Chapter 2
Linear and Nonlinear Waves

Our initial foray into the vast mathematical continent that comprises partial differential equations will begin with some basic first-order equations. In applications, first-order partial differential equations are most commonly used to describe dynamical processes, and so time, t, is one of the independent variables. Our discussion will focus on dynamical models in a single space dimension, bearing in mind that most of the methods we introduce can be extended to higher-dimensional situations. First-order partial differential equations and systems model a wide variety of wave phenomena, including transport of pollutants in fluids, flood waves, acoustics, gas dynamics, glacier motion, chromatography, traffic flow, and various biological and ecological systems.

A basic solution technique relies on an inspired change of variables, which comes from rewriting the equation in a moving coordinate frame. This naturally leads to the fundamental concept of characteristic curve, along which signals and physical disturbances propagate. The resulting method of characteristics is able to solve a first-order *linear* partial differential equation by reducing it to one or more first-order *nonlinear* ordinary differential equations.

Proceeding to the nonlinear regime, the most important new phenomenon is the possible breakdown of solutions in finite time, resulting in the formation of discontinuous shock waves. A familiar example is the supersonic boom produced by an airplane that breaks the sound barrier. Signals continue to propagate along characteristic curves, but now the curves may cross each other, precipitating the onset of a shock discontinuity. The ensuing shock dynamics is *not* uniquely specified by the partial differential equation, but relies on additional physical properties, to be specified by an appropriate conservation law along with a causality condition. A full-fledged analysis of shock dynamics becomes quite challenging, and only the basics will be developed here.

Having attained a basic understanding of first-order wave dynamics, we then focus our attention on the first of three paradigmatic second-order partial differential equations, known as the wave equation, which is used to model waves and vibrations in an elastic bar, a violin string, or a column of air in a wind instrument. Its multi-dimensional versions serve to model vibrations of membranes, solid bodies, water waves, electromagnetic waves, including light, radio waves, microwaves, acoustic waves, and many other physical phenomena. The one-dimensional wave equation is one of a small handful of physically relevant partial differential equations that has an explicit solution formula, originally discovered by the eighteenth-century French mathematician (and encyclopedist) Jean d'Alembert. His solution is the result of being able to "factorize" the second-order wave equation into a pair of first-order partial differential equations, of a type solved in the first part of this

P.J. Olver, *Introduction to Partial Differential Equations*, Undergraduate Texts in Mathematics,
DOI 10.1007/978-3-319-02099-0_2,

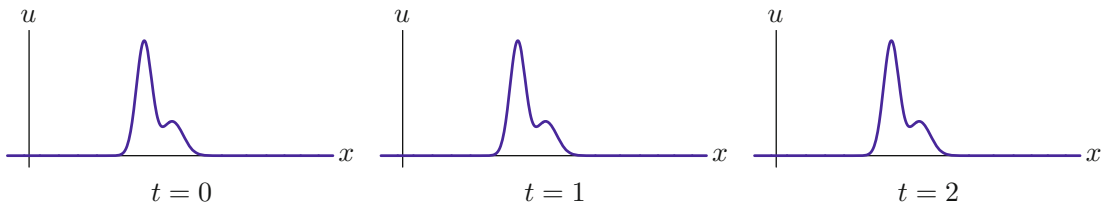


Figure 2.1. Stationary wave.

chapter. We investigate the consequences of d'Alembert's solution formula for the initial value problem on the entire real line; solutions on bounded intervals will be deferred until Chapter 4. Unfortunately, d'Alembert's method is of rather limited scope, and does not extend beyond the one-dimensional case, nor to equations modeling vibrations of nonuniform media. The analysis of the wave equation in more than one space dimension can be found in Chapters 11 and 12.

2.1 Stationary Waves

When entering a new mathematical subject — in our case, partial differential equations — one should first analyze and fully understand the very simplest examples. Indeed, mathematics is, at its core, a bootstrapping enterprise, in which one builds on one's knowledge of and experience with elementary topics — in the present case, ordinary differential equations — to make progress, first with the simpler types of partial differential equations, and then, by developing and applying each newly gained insight and technique, to more and more complicated situations.

The simplest partial differential equation, for a function $u(t, x)$ of two variables, is

$$\frac{\partial u}{\partial t} = 0. \tag{2.1}$$

It is a first-order, homogeneous, linear equation. If (2.1) were an ordinary differential equation† for a function $u(t)$ of t alone, the solution would be obvious: $u(t) = c$ must be constant. A proof of this basic fact proceeds by integrating both sides with respect to t and then appealing to the Fundamental Theorem of Calculus. To solve (2.1) as a partial differential equation for $u(t, x)$, let us similarly integrate both sides of the equation from, say, 0 to t, producing

$$0 = \int_0^t \frac{\partial u}{\partial t}\,(s, x)\,ds = u(t, x) - u(0, x).$$

Therefore, the solution takes the form

$$u(t, x) = f(x), \qquad \text{where} \qquad f(x) = u(0, x), \tag{2.2}$$

and hence is a function of the space variable x alone. The only requirement is that $f(x)$ be continuously differentiable, so $f \in \mathrm{C}^1$, in order that $u(t, x)$ be a bona fide classical

† Of course, in this situation, we would write the equation as $du/dt = 0$.

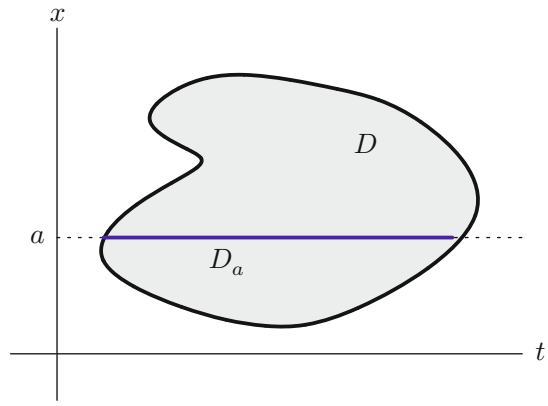


Figure 2.2. Domain for stationary-wave solution.

solution of the first-order partial differential equation (2.1). The solution (2.2) represents a *stationary wave*, meaning that it does not change in time. The initial profile stays frozen in place, and the system remains in equilibrium. Figure 2.1 plots a representative solution as a function of x at three successive times.

The preceding analysis seems very straightforward and perhaps even a little boring. But, to be completely rigorous, we need to take a bit more care. In our derivation, we implicitly assumed that the solution $u(t, x)$ was defined everywhere on $\mathbb{R}^2$. And, in fact, the solution formula (2.2) is *not* completely valid as stated if the solution $u(t, x)$ is defined only on a subdomain $D \subset \mathbb{R}^2$.

Indeed, a solution $u(t)$ to the corresponding ordinary differential equation $du/dt = 0$ is constant, *provided it is defined on a connected subinterval* $I \subset \mathbb{R}$. A solution that is defined on a disconnected subset $D \subset \mathbb{R}$ need only be constant on each connected subinterval $I \subset D$. For instance, the nonconstant function

$$u(t) = \begin{cases} 1, & t > 0, \\ -1, & t < 0, \end{cases} \qquad \text{satisfies} \qquad \frac{du}{dt} = 0$$

everywhere on its domain of definition, that is, $D = \{\, t \neq 0 \,\}$, but is constant only on the connected positive and negative half-lines.

Similar counterexamples can be constructed in the case of the partial differential equation (2.1). If the domain of definition is disconnected, then we do not expect $u(t, x)$ to depend only on x if we move from one connected component of D to another. Even that is not the full story. For example, the function

$$u(t, x) = \begin{cases} 0, & x > 0, \\ x^2, & x \leq 0, \quad t > 0, \\ -x^2, & x \leq 0, \quad t < 0, \end{cases} \tag{2.3}$$

is continuously differentiable† on its domain of definition, namely $D = \mathbb{R}^2 \setminus \{\, (0, x) \,|\, x \leq 0 \,\}$, satisfies $\partial u/\partial t = 0$ everywhere in D, but, nevertheless, is not a function of x alone, because, for example, $u(1, x) = x^2 \neq u(-1, x) = -x^2$.

† You are asked to rigorously prove differentiability in Exercise 2.1.10.

A completely correct formulation can be stated as follows: If $u(t,x)$ is a classical solution to (2.1), defined on a domain $D \subset \mathbb{R}^2$ whose intersection with any horizontal[‡] line, namely $D_a = D \cap \{\,(t,a)\,|\,t \in \mathbb{R}\,\}$, for each fixed $a \in \mathbb{R}$, is either empty or a connected interval, then $u(t,x) = f(x)$ is a function of x alone. An example of such a domain is sketched in Figure 2.2. In Exercise 2.1.9, you are asked to justify these statements.

We are thus slightly chastened in our dismissal of (2.1) as a complete triviality. The lesson is that, in future, one must *always* be careful when interpreting such "general" solution formulas — since they often rely on unstated assumptions on their underlying domain of definition.

Exercises

2.1.1. Solve the partial differential equation $\dfrac{\partial u}{\partial t} = x$ for $u(t,x)$.

2.1.2. Solve the partial differential equation $\dfrac{\partial^2 u}{\partial t^2} = 0$ for $u(t,x)$.

2.1.3. Find the general solution $u(t,x)$ to the following partial differential equations:
(a) $u_x = 0$, (b) $u_t = 1$, (c) $u_t = x - t$, (d) $u_t + 3u = 0$, (e) $u_x + t\,u = 0$, (f) $u_{tt} + 4u = 1$.

