when the data decays at infinity.
- *Half-line* $0 < x < \infty$ with one boundary: the **method of reflection / images** (Chapter 7), reducing to a whole-line problem with odd or even data.

- *Semi-infinite, with a time-dependent boundary*: the **Laplace transform** in time (Chapter 12), which builds in the initial condition and leaves an ODE in x.

- *Initial-value problem in time*, $t \geq 0$: the **Laplace transform**, especially for ODEs and for forced problems with switches (Heaviside) or impulses (delta).

B.3 Handling complications

Two features call for a preliminary step before any of the above.

- *Inhomogeneous boundary conditions*: subtract a particular function (usually the *steady state*) carrying the boundary values, reducing to homogeneous conditions for the remainder (Sections 8.6, 9.8).

- *A source term* f: use **Duhamel's principle** (Chapter 7) — evolve each instant's forcing by the homogeneous solution and integrate — or, on a bounded interval, expand both f and u in the eigenfunctions and solve mode by mode.

B.4 When no formula exists

If the geometry is irregular, the coefficients vary, or the equation is nonlinear, no closed-form method applies and one turns to **numerical methods** (Chapter 14): discretize by finite differences, check that the scheme is consistent (small truncation error) and stable (respect the CFL condition), and march forward. The choice between an explicit scheme (cheap per step, step-size restricted) and an implicit one (a linear solve per step, unconditionally stable) is itself a decision the chapter equips you to make.

B.5 A worked decision

Suppose the problem is: a rod $0 < x < \pi$, insulated at both ends, with a given initial temperature, find $u(x,t)$. The logic: the equation $u_t = ku_{xx}$ is *parabolic*; the domain is a *bounded interval*; the boundary conditions ($u_x = 0$ at both ends) are *homogeneous Neumann*. Therefore: separation of variables, expanding the initial data in the *cosine* series the Neumann conditions select, with each mode decaying as $e^{-(n\pi/l)^2kt}$ — and the constant mode surviving, so the rod relaxes to its average temperature. Every step of this reasoning is a lookup in the lists above, and that is exactly how the guide is meant to be used.

Appendix C

Solution Sketches

These sketches indicate the intended approach and answer for each exercise; a complete write-up should expand the reasoning and show all steps. Discursive exercises are answered with the key points an adequate response should contain.

Chapter 1

1.1. (a) 2nd order, linear, homogeneous. (b) 1st order, nonlinear. (c) 2nd order, linear, inhomogeneous (source $\sin t$). (d) 1st order, nonlinear. (e) 2nd order, linear, homogeneous.

1.2. (a) quasilinear (highest term u_{xx} is linear; uu_x is lower-order nonlinear); (b) semilinear (linear in u_{xx}, nonlinear in u only through u^3); (c) fully nonlinear ($u_x u_y$); (d) quasilinear (the coefficient $1 + u_x^2$ multiplies the highest derivative u_{xx}).

1.3. (a) $u_t = -k\xi^2 e^{-k\xi^2 t} \sin \xi x = k u_{xx}$. (b) $u_{tt} = c^2(f'' + g'') = c^2 u_{xx}$. (c) $u_{xx} = 2$, $u_{yy} = -2$, sum 0.

1.4. Two constants both solve $u_t + uu_x = 0$. A combination taking different values where it varies has $uu_x \neq$ the sum of separate terms; the equation is nonlinear, so $\mathcal{L}(u_1 + u_2) \neq \mathcal{L}u_1 + \mathcal{L}u_2$. Superposition needs linearity.

1.5. $u = \frac{1}{2}x^2$ gives $u_{xx} + u_{yy} = 1 + 0 = 1$. By the affine structure, the general solution is $\frac{1}{2}x^2 + h(x, y)$ with h any harmonic function.

1.6. $u_{xy} = 0 \Rightarrow u = f(x) + g(y)$; $u_{xx} = 0 \Rightarrow u = f(y) + xg(y)$. Two arbitrary functions each.

1.7. (a) elliptic; (b) hyperbolic; (c) parabolic; (d) $a = 3, b = 1, c = 1$, $b^2 - ac = -2 < 0$, elliptic; (e) $a = 1, b = 2, c = 1$, $b^2 - ac = 3 > 0$, hyperbolic.

1.8. Both governed by the sign of $b^2 - ac$, the (negative) determinant of the coefficient matrix $\left(\begin{smallmatrix} a & b \\ b & c \end{smallmatrix}\right)$; it decides the signature of the quadratic form, hence the conic's shape and the number of real characteristics.