2.1.4. Suppose $u(t,x)$ is defined for all $(t,x) \in \mathbb{R}^2$ and solves $\partial u/\partial t + 2u = 0$. Prove that $\lim_{t \to \infty} u(t,x) = 0$ for all x.

2.1.5. Write down the general solution to the partial differential equation $\partial u/\partial t = 0$ for a function of three variables $u(t,x,y)$. What assumptions should be made on the domain of definition for your solution formula to be valid?

2.1.6. Solve the partial differential equation $\dfrac{\partial^2 u}{\partial x\,\partial y} = 0$ for $u(x,y)$.

2.1.7. Answer Exercise 2.1.6 when $u(x,y,z)$ depends on the three independent variables x, y, z.

♡ 2.1.8. Let $u(t,x)$ solve the initial value problem $\dfrac{\partial u}{\partial t} + u^2 = 0$, $u(0,x) = f(x)$, where $f(x)$ is a bounded C^1 function of $x \in \mathbb{R}$. (a) Show that if $f(x) \geq 0$ for all x, then $u(t,x)$ is defined for all $t > 0$, and $\lim_{t \to \infty} u(t,x) = 0$. (b) On the other hand, if $f(x) < 0$, then the solution $u(t,x)$ is not defined for all $t > 0$, but in fact, $\lim_{t \to \tau^-} u(t,x) = -\infty$ for some $0 < \tau < \infty$. Given x, what is the corresponding value of τ? (c) Given $f(x)$ as in part (b), what is the longest time interval $0 < t < t_\star$ on which $u(t,x)$ is defined for all $x \in \mathbb{R}$?

♢ 2.1.9. Justify the claim in the text that if $u(t,x)$ is a solution of $\partial u/\partial t = 0$ that is defined on a domain $D \subset \mathbb{R}^2$ with the property that $D_a = D \cap \{\,(a,x)\,|\,x \in \mathbb{R}\,\}$ is either empty or a connected interval, then $u(t,x) = v(x)$ depends only on $x \in D$.

♢ 2.1.10. Prove that the function in (2.3) is continuously differentiable at all points (t,x) in its domain of definition.

[‡] *Important*: We will adopt the (slightly unusual) convention of displaying the (t,x)–plane with time t along the horizontal axis and space x along the vertical axis — which also conforms with our convention of writing t before x in expressions like $u(t,x)$. Later developments will amply vindicate our adoption of this convention.

2.2 Transport and Traveling Waves

In many respects, the stationary-wave equation (2.1) does not quite qualify as a partial differential equation. Indeed, the spatial variable x enters only parametrically in the solution to what is, in essence (ignoring technical difficulties with domains), a very simple ordinary differential equation.

Let us then turn to a more "genuine" example. Consider the linear, homogeneous first-order partial differential equation

$$\frac{\partial u}{\partial t} + c\,\frac{\partial u}{\partial x} = 0, \tag{2.4}$$

for a function $u(t, x)$, in which c is a fixed, nonzero constant, known as the *wave speed* for reasons that will soon become apparent. We will refer to (2.4) as the *transport equation*, because it models the transport of a substance, e.g., a pollutant, in a uniform fluid flow that is moving with velocity c. In this model, the solution $u(t, x)$ represents the concentration of the pollutant at time t and spatial position x. Other common names for (2.4) are the *first-order* or *unidirectional wave equation*. But for brevity, as well as to avoid any confusion with the second-order, bidirectional wave equation discussed extensively later on, we will stick with the designation "transport equation" here. Solving the transport equation is slightly more challenging, but, as we will see, not difficult.

Since the transport equation involves time, its solutions are distinguished by their initial values. As a first-order equation, we need only specify the value of the solution at an initial time t_0, leading to the initial value problem

$$u(t_0, x) = f(x) \qquad \text{for all} \qquad x \in \mathbb{R}. \tag{2.5}$$

As we will show, as long as $f \in \mathrm{C}^1$, i.e., is continuously differentiable, the initial conditions serve to specify a unique classical solution. Also, by replacing the time variable t by $t - t_0$, we can, without loss of generality, set $t_0 = 0$.

Uniform Transport

Let us begin by assuming that the wave speed c is constant. In general, when one is confronted with a new equation, one solution strategy is to try to convert it into an equation that you already know how to solve. In this case, we will introduce a simple change of variables that effectively rewrites the equation in a moving coordinate system, inspired by the interpretation of c as the overall transport speed.

If x represents the position of an object in a fixed coordinate frame, then

$$\xi = x - ct \tag{2.6}$$

represents the object's position relative to an observer who is uniformly moving with velocity c. Think of a passenger in a moving train to whom stationary objects appear to be moving *backwards* at the train's speed c. To formulate a physical process in the reference frame of the passenger, we replace the stationary space-time coordinates (t, x) by the moving coordinates (t, ξ).

Remark: These are the same changes of reference frame that underlie Einstein's special theory of relativity. However, unlike Einstein, we are working in a purely classical,

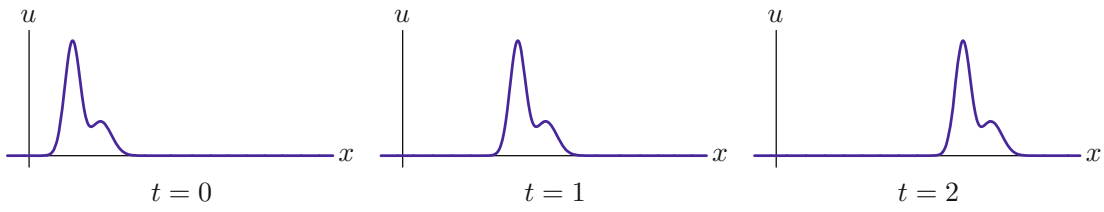


Figure 2.3. Traveling wave with $c > 0$.

nonrelativistic universe here. Such changes to moving coordinates are, in fact, of a much older vintage, and named *Galilean boosts* in honor of Galileo Galilei, who was the first to champion such "relativistic" moving coordinate systems.

Let us see what happens when we re-express the transport equation in terms of the moving coordinate frame. We rewrite

$$u(t, x) = v(t, x - ct) = v(t, \xi) \tag{2.7}$$

in terms of the *characteristic variable* $\xi = x - ct$, along with the time t. To write out the differential equation satisfied by $v(t, \xi)$, we apply the chain rule from multivariable calculus, [**8**, **108**], to express the derivatives of u in terms of those of v:

$$\frac{\partial u}{\partial t} = \frac{\partial v}{\partial t} - c\,\frac{\partial v}{\partial \xi}\,, \qquad \frac{\partial u}{\partial x} = \frac{\partial v}{\partial \xi}\,.$$

Therefore,

$$\frac{\partial u}{\partial t} + c\,\frac{\partial u}{\partial x} = \frac{\partial v}{\partial t} - c\,\frac{\partial v}{\partial \xi} + c\,\frac{\partial v}{\partial \xi} = \frac{\partial v}{\partial t}\,. \tag{2.8}$$

We deduce that $u(t, x)$ solves the transport equation (2.4) if and only if $v(t, \xi)$ solves the stationary-wave equation

$$\frac{\partial v}{\partial t} = 0. \tag{2.9}$$

Thus, the effect of using a moving coordinate system is to convert a wave moving with velocity c into a stationary wave. Think again of the passenger in the train — a second train moving at the same speed appears as if it were stationary.

According to our earlier discussion, the solution $v = v(\xi)$ to the stationary-wave equation (2.9) is a function of the characteristic variable alone. (For simplicity, we assume that $v(t, \xi)$ has an appropriate domain of definition, e.g., it is defined everywhere on $\mathbb{R}^2$.) Recalling (2.7), we conclude that the solution

$$u = v(\xi) = v(x - ct)$$

to the transport equation must be a function of the characteristic variable only. We have therefore proved the following result:

Proposition 2.1. *If $u(t, x)$ is a solution to the partial differential equation*

$$u_t + c\,u_x = 0, \tag{2.10}$$

which is defined on all of $\mathbb{R}^2$, then

$$u(t, x) = v(x - ct), \tag{2.11}$$

where $v(\xi)$ is a C^1 function of the characteristic variable $\xi = x - ct$.

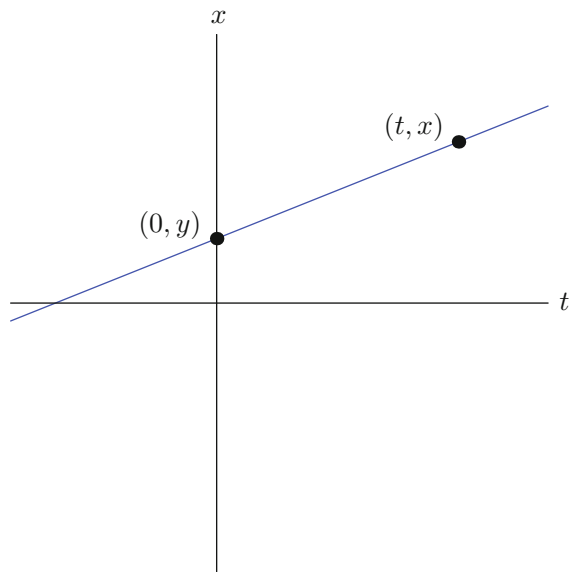


Figure 2.4. Characteristic line.

In other words, *any* (reasonable) function of the characteristic variable, e.g., ξ^2+1, or $\cos\xi$, or e^{ξ}, will produce a corresponding solution, $(x-ct)^2+1$, or $\cos(x-ct)$, or e^{x-ct}, to the transport equation with constant wave speed c. And, in accordance with the counting principle of Chapter 1, the general solution to this first-order partial differential equation in two independent variables depends on one arbitrary function of a single variable.

To a stationary observer, the solution (2.11) appears as a *traveling wave* of unchanging form moving at constant velocity c. When $c>0$, the wave translates to the right, as illustrated in Figure 2.3. When $c<0$, the wave translates to the left, while $c=0$ corresponds to a stationary wave form that remains fixed at its original location, as in Figure 2.1.

At $t=0$, the wave has the initial profile

$$u(0,x)=v(x), \tag{2.12}$$

and so (2.11) provides the (unique) solution to the initial value problem (2.4, 12). For example, the solution to the particular initial value problem

$$u_t+2\,u_x=0, \qquad u(0,x)=\frac{1}{1+x^2}\,, \qquad \text{is} \qquad u(t,x)=\frac{1}{1+(x-2t)^2}\,.$$

Since it depends only on the characteristic variable $\xi=x-ct$, every solution to the transport equation is constant on the *characteristic lines* of slope[†] c, namely

$$x=ct+k, \tag{2.13}$$

where k is an arbitrary constant. At any given time t, the value of the solution at position x depends only on its original value on the characteristic line passing through (t,x).

[†] This makes use of our convention that the t–axis is horizontal and the x–axis is vertical. Reversing the axes will replace the slope by its reciprocal.

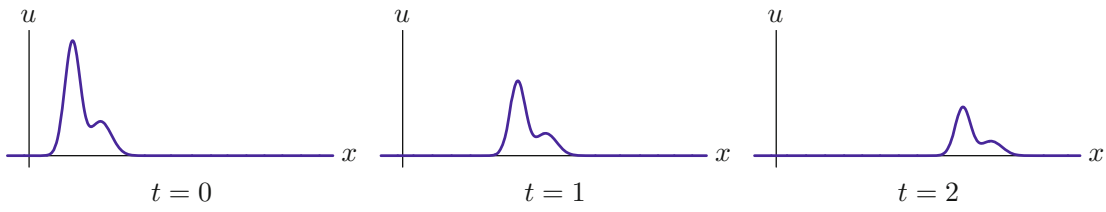


Figure 2.5. Decaying traveling wave.

This is indicative of a general fact concerning such wave models: *Signals propagate along characteristics*. Indeed, a disturbance at an initial point $(0, y)$ only affects the value of the solution at points (t, x) that lie on the characteristic line $x = ct + y$ emanating therefrom, as illustrated in Figure 2.4.

Transport with Decay

Let $a > 0$ be a positive constant, and c an arbitrary constant. The homogeneous linear first-order partial differential equation

$$\frac{\partial u}{\partial t} + c\,\frac{\partial u}{\partial x} + a\,u = 0 \tag{2.14}$$

models the transport of, say, a radioactively decaying solute in a uniform fluid flow with wave speed c. The coefficient a governs the rate of decay. We can solve this variant of the transport equation by the self-same change of variables to a uniformly moving coordinate system.

Rewriting $u(t, x)$ in terms of the characteristic variable, as in (2.7), and then recalling our chain rule calculation (2.8), we find that $v(t, \xi) = u(t, \xi + ct)$ satisfies the partial differential equation

$$\frac{\partial v}{\partial t} + a\,v = 0.$$

The result is, effectively, a homogeneous linear first-order ordinary differential equation, in which the characteristic variable ξ enters only parametrically. The standard solution technique learned in elementary ordinary differential equations, [**20**, **23**], tells us to multiply the equation by the exponential *integrating factor* $e^{a\,t}$, leading to

$$e^{a\,t}\left(\frac{\partial v}{\partial t} + a\,v\right) = \frac{\partial}{\partial t}\,(e^{a\,t}v) = 0.$$

We conclude that $w = e^{a\,t}v$ solves the stationary-wave equation (2.1). Thus,

$$w = e^{a\,t}v = f(\xi), \qquad \text{and hence} \qquad v(t, \xi) = f(\xi)\,e^{-a\,t},$$

where $f(\xi)$ is an arbitrary function of the characteristic variable. Reverting to physical coordinates, we produce the solution formula

$$u(t, x) = f(x - ct)\,e^{-a\,t}, \tag{2.15}$$

which solves the initial value problem $u(0, x) = f(x)$. It represents a wave that is moving along with fixed velocity c while simultaneously decaying at an exponential rate as prescribed by the coefficient $a > 0$. A typical solution, for $c > 0$, is plotted at three successive

times in Figure 2.5. While the solution (2.15) is no longer constant on the characteristics, signals continue to propagate along them, since a solution's initial value at a point $(0, y)$ will only affect its subsequent (decaying) values on the associated characteristic line $x = ct + y$.

Exercises

2.2.1. Find the solution to the initial value problem $u_t + u_x = 0$, $u(1, x) = x/(1 + x^2)$.

2.2.2. Solve the following initial value problems and graph the solutions at times $t = 1, 2$, and 3:
(a) $u_t - 3u_x = 0$, $u(0, x) = e^{-x^2}$; (b) $u_t + 2u_x = 0$, $u(-1, x) = x/(1 + x^2)$;
(c) $u_t + u_x + \frac{1}{2}u = 0$, $u(0, x) = \tan^{-1} x$; (d) $u_t - 4u_x + u = 0$, $u(0, x) = 1/(1 + x^2)$.

2.2.3. Graph some of the characteristic lines for the following equations, and write down a formula for the general solution:
(a) $u_t - 3u_x = 0$, (b) $u_t + 5u_x = 0$, (c) $u_t + u_x + 3u = 0$, (d) $u_t - 4u_x + u = 0$.

2.2.4. Solve the initial value problem $u_t + 2u_x = 1$, $u(0, x) = e^{-x^2}$.
Hint: Use characteristic coordinates.

2.2.5. Answer Exercise 2.2.4 for the initial value problem $u_t + 2u_x = \sin x$, $u(0, x) = \sin x$.

♢ 2.2.6. Let c be constant. Suppose that $u(t, x)$ solves the initial value problem $u_t + cu_x = 0$, $u(0, x) = f(x)$. Prove that $v(t, x) = u(t - t_0, x)$ solves the initial value problem $v_t + cv_x = 0$, $v(t_0, x) = f(x)$.

2.2.7. Is Exercise 2.2.6 valid when the transport equation is replaced by the damped transport equation (2.14)?

2.2.8. Let $c \neq 0$. (a) Prove that if the initial data satisfies $u(0, x) = v(x) \to 0$ as $x \to \pm\infty$, then, for each fixed x, the solution to the transport equation (2.4) satisfies $u(t, x) \to 0$ as $t \to \infty$. (b) Is the convergence uniform in x?

2.2.9. (a) Prove that if the initial data is bounded, $|f(x)| \leq M$ for all $x \in \mathbb{R}$, then the solution to the damped transport equation (2.14) with $a > 0$ satisfies $u(t, x) \to 0$ as $t \to \infty$. (b) Find a solution to (2.14) that is defined for all (t, x) but does not satisfy $u(t, x) \to 0$ as $t \to \infty$.

2.2.10. Let $F(t, x)$ be a C^1 function of $(t, x) \in \mathbb{R}^2$. (a) Write down a formula for the general solution $u(t, x)$ to the inhomogeneous partial differential equation $u_t = F(t, x)$.
(b) Solve the inhomogeneous transport equation $u_t + c\,u_x = F(t, x)$.

♡ 2.2.11. (a) Write down a formula for the general solution to the nonlinear partial differential equation $u_t + u_x + u^2 = 0$. (b) Show that if the initial data is positive and bounded, $0 \leq u(0, x) = f(x) \leq M$, then the solution exists for all $t > 0$, and $u(t, x) \to 0$ as $t \to \infty$. (c) On the other hand, if the initial data is negative at some x, then the solution *blows up* at x in finite time: $\lim_{t \to \tau^-} u(t, x) \to -\infty$ for some $\tau > 0$. (d) Find a formula for the earliest blow-up time $\tau_\star > 0$.

2.2.12. A sensor situated at position $x = 1$ monitors the concentration of a pollutant $u(t, 1)$ as a function of t for $t \geq 0$. Assuming that the pollutant is transported with wave speed $c = 3$, at what locations x can you determine the initial concentration $u(0, x)$?

2.2.13. Write down a solution to the transport equation $u_t + 2u_x = 0$ that is defined on a connected domain $D \subset \mathbb{R}^2$ and that is *not* a function of the characteristic variable alone.