1.9. Hyperbolic: two real families (e.g. $x \pm ct$ for the wave equation); parabolic: one ($t = \text{const}$ for heat); elliptic: none. The count equals the number of real roots of the characteristic quadratic, governed by $\operatorname{sign}(b^2 - ac)$.

1.10. $a = y$, $b = 0$, $c = 1$, so $b^2 - ac = -y$: elliptic $y > 0$, parabolic $y = 0$, hyperbolic $y < 0$.

1.11. Setting $u_t = 0$ gives $u_{xx} + u_{yy} = 0$. It describes the steady-state (equilibrium) temperature distribution.

Chapter 2

2.1. Characteristics $x - 3t = \text{const}$; $u = 1/(1 + (x - 3t)^2)$. Peak at $x = 3t$; at $t = 4$, $x = 12$.

2.2. $u = \cos(x + 2t)$, moving left at speed 2.

2.3. Characteristic direction $(2, 3)$, so characteristics $3x - 2t = \text{const}$ and $u(x, t) = \phi(3x - 2t)$ (equivalently ϕ of $x - \frac{2}{3}t$).

2.4. $u_t = -c\phi'$, $u_x = \phi'$, so $u_t + cu_x = 0$.

2.5. $X' = 2X \Rightarrow X = x_0 e^{2t}$, $x_0 = xe^{-2t}$, so $u = \phi(xe^{-2t})$.

2.6. $X(t) = x_0 + \frac{1}{2}t^2$, so $u = \phi(x - \frac{1}{2}t^2)$; characteristics are congruent parabolas.

2.7. On $x = ct + x_0$: $\frac{d}{dt}u = u$, so $u = e^t\phi(x - ct)$.

2.8. Characteristics $X = x_0 + t$; $\frac{d}{dt}u = x_0 + t$ gives $u = x_0 t + \frac{1}{2}t^2 = xt - \frac{1}{2}t^2$ (with $x_0 = x - t$).

2.9. Characteristics $x = x_0 e^t$ diverge as t grows, so a bump stretches out.

2.10. $\phi = -x$, $\phi' = -1$, $t^\star = -1/(-1) = 1$. Characteristics cross at $t = 1$: the solution becomes multivalued and a shock forms.

2.11. The PDE fixes the derivative of u along the characteristic; integrating that derivative along each characteristic (an ODE) recovers u, the data fixing the constant.

Chapter 3

3.1. (a) one (first order in t); (b) two (second order in t); (c) two (the u_{xxxx} is spatial; the equation is 2nd order in t).

3.2. (a) Dirichlet; (b) Neumann; (c) Robin; (d) periodic.

3.3. (a) end held at 0°; (b) insulated end; (c) end radiating heat to a 0° environment proportional to its temperature.

3.4. $X = A\cos\beta x + B\sin\beta x$; $X(0) = 0 \Rightarrow A = 0$; $X'(l) = 0 \Rightarrow \cos\beta l = 0 \Rightarrow \beta = (2n - 1)\pi/2l$. So $X_n = \sin\frac{(2n-1)\pi x}{2l}$, $\lambda_n = ((2n - 1)\pi/2l)^2$.

3.5. (a) initial condition plus one boundary condition at each end for $t > 0$; (b) boundary data on the entire boundary circle; (c) two initial conditions, no boundary conditions.

3.6. Two initial conditions (u, u_t at $t = 0$) and one boundary condition at each end. One fewer: solution not determined (uniqueness fails). One more (e.g. data at $t = T$): generally over-determined, no solution (existence fails).

3.7. (a) straight line from T_1 to T_2; (b) the constant equal to the average initial temperature (insulated ends conserve $\int u\,dx$).

3.8. Equation alone leaves the flux at the ends free, so many solutions; one boundary condition per end removes the freedom.

3.9. Existence (model not over-determined), uniqueness (enough data), stability (measurement error does not ruin prediction).

3.10. No — stability fails. A solution swinging wildly under tiny data changes cannot be computed from noisy data.

3.11. $u_{xx} = -\sin(nx)\sinh(ny)$, $u_{yy} = \sin(nx)\sinh(ny)$, sum 0; $u(x, 0) = 0$, $u_y(x, 0) = \frac{1}{n}\sin(nx)$. Data $\to 0$ but $\frac{1}{n^2}\sinh(ny)$ blows up for $y > 0$. *Stability* fails; elliptic Cauchy problems are ill-posed.

Chapter 4

4.1. $\partial_t = c(\partial_\xi - \partial_\eta)$, $\partial_x = \partial_\xi + \partial_\eta$, giving $u_{tt} - c^2u_{xx} = -4c^2u_{\xi\eta}$, so $u_{\xi\eta} = 0$ and $u = F(\xi) + G(\eta)$.