2.2.14. Let $c > 0$. Consider the uniform transport equation $u_t + cu_x = 0$ restricted to the quarter-plane $Q = \{ x > 0,\ t > 0 \}$ and subject to initial conditions $u(0, x) = f(x)$ for $x \geq 0$, along with boundary conditions $u(t, 0) = g(t)$ for $t \geq 0$. (*a*) For which initial and boundary conditions does a classical solution to this initial-boundary value problem exist? Write down a formula for the solution. (*b*) On which regions are the effects of the initial conditions felt? What about the boundary conditions? Is there any interaction between the two?

2.2.15. Answer Exercise 2.2.14 when $c < 0$.

Nonuniform Transport

Slightly more complicated, but still linear, is the *nonuniform transport equation*

$$\frac{\partial u}{\partial t} + c(x)\,\frac{\partial u}{\partial x} = 0, \tag{2.16}$$

where the wave speed $c(x)$ is now allowed to depend on the spatial position. Characteristics continue to guide the behavior of solutions, but when the wave speed is not constant, we can no longer expect them to be straight lines. To adapt the method of characteristics, let us look at how the solution varies along a prescribed curve in the (t, x)–plane. Assume that the curve is identified with the graph of a function $x = x(t)$, and let

$$h(t) = u\big(t, x(t)\big)$$

be the value of the solution on it. We compute the rate of change in the solution along the curve by differentiating h with respect to t. Invoking the multivariable chain rule, we obtain

$$\frac{dh}{dt} = \frac{d}{dt}\,u\big(t, x(t)\big) = \frac{\partial u}{\partial t}\big(t, x(t)\big) + \frac{\partial u}{\partial x}\big(t, x(t)\big)\ \frac{dx}{dt}\,. \tag{2.17}$$

In particular, if $x(t)$ satisfies

$$\frac{dx}{dt} = c\big(x(t)\big), \qquad \text{then} \qquad \frac{dh}{dt} = \frac{\partial u}{\partial t}\big(t, x(t)\big) + c\big(x(t)\big)\ \frac{\partial u}{\partial x}\big(t, x(t)\big) = 0,$$

since we are assuming that $u(t, x)$ solves the transport equation (2.16) for all values of (t, x), including those points $\big(t, x(t)\big)$ on the curve. Since its derivative is zero, $h(t)$ must be a constant, which motivates the following definition.

Definition 2.2. The graph of a solution $x(t)$ to the autonomous ordinary differential equation

$$\frac{dx}{dt} = c(x) \tag{2.18}$$

is called a *characteristic curve* for the transport equation with wave speed $c(x)$.

In other words, at each point (t, x), the slope of the characteristic curve equals the wave speed $c(x)$ there. In particular, if c is constant, the characteristic curves are straight lines of slope c, in accordance with our earlier construction.

Proposition 2.3. *Solutions to the linear transport equation* (2.16) *are constant along characteristic curves.*

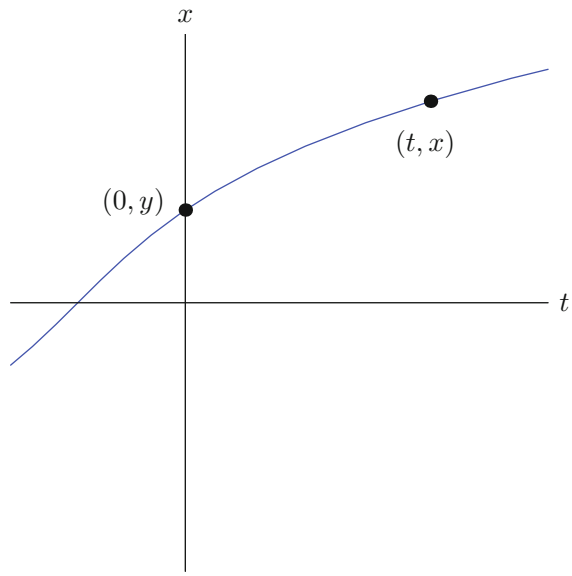


Figure 2.6. Characteristic curve.

The characteristic curve equation (2.18) is an autonomous first-order ordinary differential equation. As such, it can be immediately solved by separating variables, [**20**, **23**]. Assuming $c(x) \neq 0$, we divide both sides of the equation by $c(x)$, and then integrate the resulting equation:

$$\frac{dx}{c(x)} = dt, \qquad \text{whereby} \qquad \beta(x) := \int \frac{dx}{c(x)} = t + k, \tag{2.19}$$

with k denoting the integration constant. For each fixed value of k, (2.19) serves to implicitly define a characteristic curve, namely,

$$x(t) = \beta^{-1}(t + k),$$

with β^{-1} denoting the inverse function. On the other hand, if $c(x_\star) = 0$, then $x_\star$ is a *fixed point* for the ordinary differential equation (2.18), and the horizontal line $x \equiv x_\star$ is a stationary characteristic curve.

Since the solution $u(t, x)$ is constant along the characteristic curves, it must therefore be a function of the *characteristic variable*

$$\xi = \beta(x) - t \tag{2.20}$$

alone, and hence of the form

$$u(t, x) = v\big(\beta(x) - t\big), \tag{2.21}$$

where $v(\xi)$ is an arbitrary C^1 function. Indeed, it is easy to check directly that, provided $\beta(x)$ is defined by (2.19), $u(t, x)$ solves the partial differential equation (2.16) for *any* choice of C^1 function $v(\xi)$. (But keep in mind that the algebraic solution formula (2.21) may fail to be valid at points where the wave speed vanishes: $c(x_\star) = 0$.)

Warning: The definition of characteristic variable used here is slightly different from that in the constant wave speed case, which, by (2.20), would be $\xi = x/c - t = (x - ct)/c$. Clearly, rescaling the characteristic variable by $1/c$ is an inessential modification of our original definition.

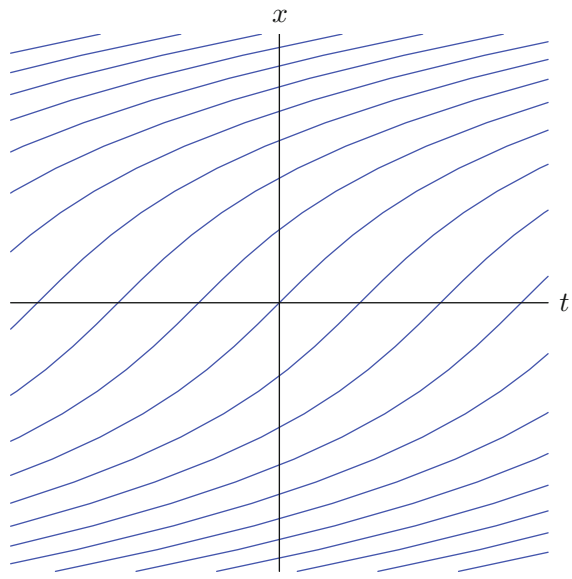


Figure 2.7. Characteristic curves for $u_t + (x^2+1)^{-1}u_x = 0$.

To find the solution that satisfies the prescribed initial conditions

$$u(0,x) = f(x), \tag{2.22}$$

we merely substitute the general solution formula (2.21). This leads to the implicit equation $v(\beta(x)) = f(x)$ for the function $v(\xi) = f \circ \beta^{-1}(\xi)$. The resulting solution formula

$$u(t,x) = f \circ \beta^{-1}\big(\beta(x) - t\big) \tag{2.23}$$

is not particularly enlightening, but it does have a simple graphical interpretation: To find the value of the solution $u(t,x)$, we look at the characteristic curve passing through the point (t,x). If this curve intersects the x–axis at the point $(0,y)$, as in Figure 2.6, then $u(t,x) = u(0,y) = f(y)$, since the solution must be constant along the curve. On the other hand, if the characteristic curve through (t,x) doesn't intersect the x–axis, the solution value $u(t,x)$ is *not* prescribed by the initial data.

Example 2.4. Let us solve the nonuniform transport equation

$$\frac{\partial u}{\partial t} + \frac{1}{x^2+1}\,\frac{\partial u}{\partial x} = 0 \tag{2.24}$$

by the method of characteristics. According to (2.18), the characteristic curves are the graphs of solutions to the first-order ordinary differential equation

$$\frac{dx}{dt} = \frac{1}{x^2+1}\,.$$

Separating variables and integrating, we obtain

$$\beta(x) = \int (x^2+1)\,dx = \tfrac{1}{3}x^3 + x = t + k, \tag{2.25}$$

where k is the integration constant. Representative curves are plotted in Figure 2.7. (In this case, inverting the function β, i.e., solving (2.25) for x as a function of t, is not particularly enlightening.)

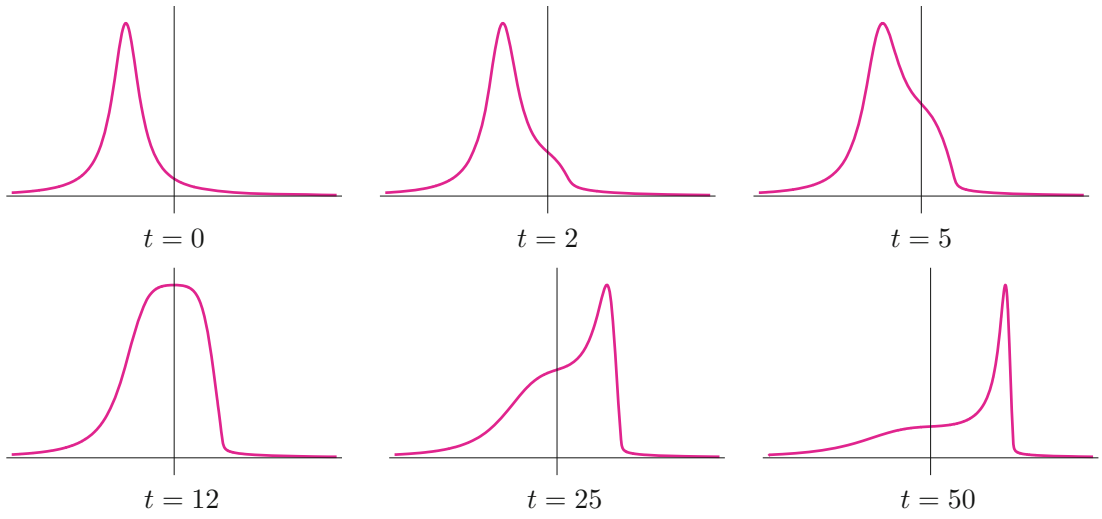


Figure 2.8. Solution to $u_t + \dfrac{1}{x^2+1}\, u_x = 0$.

According to (2.20), the characteristic variable is $\xi = \frac{1}{3}x^3 + x - t$, and hence the general solution to the equation takes the form

$$u = v\big(\tfrac{1}{3}x^3 + x - t\big), \tag{2.26}$$

where $v(\xi)$ is an arbitrary C^1 function. A typical solution, corresponding to initial data

$$u(0,x) = \frac{1}{1+(x+3)^2}\,, \tag{2.27}$$

is plotted† at the indicated times in Figure 2.8. Although the solution remains constant along each individual curve, a stationary observer will witness a dynamically changing profile as the wave moves through the nonuniform medium. In this example, since $c(x) > 0$ everywhere, the wave always moves from left to right; its speed as it passes through a point x determined by the magnitude of $c(x) = (x^2+1)^{-1}$, with the consequence that each part accelerates as it approaches the origin from the left, and then slows back down once it passes by and $c(x)$ decreases in magnitude. To a stationary observer, the wave spreads out as it speeds through the origin, and then becomes progressively narrower and slower as it gradually moves off to $+\infty$.

Example 2.5. Consider the nonuniform transport equation

$$u_t + (x^2-1)\,u_x = 0. \tag{2.28}$$

† The required function $v(\xi)$ in (2.26) is implicitly given by the equation $v\big(\frac{1}{3}x^3 + x\big) = u(0,x)$, and so the explicit formula for $u(t,x)$ is not very instructive or useful. Indeed, to make the plots, we instead sampled the initial data (2.27) at a collection of uniformly spaced points $y_1 < y_2 < \cdots < y_n$. Since the solution is constant along the characteristic curve (2.25) passing through each sample point $(0, y_i)$, we can find nonuniformly spaced sample values for $u(t, x_i)$ at any later time. The smooth solution curve $u(t,x)$ is then approximated using spline interpolation, [**89**; §11.4], on these sample values.

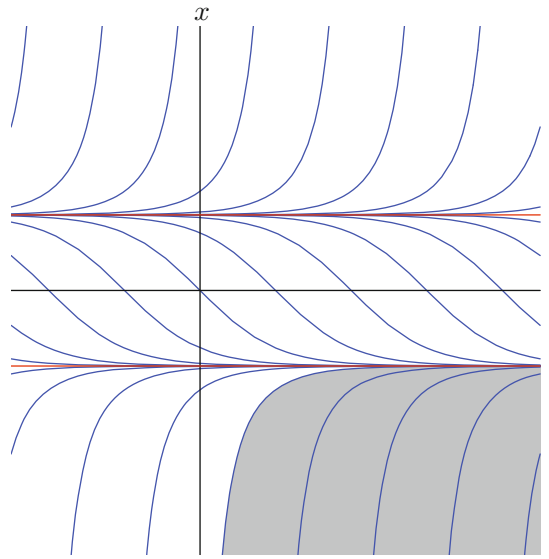


Figure 2.9. Characteristic curves for $u_t + (x^2 - 1)u_x = 0$.

In this case, the characteristic curves are the solutions to

$$\frac{dx}{dt} = x^2 - 1,$$

and so

$$\beta(x) = \int \frac{dx}{x^2 - 1} = \frac{1}{2} \log \left| \frac{x-1}{x+1} \right| = t + k. \tag{2.29}$$

One must also include the horizontal lines $x = x_\pm = \pm 1$ corresponding to the roots of $c(x) = x^2 - 1$. The curves are graphed in Figure 2.9. Note that those curves starting below $x_+ = 1$ converge to $x_- = -1$ as $t \to \infty$, while those starting above $x_+ = 1$ veer off to ∞ in finite time. Owing to the sign of $c(x) = x^2 - 1$, points on the graph of $u(0, x)$ lying over $|\, x \,| < 1$ will move to the left, while those over $|\, x \,| > 1$ will move to the right.

In Figure 2.10, we graph several snapshots of the solution whose initial value is a bell-shaped Gaussian profile

$$u(0, x) = e^{-x^2}.$$

The initial conditions uniquely prescribe the value of the solution along the characteristic curves that intersect the x–axis. On the other hand, if

$$x \le \frac{1 + e^{2t}}{1 - e^{2t}} \qquad \text{for} \qquad t > 0,$$

the characteristic curve through (t, x) does not intersect the x–axis, and hence the value of the solution at such points, lying in the shaded region in Figure 2.9, is *not* prescribed by the initial data. Let us arbitrarily assign the solution to be $u(t, x) = 0$ at such points. At other values of (t, x) with $t \ge 0$, the solution (2.23) is

$$u(t, x) = \exp\left[- \left(\frac{x + 1 + (x-1)\, e^{-2t}}{x + 1 - (x-1)\, e^{-2t}} \right)^2 \right]. \tag{2.30}$$

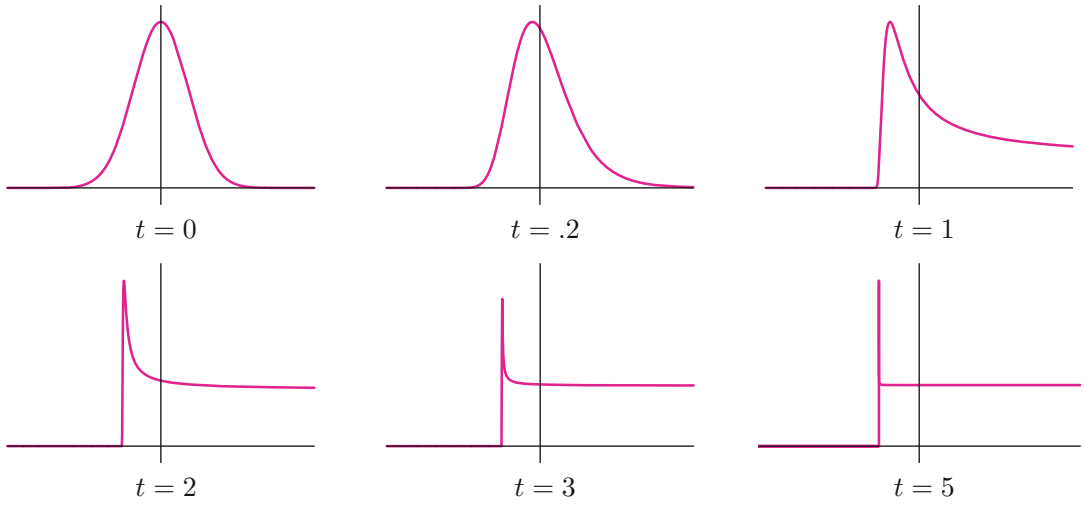


Figure 2.10. Solution to $u_t + (x^2 - 1)u_x = 0$.

(The derivation of this solution formula is left as Exercise 2.2.23.) As t increases, the solution's peak becomes more and more concentrated near $x_- = -1$, while the section of the wave above $x > x_+ = 1$ rapidly spreads out to ∞. In the long term, the solution converges (albeit nonuniformly) to a step function of height $1/e$:

$$u(t,x) \longrightarrow s(x) = \begin{cases} 1/e \approx .367879, & x \geq -1, \\ 0, & x < -1, \end{cases} \qquad \text{as} \qquad t \longrightarrow \infty.$$

Let us finish by making a few general observations concerning the characteristic curves of transport equations whose wave speed $c(x)$ depends only on the position x. Using the basic existence and uniqueness theory for such autonomous ordinary differential equations, [**20**, **23**, **52**], and assuming that $c(x)$ is continuously differentiable:†

- There is a unique characteristic curve passing through each point $(t,x) \in \mathbb{R}^2$.
- Characteristic curves cannot cross each other.
- If $t = \beta(x)$ is a characteristic curve, then so are all its horizontal translates:
  $$t = \beta(x) + k \text{ for any } k.$$
- Each non-horizontal characteristic curve is the graph of a strictly monotone function. Thus, each point on a wave always moves in the same direction, and can never reverse its direction of propagation.
- As t increases, the characteristic curve either tends to a fixed point, $x(t) \to x_\star$ as $t \to \infty$, with $c(x_\star) = 0$, or goes off to $\pm\infty$ in either finite or infinite time.