4.2. First: $u = \frac{1}{2}[\sin(x + t) + \sin(x - t)] = \sin x\cos t$. Second: $u = \frac{1}{2}\int_{x-t}^{x+t}\cos s\,ds = \cos x\sin t$.

4.3. $u = \frac{1}{2}[e^{-(x+t)^2} + e^{-(x-t)^2}] + \frac{1}{2}\int_{x-t}^{x+t} 1\,ds = \frac{1}{2}[e^{-(x+t)^2} + e^{-(x-t)^2}] + t$. The constant initial velocity contributes a term t growing linearly in time.

4.4. $u = \frac{1}{4}\int_{x-2t}^{x+2t}\psi$, equal to $\frac{1}{4}$ times the overlap length of $[x-2t, x+2t]$ with $(-1,1)$; for large t a plateau of height $\frac{1}{2}$ between the fronts.

4.5. The plucked string's displacement is $\frac{1}{2}[\phi(x+ct)+\phi(x-ct)]$, which returns to 0 behind each pulse; the struck string's is $\frac{1}{2c}\int\psi$, which accumulates the (one-signed) velocity and stays displaced.

4.6. Depends on ϕ at $x_0 \pm ct_0$, ψ on $[x_0-ct_0, x_0+ct_0]$. Changing data on $|x|<1$ affects (x_0,t_0) iff that interval meets $[x_0-ct_0, x_0+ct_0]$.

4.7. With $u = F(x+ct)+G(x-ct)$, opposite vertices of a characteristic parallelogram share the same $x+ct$ (one pair) and $x-ct$ (other pair) values, so $u(A)+u(C)=u(B)+u(D)$.

4.8. Data supported in $|x|<a$: $\phi(x_0 \pm ct) = 0$ and the velocity integral is empty until $[x_0-ct, x_0+ct]$ reaches $|x|<a$, at $t = (|x_0|-a)/c$. Diffusion spreads instantly.

4.9. $\frac{dE}{dt} = \int(u_t u_{tt} + c^2 u_x u_{xt})$; parts on the second term (boundary terms vanish by decay) gives $\int u_t(u_{tt}-c^2u_{xx}) = 0$ by the equation.

4.10. On $0<x<l$ with $u=0$ at the ends, $u_t=0$ there, so the boundary term $[c^2u_xu_t]_0^l = 0$; energy conserved; the difference has $E(0)=0$, so $E\equiv 0$ and the difference is $\equiv 0$.

4.11. $\frac{dE}{dt} = -a\int u_t^2 \le 0$; the vibration loses energy (damps); uniqueness follows since the difference has $E(0)=0$ and E non-increasing.

Chapter 5

5.1. Conservation of heat gives $\frac{d}{dt}\int_a^b u = q(a)-q(b)$; Fourier's law $q=-Ku_x$ supplies the flux. For a chemical, Fick's law replaces Fourier's.

5.2. At an interior max in x, $u_{xx}\le 0$, so $u_t = ku_{xx}\le 0$: the hot spot cools.

5.3. At most 1: the max over the rectangle is on the parabolic boundary (initial ≤ 1, ends $=0$), so $u\le 1$.

5.4. Max ≤ 10 and min ≥ 0 on the parabolic boundary (initial in $[0,10]$, ends 0), so $0\le u\le 10$ everywhere. The bounds are attained initially (the max) and at the ends (the min 0).

5.5. Comparison principle: if $u\le v$ on the parabolic boundary then $u\le v$ everywhere (apply the maximum principle to $u-v$). A rod starting and held cooler stays cooler.

5.6. $\frac{dE}{dt} = -k\int_0^l u_x^2 \le 0$: dissipative. E increases only if $u_x\equiv 0$ (spatially constant), where it is constant.

5.7. Energy: difference w has $E(0)=0$, $E\ge 0$ non-increasing, so $w\equiv 0$. Maximum principle: $w=0$ on the parabolic boundary, so $w\equiv 0$. The maximum-principle proof is shorter.

5.8. Boundary term $k[ww_x]_0^l = k[-\beta w(l)^2 - \alpha w(0)^2] \le 0$; so $\frac{dE}{dt} = -k\int w_x^2 - k\beta w(l)^2 - k\alpha w(0)^2 \le 0$. Signs $\alpha,\beta\ge 0$ make these terms non-positive.

5.9. $\frac{dE}{dt} = -k\int w_x^2 - \int c\,w^2 \le 0$ when $c\ge 0$. If $c<0$ somewhere, $-\int cw^2$ can be positive and E may grow.

5.10. $u_t = kn^2e^{kn^2t}\sin nx = -ku_{xx}$, data $\sin nx$ bounded, solution grows like e^{kn^2t}, unbounded as $n\to\infty$. Stability fails; reflects irreversibility.

Chapter 6

6.1. $S_t = S(-\frac{1}{2t}+\frac{x^2}{4kt^2})$, $S_{xx} = S(-\frac{1}{2kt}+\frac{x^2}{4k^2t^2})$; then $kS_{xx}=S_t$.

6.2. $\int S\,dx = (4\pi kt)^{-1/2}\sqrt{4\pi kt} = 1$; constant total heat = conservation.

6.3. Chain rule: $u(\lambda x, \lambda^2 t)$ solves the equation. $x/\sqrt{t}$ is invariant, motivating a solution of that variable, leading to the Gaussian.

6.4. The substitution gives $\frac{1}{2}(1+\operatorname{erf}(x/\sqrt{4kt}))$; transition width $\sim\sqrt{4kt}$.

6.5. erf is odd since the integrand e^{-s^2} is even; $\mathrm{erf}(0)=0$; and $\mathrm{erf}(\infty)=\frac{2}{\sqrt{\pi}}\cdot\frac{\sqrt{\pi}}{2}=1$, which is why $2/\sqrt{\pi}$ is the right constant.

6.6. Convolving e^{-x^2} with the Gaussian kernel yields a Gaussian of variance the sum of the variances, hence widening with t.

6.7. x-derivatives fall on the smooth kernel S, infinitely differentiable for $t>0$, so u is too, regardless of ϕ.

6.8. $S(x-y,t)>0$ for all $x,y,t>0$, so $\int S\phi>0$ everywhere when $\phi\geq 0$ is positive somewhere. The wave equation's domain of influence is bounded.

6.9. A sum of many independent small steps is approximately Gaussian (central limit theorem); the standard deviation grows like $\sqrt{n}\propto\sqrt{t}$, giving the $\sqrt{t}$ spread and matching the heat kernel.

6.10. Depth $\sim\sqrt{kt}$; yearly t is $\sim 365\times$ daily, so depth larger by $\sqrt{365}\approx 19$.

Chapter 7

7.1. $u=\int_0^\infty[S(x-y,t)-S(x+y,t)]\phi(y)\,dy$; odd in x, so $u(0,t)=0$.

7.2. Same with a plus sign; even in x, so $u_x(0,t)=0$.

7.3. Real source at a, image *sink* at $-a$: $u=S(x-a,t)-S(x+a,t)$, zero at $x=0$. An insulated wall uses an image source of the same sign (a plus).

7.4. The odd extension sends an inverted image toward $x=0$; the reflected pulse returns *inverted.*

7.5. Even extension; reflected pulse returns *upright.*

7.6. Homogeneous solution starting at s is $e^{a(t-s)}f(s)$; Duhamel gives $u=\int_0^t e^{a(t-s)}f(s)\,ds$, the integrating-factor result.

7.7. For $u_{tt}-c^2u_{xx}=f$ with zero data, forcing enters as an initial velocity (the equation is 2nd order in t); the d'Alembert velocity term $\frac{1}{2c}\int_{x-c(t-s)}^{x+c(t-s)}f(y,s)\,dy$ accumulated over s gives the characteristic-triangle formula.

7.8. Differentiating $u=\int_0^t U\,ds$: upper-limit term $U(x,t;t)=f$; integral of $U_t=kU_{xx}$ gives ku_{xx}; so $u_t=f+ku_{xx}$, $u(x,0)=0$.

7.9. Split $u=v+w$: v solves the equation with the source and zero initial data, w the homogeneous equation with the initial data; superposition gives u.

7.10. Steady state solves $-ku_{xx}=f(x)$ with $u(0)=u(l)=0$: source heating balanced by conduction out the ends.

Chapter 8

8.1. $XT'=kX''T\Rightarrow T'/(kT)=X''/X$; left depends on t, right on x, so both equal a constant, giving the two ODEs.

8.2. $\lambda\leq 0$ forces $X\equiv 0$; $\lambda=\beta^2>0$ with $X(0)=0$ gives $X=B\sin\beta x$, and $X(l)=0$ gives $\beta=n\pi/l$. Hence $\lambda_n=(n\pi/l)^2$, $X_n=\sin(n\pi x/l)$.