Proofs of these statements are assigned to the reader in Exercise 2.2.25.

† For those who know about such things, [**18**, **52**], this assumption can be weakened to just Lipschitz continuity.

Exercises

2.2.16. (a) Find the general solution to the first-order equation $u_t + \frac{3}{2}\, u_x = 0$.
(b) Find a solution satisfying the initial condition $u(1, x) = \sin x$. Is your solution unique?

2.2.17. (a) Solve the initial value problem $u_t - x\, u_x = 0$, $u(0, x) = (x^2 + 1)^{-1}$.
(b) Graph the solution at times $t = 0, 1, 2, 3$. (c) What is $\lim_{t \to \infty} u(t, x)$?

2.2.18. Suppose the initial data $u(0, x) = f(x)$ of the nonuniform transport equation (2.28) is continuous and satisfies $f(x) \to 0$ as $|\, x \,| \to \infty$. What is the limiting solution profile $u(t, x)$ as (a) $t \to \infty$? (b) $t \to -\infty$?

♡ 2.2.19. (a) Find and graph the characteristic curves for the equation $u_t + (\sin x)\, u_x = 0$.
(b) Write down the solution with initial data $u(0, x) = \cos \frac{1}{2}\pi x$. (c) Graph your solution at times $t = 0, 1, 2, 3, 5$, and 10. (d) What is the limiting solution profile as $t \to \infty$?

2.2.20. Consider the linear transport equation $u_t + (1 + x^2)\, u_x = 0$. (a) Find and sketch the characteristic curves. (b) Write down a formula for the general solution. (c) Find the solution to the initial value problem $u(0, x) = f(x)$ and discuss its behavior as t increases.

2.2.21. Prove that, for $t \gg 0$, the speed of the wave in Example 2.4 is asymptotically proportional to $t^{-2/3}$.

2.2.22. Verify directly that formula (2.21) defines a solution to the differential equation (2.16).

♢ 2.2.23. Explain how to derive the solution formula (2.30). Justify that it defines a solution to equation (2.28).

2.2.24. Let $c(x)$ be a bounded C^1 function, so $|\, c(x) \,| \leq c_\star < \infty$ for all x. Let $f(x)$ be any C^1 function. Prove that the solution $u(t, x)$ to the initial value problem $u_t + c(x)\, u_x = 0$, $u(0, x) = f(x)$, is uniquely defined for all $(t, x) \in \mathbb{R}^2$.

♡ 2.2.25. Suppose that $c(x) \in \mathrm{C}^1$ is continuously differentiable for all $x \in \mathbb{R}$. (a) Prove that the characteristic curves of the transport equation (2.16) cannot cross each other. (b) A point where $c(x_\star) = 0$ is known as a *fixed point* for the characteristic equation $dx/dt = c(x)$. Explain why the characteristic curve passing through a fixed point $(t, x_\star)$ is a horizontal straight line. (c) Prove that if $x = g(t)$ is a characteristic curve, then so are all the horizontally translated curves $x = g(t + \delta)$ for any δ. (d) *True or false*: Every characteristic curve has the form $x = g(t + \delta)$, for some fixed function $g(t)$. (e) Prove that each non-horizontal characteristic curve is the graph $x = g(t)$ of a strictly monotone function. (f) Explain why a wave cannot reverse its direction. (g) Show that a non-horizontal characteristic curve starts, in the distant past, $t \to -\infty$, at either a fixed point or at $-\infty$ and ends, as $t \to +\infty$, at either the next-larger fixed point or at $+\infty$.

♡ 2.2.26. Consider the transport equation $\dfrac{\partial u}{\partial t} + c(t, x)\, \dfrac{\partial u}{\partial x} = 0$ with time-varying wave speed. Define the corresponding characteristic ordinary differential equation to be $\dfrac{dx}{dt} = c(t, x)$, the graphs of whose solutions $x(t)$ are the *characteristic curves*. (a) Prove that any solution $u(t, x)$ to the partial differential equation is constant on each characteristic curve. (b) Suppose that the general solution to the characteristic equation is written in the form $\xi(t, x) = k$, where k is an arbitrary constant. Prove that $\xi(t, x)$ defines a *characteristic variable*, meaning that $u(t, x) = f(\xi(t, x))$ is a solution to the time-varying transport equation for any continuously differentiable scalar function $f \in \mathrm{C}^1$.

2.2.27. (a) Apply the method in Exercise 2.2.26 to find the characteristic curves for the equation $u_t + t^2\, u_x = 0$. (b) Find the solution to the initial value problem $u(0, x) = e^{-x^2}$, and discuss its dynamic behavior.

2.2.28. Solve Exercise 2.2.27 for the equation $u_t + (x - t)\, u_x = 0$.

♡ 2.2.29. Consider the first-order partial differential equation $u_t + (1 - 2t)\, u_x = 0$. Use Exercise 2.2.26 to: (a) Find and sketch the characteristic curves. (b) Write down the general solution. (c) Solve the initial value problem with $u(0, x) = \dfrac{1}{1 + x^2}$. (d) Describe the behavior of your solution $u(t, x)$ from part (c) as $t \to \infty$. What about $t \to -\infty$?

2.2.30. Discuss which of the conclusions of Exercise 2.2.25 are valid for the characteristic curves of the transport equation with time-varying wave speed, as analyzed in Exercise 2.2.26.

♢ 2.2.31. Consider the two-dimensional transport equation $\dfrac{\partial u}{\partial t} + c(x, y)\, \dfrac{\partial u}{\partial x} + d(x, y)\, \dfrac{\partial u}{\partial y} = 0$, whose solution $u(t, x, y)$ depends on time t and space variables x, y. (a) Define a characteristic curve, and prove that the solution is constant along it. (b) Apply the method of characteristics to solve the initial value problem $u_t + y\, u_x - x\, u_y$, $u(0, x, y) = e^{-(x-1)^2 - (y-1)^2}$. (c) Describe the behavior of your solution.

2.3 Nonlinear Transport and Shocks

The first-order nonlinear partial differential equation

$$u_t + u\, u_x = 0 \tag{2.31}$$

has the form of a transport equation (2.4), but the wave speed $c = u$ now depends, not on the position x, but rather on the size of the disturbance u. Larger waves will move faster, and overtake smaller, slower-moving waves. Waves of elevation, where $u > 0$, move to the right, while waves of depression, where $u < 0$, move to the left. This equation is considerably more challenging than the linear transport models analyzed above, and was first systematically studied in the early nineteenth century by the influential French mathematician Siméon–Denis Poisson and the great German mathematician Bernhard Riemann.† It and its multi-dimensional and multi-component generalizations play a crucial role in the modeling of gas dynamics, acoustics, shock waves in pipes, flood waves in rivers, chromatography, chemical reactions, traffic flow, and so on. Although we will be able to write down a solution formula, the complete analysis is far from trivial, and will require us to confront the possibility of discontinuous shock waves. Motivated readers are referred to Whitham's book, [**122**], for further details.

Fortunately, the method of characteristics that was developed for linear transport equations also works in the present context and leads to a complete mathematical solution. Mimicking our previous construction, (2.18), but now with wave speed $c = u$, let us define a *characteristic curve* of the nonlinear wave equation (2.31) to be the graph of a solution $x(t)$ to the ordinary differential equation

$$\frac{dx}{dt} = u(t, x). \tag{2.32}$$

† In addition to his fundamental contributions to partial differential equations, complex analysis, and number theory, Riemann also was the inventor of Riemannian geometry, which turned out to be absolutely essential for Einstein's theory of general relativity some 70 years later!

As such, the characteristics depend upon the solution u, which, in turn, is to be specified by its characteristics. We appear to be trapped in a circular argument.

The resolution of the conundrum is to argue that, as in the linear case, the solution $u(t,x)$ remains constant along its characteristics, and this fact will allow us to simultaneously specify both. To prove this claim, suppose that $x = x(t)$ parametrizes a characteristic curve associated with the given solution $u(t,x)$. Our task is to show that $h(t) = u\big(t, x(t)\big)$, which is obtained by evaluating the solution along the curve, is constant, which, as usual, is proved by checking that its derivative is identically zero. Repeating our chain rule computation (2.17), and using (2.32), we deduce that

$$\frac{dh}{dt} = \frac{d}{dt}\,u\big(t,x(t)\big) = \frac{\partial u}{\partial t}\big(t,x(t)\big) + \frac{dx}{dt}\,\frac{\partial u}{\partial x}\big(t,x(t)\big) = \frac{\partial u}{\partial t}\big(t,x(t)\big) + u\big(t,x(t)\big)\frac{\partial u}{\partial x}\big(t,x(t)\big) = 0,$$

since u is assumed to solve the nonlinear transport equation (2.31) at all values of (t,x), including those on the characteristic curve. We conclude that $h(t)$ is constant, and hence u is indeed constant on the characteristic curve.

Now comes the clincher. We know that the right-hand side of the characteristic ordinary differential equation (2.32) is a constant whenever $x = x(t)$ defines a characteristic curve. This means that the derivative dx/dt is a constant — namely the fixed value of u on the curve. Therefore, the characteristic curve must be a *straight line*,

$$x = u\,t + k, \tag{2.33}$$

whose slope equals the value assumed by the solution u on it.