8.3. $X'(0)=0$ kills the sine; $X'(l)=0$ gives $\beta=n\pi/l$, $X_n=\cos(n\pi x/l)$; $n=0$ gives the constant $X_0=1$.

8.4. Only $n=3$: $u=5e^{-9t}\sin 3x$; at $t=1$, $u=5e^{-9}\sin 3x\approx 6.17\times 10^{-4}\sin 3x$.

8.5. $u=2e^{-t}\sin x-e^{-16t}\sin 4x$; the first mode dominates for large t (rate 1 vs 16).

8.6. $b_n=\frac{2}{\pi}\int_0^\pi x(\pi-x)\sin nx\,dx=\frac{4}{\pi n^3}(1-(-1)^n)$, i.e. $8/(\pi n^3)$ for odd n. So $u=\frac{8}{\pi}\sum_{\text{odd}}n^{-3}e^{-n^2t}\sin nx$; at $t=1$ the $n=3$ term is $\sim e^{-9}$ of the first, so essentially $\frac{8}{\pi}e^{-t}\sin x$.

8.7. $u = 3 + e^{-4t}\cos 2x$; relaxes to 3 (average initial temperature), not zero, because the constant mode does not decay (insulated ends conserve heat).

8.8. $v(x) = 2 + 3x$ is the steady state, and here $u(x,0) = 2 + 3x = v(x)$, so the transient $w = u - v$ has zero initial data and zero boundary data, giving $w \equiv 0$ and $u(x,t) = 2 + 3x$ for all t (already steady).

8.9. $u = \cos(\pi ct/l)\sin(\pi x/l)$, frequency $\pi c/l$ (fundamental).

8.10. Expand the right- and left-moving sines and add to get $2\sin(n\pi x/l)\cos(n\pi ct/l)$; a standing wave is two opposite travelling waves.

8.11. The nth mode decays like $e^{-(n\pi/l)^2kt}$; larger n faster. High modes carry the rough (rapidly varying) features of ϕ, which vanish first — smoothing.

Chapter 9

9.1. $\sin\frac{m\pi x}{l}\sin\frac{n\pi x}{l} = \frac{1}{2}[\cos\frac{(m-n)\pi x}{l} - \cos\frac{(m+n)\pi x}{l}]$; both integrate to 0 over $(0,l)$ for $m \neq n$; for $m = n$, $\int_0^l \sin^2 = l/2$.

9.2. $b_n = \frac{2l}{n\pi}(-1)^{n+1}$, decaying like $1/n$, because the odd periodic extension of x has jumps at $\pm l$.

9.3. $a_0/2 = l/2$, $a_n = \frac{2l}{n^2\pi^2}((-1)^n - 1)$, decaying like $1/n^2$; the even extension is continuous (triangle wave), hence faster convergence at the endpoints.

9.4. $b_n = 4/(n\pi)$ (odd n). At $x = l/2$, $\sin(n\pi/2) = \pm 1$ for odd n, giving $1 = \frac{4}{\pi}(1 - \frac{1}{3} + \frac{1}{5} - \cdots)$, i.e. $\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \cdots$.

9.5. x is odd, so only sine terms: $c_n = -ib_n/2 = i\frac{l}{n\pi}(-1)^n$ for $n \neq 0$ (and $c_0 = 0$), where $b_n = \frac{2l}{n\pi}(-1)^{n+1}$ from Exercise 9.2.

9.6. At $x = 0$ the odd extension jumps from -1 to 1; Dirichlet's theorem gives the midpoint 0, which the sine series indeed takes.

9.7. Differentiating multiplies the nth coefficient by $n\pi/l$; since $b_n \sim 1/n$, the differentiated coefficients do not tend to zero and the series diverges. The decay-to-zero of coefficients fails.

9.8. Pointwise (to the midpoint at a jump; needs piecewise smoothness); uniform (everywhere at once; needs continuity); mean-square ($\int |\text{error}|^2 \to 0$; needs only square-integrability) — the least.

9.9. $\int_0^l 1\,dx = l = \frac{l}{2}\sum b_n^2 = \frac{l}{2}\cdot\frac{16}{\pi^2}\sum_{\text{odd}}\frac{1}{n^2}$, giving $\sum_{\text{odd}}\frac{1}{n^2} = \frac{\pi^2}{8}$.

9.10. The overshoot stays $\approx 9\%$ of the jump; more terms only narrow its location. Uniform convergence fails across the jump; pointwise and mean-square still hold.

9.11. $v(x) = T_1 + (T_2 - T_1)x/l$; $u - v$ solves the heat equation with homogeneous Dirichlet data and initial data $\phi - v$.

Chapter 10

10.1. $p = 1$, $q = 0$, $w = 1$; regular on a finite interval.

10.2. Multiply by e^{2x}: $(e^{2x}X')' + \lambda e^{2x}X = 0$, so $p = e^{2x}$, $q = 0$, $w = e^{2x}$.

10.3. $\sin(n\pi x/l)$ vanishes at $x = kl/n$, $k = 1,\ldots,n-1$: exactly $n-1$ interior zeros.

10.4. $\int_a^b X_mX_nx\,dx = 0$ $(m \neq n)$; $c_n = \int \phi X_nx\,dx / \int X_n^2x\,dx$.

10.5. Two integrations by parts of $\int(X_m(-(pX_n')') - X_n(-(pX_m')'))$ produce $[p(X_nX_m' - X_mX_n')]_a^b$; the qX terms cancel.

10.6. Dirichlet ($X = 0$) or Neumann ($X' = 0$) at both ends makes the boundary term vanish.

10.7. Rayleigh quotient $R = \int_0^l (X')^2\,dx / \int_0^l X^2\,dx$. With $X = x(l-x)$: $\int (X')^2 = l^3/3$ and $\int X^2 = l^5/30$, so $R = 10/l^2$. This bounds $\lambda_1 = \pi^2/l^2 \approx 9.87/l^2$ from above; the estimate is about 1.3% high.

10.8. With X_n and $\overline{X_n}$, Lagrange's identity gives $(\lambda_n - \overline{\lambda_n}) \int |X_n|^2 w = 0$; since the integral is positive, λ_n is real.

10.9. Eigenvalues satisfy $\sqrt{\lambda}\tan(\sqrt{\lambda}\,l) = \beta$ (or equivalent). The S–L theorem guarantees real discrete eigenvalues with orthogonal complete eigenfunctions regardless of closed-form solvability.

10.10. Weight $w(r) = r$; orthogonality $\int_0^a J_n(\sqrt{\lambda_j}r)J_n(\sqrt{\lambda_k}r)\,r\,dr = 0$ for $j \neq k$. The weight r comes from the factor $p = w = r$ in the polar Laplacian (the area element $r\,dr$).

Chapter 11

11.1. $\hat{f}(\xi) = \frac{1}{a-i\xi} + \frac{1}{a+i\xi} = \frac{2a}{a^2+\xi^2}$ — a Lorentzian, not a Gaussian.

11.2. Parts on $\int f' e^{-i\xi x}$ moves the derivative onto $e^{-i\xi x}$, giving $i\xi\hat{f}$; boundary terms vanish by decay.

11.3. Substituting $x - a \to x$ gives $e^{-ia\xi}\hat{f}$; substituting $ax \to x$ gives $\frac{1}{|a|}\hat{f}(\xi/a)$.

11.4. $\hat{f}' = -\frac{\xi}{2a}\hat{f}$ (differentiate under the integral and integrate by parts); solving with $\hat{f}(0) = \sqrt{\pi/a}$ gives $\hat{f} = \sqrt{\pi/a}\,e^{-\xi^2/4a}$.

11.5. Substitute $z = x - y$ and factor the double integral into $\hat{f}(\xi)\hat{g}(\xi)$.

11.6. Energy in physical space equals (up to 2π) energy in frequency space; the transform redistributes energy among frequencies without changing the total. It is the continuous analogue of Parseval for series.

11.7. $\hat{u}_t = -k\xi^2\hat{u} \Rightarrow \hat{u} = \hat{\phi}e^{-k\xi^2 t}$; inverting with $\mathcal{F}^{-1}[e^{-k\xi^2 t}] = S(\cdot, t)$ gives the heat-kernel convolution.

11.8. $\hat{u} = \hat{\phi}\cos(c\xi t) + \hat{\psi}\frac{\sin(c\xi t)}{c\xi}$; $\cos(c\xi t) = \frac{1}{2}(e^{ic\xi t} + e^{-ic\xi t})$ and the shift rule turn the $\hat{\phi}$ term into $\frac{1}{2}[\phi(x+ct) + \phi(x-ct)]$.

11.9. Rate $k\xi^2$ grows with $|\xi|$, so high frequencies (sharp features) decay fastest — the smoothing.

11.10. $\hat{u}_t = -ic\xi\hat{u} \Rightarrow \hat{u} = \hat{\phi}e^{-ic\xi t}$; the shift rule (multiplication by $e^{-ic\xi t}$ is translation by ct) gives $u = \phi(x - ct)$.

11.11. A narrow Gaussian (large a) transforms to a wide one ($e^{-\xi^2/4a}$, small $1/4a$). Narrowing f widens $\hat{f}$.