And, as before, since the solution is constant along each characteristic line, it must be a function of the *characteristic variable*

$$\xi = x - t\,u \tag{2.34}$$

alone, and so

$$u = f(x - t\,u), \tag{2.35}$$

where $f(\xi)$ is an arbitrary C^1 function. Formula (2.35) should be viewed as an algebraic equation that implicitly defines the solution $u(t,x)$ as a function of t and x. Verification that ther resulting function is indeed a solution to (2.31) is the subject of Exercise 2.3.14.

Example 2.6. Suppose that

$$f(\xi) = \alpha\,\xi + \beta,$$

with α, β constant. Then (2.35) becomes

$$u = \alpha(x - t\,u) + \beta, \qquad \text{and hence} \qquad u(t,x) = \frac{\alpha\,x + \beta}{1 + \alpha\,t} \tag{2.36}$$

is the corresponding solution to the nonlinear transport equation. At each fixed t, the graph of the solution is a straight line. If $\alpha > 0$, the solution flattens out: $u(t,x) \to 0$ as $t \to \infty$. On the other hand, if $\alpha < 0$, the straight line rapidly steepens to vertical as t approaches the critical time $t_\star = -1/\alpha$, at which point the solution ceases to exist. Figure 2.11 graphs two representative solutions. The top row shows the solution with $\alpha = 1$, $\beta = .5$, plotted at times $t = 0, 1, 5$, and 20; the bottom row takes $\alpha = -.2$, $\beta = .1$, and plots the solution at times $t = 0, 3, 4$, and 4.9. In the second case, the solution *blows up* by becoming vertical as $t \to 5$.

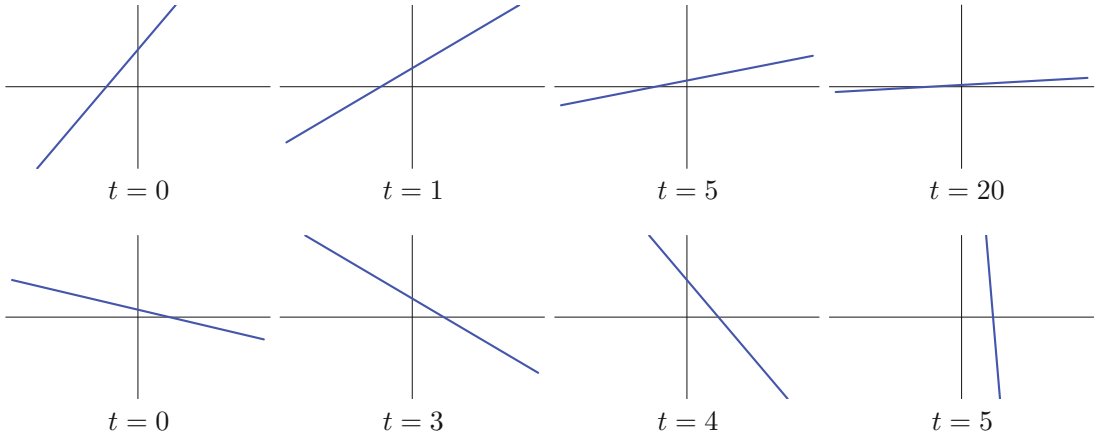


Figure 2.11. Two solutions to $u_t + u\,u_x = 0$.

Remark: Although (2.36) remains a valid solution formula after the blow-up time, $t > 5$, this is *not* to be viewed as a part of the original solution. With the appearance of such a singularity, the physical solution has broken down, and we stop tracking it.

To solve the general initial value problem

$$u(0, x) = f(x), \tag{2.37}$$

we note that, at $t = 0$, the implicit solution formula (2.35) reduces to (2.37), and hence the function f coincides with the initial data. However, because our solution formula (2.35) is an implicit equation, it is not immediately evident

(*a*) whether it can be solved to give a well-defined function $u(t, x)$, and,

(*b*) even granted this, how to describe the resulting solution's qualitative features and dynamical behavior.

A more instructive approach is founded on the following geometrical construction. Through each point $(0, y)$ on the x–axis, draw the characteristic line

$$x = t\,f(y) + y \tag{2.38}$$

whose slope, namely $f(y) = u(0, y)$, equals the value of the initial data (2.37) at that point. According to the preceding discussion, the solution will have the same value on the entire characteristic line (2.38), and so

$$u(t, t\,f(y) + y) = f(y) \qquad \text{for all } \ t. \tag{2.39}$$

For example, if $f(y) = y$, then $u(t, x) = y$ whenever $x = t\,y + y$; eliminating y, we find $u(t, x) = x/(t + 1)$, which agrees with one of our straight line solutions (2.36).

Now, the problem with this construction is immediately apparent from Figure 2.12, which plots the characteristic lines associated with the initial data

$$u(0, x) = \tfrac{1}{2}\pi - \tan^{-1} x.$$

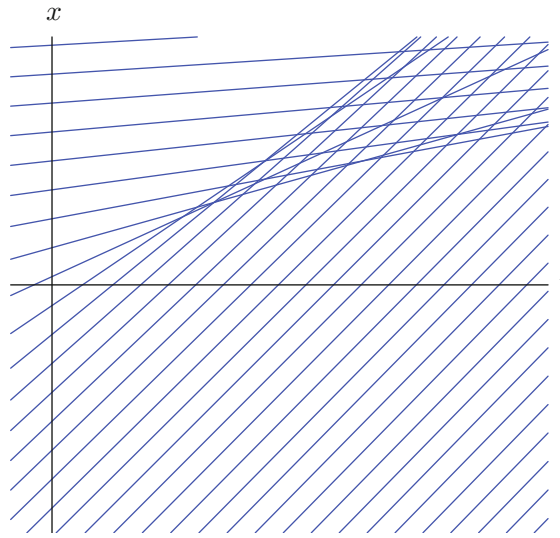


Figure 2.12. Characteristics lines for $u(0,x) = \frac{1}{2}\pi - \tan^{-1} x$.

Two characteristic lines that are not parallel must cross each other somewhere. The value of the solution is supposed to equal the slope of the characteristic line passing through the point. Hence, at a crossing point, the solution is required to assume two *different* values, one corresponding to each line. Something is clearly amiss, and we need to resolve this apparent paradox.

There are three principal scenarios. The first, trivial, situation occurs when all the characteristic lines are parallel, and so the difficulty does not arise. In this case, they all have the same slope, say c, which means that the solution has the same value on each one. Therefore, $u(t,x) \equiv c$ is a constant solution.

The next-simplest case occurs when the initial data is everywhere *nondecreasing*, so $f(x) \le f(y)$ whenever $x \le y$, which is assured if its derivative is never negative: $f'(x) \ge 0$. In this case, as sketched in Figure 2.13, the characteristic lines emanating from the x axis fan out into the right half-plane, and so never cross each other at any future time $t > 0$. Each point (t,x) with $t \ge 0$ lies on a unique characteristic line, and the value of the solution at (t,x) is equal to the slope of the line. We conclude that the solution $u(t,x)$ is well defined at all future times $t \ge 0$. Physically, such solutions represent *rarefaction waves*, which spread out as time progresses. A typical example, corresponding to initial data

$$u(0,x) = \tfrac{1}{2}\pi + \tan^{-1}(3\,x),$$

has its characteristic lines plotted in Figure 2.13, while Figure 2.14 graphs some representative solution profiles.

The more interesting case occurs when the initial data is a decreasing function, and so $f'(x) < 0$. Now, as in Figure 2.12, some of the characteristic lines starting at $t = 0$ will cross at some point in the future. If a point (t,x) lies on two or more distinct characteristic lines, the value of the solution $u(t,x)$, which should equal the characteristic slope, is no longer uniquely determined. Although, in a purely mathematical context, one might be tempted to allow such multiply valued solutions, from a physical standpoint this is unacceptable. The solution $u(t,x)$ is supposed to represent a measurable quantity, e.g., concentration,

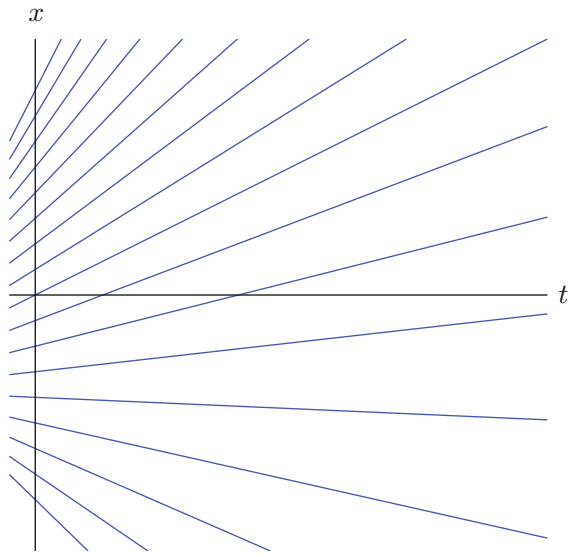


Figure 2.13. Characteristic lines for a rarefaction wave.