Chapter 12

12.1. (a) $1/s$, $s > 0$; (b) $1/(s-a)$, $s > a$; (c) $1/s^2$, $s > 0$.

12.2. $\frac{3}{s} + \frac{2}{s+1} - \frac{2}{s^2+4}$.

12.3. s-shift: $\mathcal{L}\{e^{-t}\cos 3t\} = \frac{s+1}{(s+1)^2+9}$. t-shift: $\mathcal{L}\{H(t-2)\} = e^{-2s}/s$.

12.4. $(s-2)U = 5 \Rightarrow U = 5/(s-2) \Rightarrow u = 5e^{2t}$.

12.5. $(s+1)U = 1/s^2 \Rightarrow U = \frac{1}{s^2} - \frac{1}{s} + \frac{1}{s+1}$, so $u = t - 1 + e^{-t}$.

12.6. $(s^2+4)U = 6 \Rightarrow U = \frac{6}{s^2+4} = 3 \cdot \frac{2}{s^2+4}$, so $u = 3\sin 2t$.

12.7. $U = \frac{s+2}{(s+1)^2} = \frac{1}{s+1} + \frac{1}{(s+1)^2}$, so $u = e^{-t} + te^{-t}$. The repeated root $s = -1$ gives the factor t (critical damping).

12.8. $\frac{1}{s(s+1)} = \frac{1}{s} - \frac{1}{s+1}$, inverse $1 - e^{-t}$; as a convolution, $1 * e^{-t} = \int_0^t e^{-\tau}d\tau = 1 - e^{-t}$. Agree.

12.9. $\int_0^\infty f'e^{-st} = [fe^{-st}]_0^\infty + s\int fe^{-st} = -f(0) + sF$. Applying to f' gives $s^2F - sf(0) - f'(0)$.

12.10. $\frac{s+3}{(s+1)(s+2)} = \frac{2}{s+1} - \frac{1}{s+2}$, inverse $2e^{-t} - e^{-2t}$.

12.11. (a) Fourier (unbounded space, decay); (b) Laplace (initial-value problem in t); (c) Laplace in t (initial condition plus a boundary condition at $x = 0$).

Chapter 13

13.1. (a) $2 - 2 = 0$; (b) $0 + 0 = 0$; (c) $e^x \cos y - e^x \cos y = 0$; (d) direct computation 0 away from origin.

13.2. At a steady state heat flows in and out in balance, so the temperature at a point equals the average of its surroundings — the mean value property.

13.3. By the maximum principle, $2 \le u \le 4$ throughout (boundary min 2, max 4).

13.4. Difference w harmonic, $w = 0$ on the boundary; maximum principle gives max and min 0, so $w \equiv 0$. Harmonicity is used to invoke the principle.

13.5. Integrating $\Delta u = 0$ over Ω and applying the divergence theorem gives $\oint \partial u/\partial n = 0$, so $\oint g = 0$ is necessary. The solution is determined up to a constant because adding a constant changes neither Δu nor $\partial u/\partial n$.

13.6. $u = \frac{\sinh y}{\sinh \pi} \sin x$ (only $n = 1$). Largest near the hot top edge $y = \pi$.

13.7. Angular equation $\Theta'' + n^2\Theta = 0$, periodicity forces integer n; radial Euler equation gives r^n, r^{-n}, and r^{-n} blows up at $r = 0$, so dropped.

13.8. At $r = 0$ the kernel is $R^2/R^2 = 1$, so $u(0) = \frac{1}{2\pi}\int_0^{2\pi} g\,d\phi$ — the mean value property.

13.9. $u = 2r\cos\theta = 2x$ (the $n = 1$ cosine term, $a_1 = 2$).

13.10. Match term by term: $u = 1 + r\cos\theta + r^2 \sin 2\theta$ (constant, $n = 1$ cosine, $n = 2$ sine). On $r = 1$ it equals g.

Chapter 14

14.1. Forward $\frac{u(x+h)-u(x)}{h}$ ($O(h)$); backward $\frac{u(x)-u(x-h)}{h}$ ($O(h)$); centered $\frac{u(x+h)-u(x-h)}{2h}$ ($O(h^2)$); second $\frac{u(x+h)-2u(x)+u(x-h)}{h^2}$ ($O(h^2)$).

14.2. Forward $= u' + \frac{h}{2}u'' + \cdots = u' + O(h)$. Centered: the $\frac{h}{2}u''$ terms cancel, leaving $u' + O(h^2)$.

14.3. Adding $u(x \pm h)$ cancels odd terms; dividing by h^2 gives $u'' + \frac{h^2}{12}u'''' + \cdots$, error $O(h^2)$, leading term $\frac{h^2}{12}u''''$.

14.4. True $-\sin 1 \approx -0.8415$; centered second difference gives ≈ -0.8408 ($h = 0.1$) and ≈ -0.8413 ($h = 0.05$), errors 7.0×10^{-4} and 1.75×10^{-4}, dropping by ≈ 4.

14.5. $u_j^{n+1} = u_j^n + r(u_{j+1}^n - 2u_j^n + u_{j-1}^n)$, $r = \Delta t/h^2$. With $h = 0.1$: $\Delta t \le \frac{1}{2}(0.01) = 0.005$.

14.6. Substituting $u_j^n = G^n e^{i\beta jh}$ gives $G = 1 - 4r\sin^2(\beta h/2)$; $|G| \le 1$ needs $0 \le r \le \frac{1}{2}$. The most unstable frequency is $\beta h = \pi$ (the sawtooth, $\sin^2 = 1$), giving $G = 1 - 4r$.

14.7. $r = k\Delta t/h^2$, so halving h multiplies r by 4; to keep $r \le \frac{1}{2}$, Δt must be quartered. The step count grows like $1/h^2$.

14.8. $G = 1/(1 + 4r\sin^2(\beta h/2))$, which is ≤ 1 for every $r > 0$: unconditionally stable. Cost: a (tridiagonal) linear solve per step.

14.9. $c\Delta t/h \le 1$: in one step a signal travels $\le h$, so the numerical domain of dependence contains the true one (the characteristics). Otherwise the scheme misses the data determining the solution and is unstable.

14.10. Explicit: cheap per step, CFL-restricted, $O(\Delta t)$. Implicit: linear solve per step, unconditionally stable, $O(\Delta t)$. Crank–Nicolson: linear solve per step, unconditionally stable, $O(\Delta t^2)$. Crank–Nicolson is preferable when both stability and time-accuracy matter.

14.11. Truncation (from differencing; $O(h^p)$, shrinks with h); stability (amplification of existing errors; refining h without respecting CFL makes it worse); round-off (finite precision; a floor unaffected by h).

Further Reading

These notes are an introduction, and a student who wishes to go further — or who wants a second account of a difficult point — has excellent options. The suggestions below are organized by purpose rather than presented as a formal bibliography; specific editions and full citation details can be filled in as the course adopts them.

The primary companion. The natural text to read alongside these notes is *Strauss, Partial Differential Equations: An Introduction*, whose notation and ordering we have followed throughout. Its treatment of the energy method, the maximum principle, and separation of variables is close to ours, and its exercises are a rich source of additional practice. A student should regard Strauss as the main reference and these notes as a streamlined path through it.

For applied and computational emphasis. *Haberman, Applied Partial Differential Equations with Fourier Series and Boundary Value Problems* offers a more leisurely, application-oriented account, with especially thorough chapters on Fourier series, Sturm–Liouville theory, and the Laplace transform, and a fuller treatment of problems in cylindrical and spherical geometry. It complements Strauss well for readers who want more worked examples and physical motivation. For the numerical chapter, *LeVeque, Finite Difference Methods for Ordinary and Partial Differential Equations* develops the consistency–stability–convergence framework with care and rigor.

For deeper theory. A student continuing to a graduate course will encounter *Evans, Partial Differential Equations*, the standard modern reference, which develops the theory of the three equation types, Sobolev spaces, and the modern existence theory far beyond the level of an undergraduate course. It is demanding but authoritative, and the early chapters on the wave, heat, and Laplace equations reward a look even now. For the classical theory of Fourier series and their convergence, *Stein and Shakarchi, Fourier Analysis: An Introduction* is lucid and self-contained.

On the connections to modern computation. The methods of this course — separation of variables, transforms, finite differences — are the classical foundation on which contemporary scientific computing is built. A reader curious about how these ideas extend to the machine-learning methods now used for PDEs (neural-network solvers, operator learning, equation discovery from data) will find that every such method rests on the classical understanding developed here: what a well-posed problem is, why the type of an equation governs its behavior, and how solutions depend on their data. The classical theory is not superseded by the computational one; it is its prerequisite.

A closing word. The best way to learn partial differential equations is to solve them. The exercises in these notes, and the far larger collections in the texts above, are where the understanding is actually built. A formula read is quickly forgotten; a problem worked is retained. The reader is encouraged to treat the solution sketches in the appendix as a check on completed work, not a substitute for it.