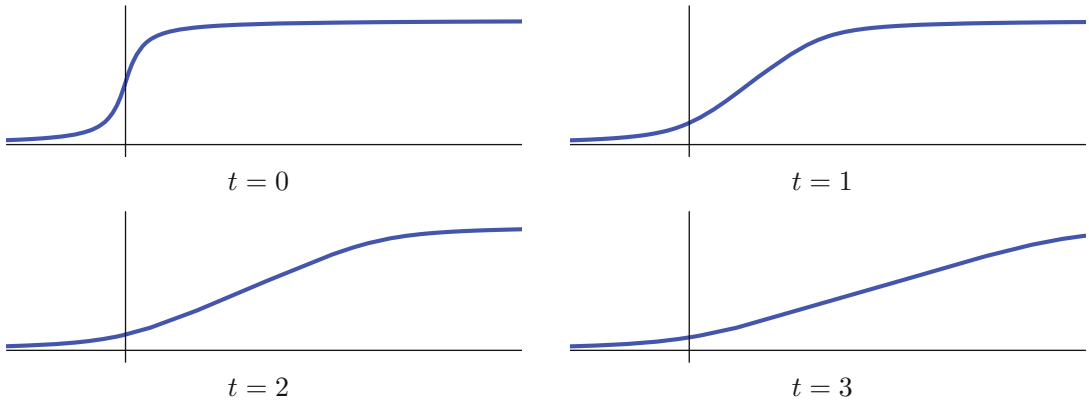


Figure 2.14. Rarefaction wave.

velocity, pressure, and must therefore assume a unique value at each point. In effect, the mathematical model has broken down and no longer conforms to physical reality.

However, before confronting this difficulty, let us first, from a purely theoretical standpoint, try to understand what happens if we mathematically continue the solution as a multiply valued function. For specificity, consider the initial data

$$u(0, x) = \tfrac{1}{2}\pi - \tan^{-1} x, \tag{2.40}$$

appearing in the first graph in Figure 2.15. The corresponding characteristic lines are displayed in Figure 2.12. Initially, they do not cross, and the solution remains a well-defined, single-valued function. However, after a while one reaches a critical time, $t_\star > 0$, when the first two characteristic lines cross each other. Subsequently, a wedge-shaped region appears in the (t, x)–plane, consisting of points that lie on the intersection of three

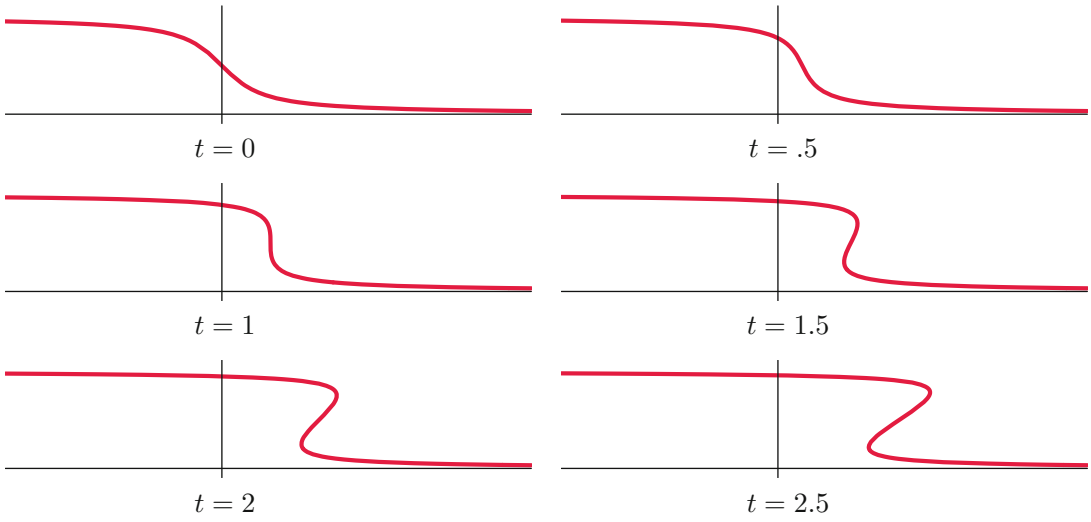


Figure 2.15. Multiply valued compression wave.

distinct characteristic lines with different slopes; at such points, the mathematical solution achieves three distinct values. Points outside the wedge lie on a single characteristic line, and the solution remains single-valued there. The boundary of the wedge consists of points where precisely two characteristic lines cross.

To fully appreciate what is going on, look now at the sequence of pictures of the multiply valued solution in Figure 2.15, plotted at six successive times. Since the initial data is positive, $f(x) > 0$, all the characteristic slopes are positive. As a consequence, every point on the solution curve moves to the right, at a speed equal to its height. Since the initial data is a decreasing function, points on the graph lying to the left will move faster than those to the right and eventually overtake them. At first, the solution merely steepens into a *compression wave*. At the critical time $t_\star$ when the first two characteristic lines cross, say at position $x_\star$, so that $(t_\star, x_\star)$ is the tip of the aforementioned wedge, the solution graph has become vertical:

$$\frac{\partial u}{\partial x}(t, x_\star) \longrightarrow \infty \qquad \text{as} \qquad t \longrightarrow t_\star,$$

and $u(t, x)$ is no longer a classical solution. Once this occurs, the solution graph ceases to be a single-valued function, and its overlapping lobes lie over the points (t, x) belonging to the wedge.

The critical time $t_\star$ can, in fact, be determined from the implicit solution formula (2.35). Indeed, if we differentiate with respect to x, we obtain

$$\frac{\partial u}{\partial x} = \frac{\partial}{\partial x} f(\xi) = f'(\xi) \frac{\partial \xi}{\partial x} = f'(\xi) \left(1 - t \frac{\partial u}{\partial x} \right), \qquad \text{where} \qquad \xi = x - t\, u.$$

Solving for

$$\frac{\partial u}{\partial x} = \frac{f'(\xi)}{1 + t\, f'(\xi)},$$

we see that the slope blows up:

$$\frac{\partial u}{\partial x} \longrightarrow \infty \qquad \text{as} \qquad t \longrightarrow -\frac{1}{f'(\xi)}\,.$$

In other words, if the initial data has negative slope at position x, so $f'(x) < 0$, then the solution along the characteristic line emanating from the point $(0, x)$ will fail to be smooth at the time $-1/f'(x)$. The earliest critical time is, thus,

$$t_\star := \min\left\{ -\frac{1}{f'(x)} \;\middle|\; f'(x) < 0 \right\}. \tag{2.41}$$

If x_0 is the value of x that produces the minimum $t_\star$, then the slope of the solution profile will first become infinite at the location where the characteristic starting at x_0 is at time $t_\star$, namely

$$x_\star = x_0 + f(x_0)\, t_\star. \tag{2.42}$$

For instance, for the particular initial configuration (2.40) represented in Figure 2.15,

$$f(x) = \frac{\pi}{2} - \tan^{-1} x, \qquad\qquad f'(x) = -\frac{1}{1+x^2}\,,$$

and so the critical time is

$$t_\star = \min\{\, 1 + x^2 \,\} = 1, \qquad \text{with} \qquad x_\star = f(0)\, t_\star = \tfrac{1}{2}\,\pi,$$

since the minimum value occurs at $x_0 = 0$.

Now, while mathematically plausible, such a multiply valued solution is physically untenable. So what really happens after the critical time $t_\star$? One needs to decide which (if any) of the possible solution values is physically appropriate. The mathematical model, in and of itself, is incapable of resolving this quandary. We must therefore revisit the underlying physics, and ask what sort of phenomenon we are trying to model.

Shock Dynamics

To be specific, let us regard the transport equation (2.31) as a model of compressible fluid flow in a single space variable, e.g., the motion of gas in a long pipe. If we push a piston into the pipe, then the gas will move ahead of it and thereby be compressed. However, if the piston moves too rapidly, then the gas piles up on top of itself, and a shock wave forms and propagates down the pipe. Mathematically, the shock is represented by a discontinuity where the solution abruptly changes value. The formulas (2.41) and (2.42) determine the time and position for the onset of the shock-wave discontinuity. Our goal now is to predict its subsequent behavior, and this will be based on use of a suitable physical conservation law. Indeed, one expects mass to be conserved – even through a shock discontinuity — since gas atoms can neither be created nor destroyed. And, as we will see, conservation of mass (almost) suffices to prescribe the subsequent motion of the shock wave.

Before investigating the implications of conservation of mass, let us first convince ourselves of its validity for the nonlinear transport model. (Just because a mathematical equation models a physical system does not automatically imply that it inherits any of its

physical conservation laws.) If $u(t,x)$ represents density, then, at time t, the total mass lying in an interval $a \le x \le b$ is calculated by integration:

$$M_{a,b}(t) = \int_a^b u(t,x)\,dx. \tag{2.43}$$

Assuming that $u(t,x)$ is a classical solution to the nonlinear transport equation (2.31), we can determine the rate of change of mass on this interval by differentiation:

$$\begin{aligned} \frac{dM_{a,b}}{dt} &= \frac{d}{dt}\int_a^b u(t,x)\,dx = \int_a^b \frac{\partial u}{\partial t}(t,x)\,dx = -\int_a^b u(t,x)\,\frac{\partial u}{\partial x}(t,x)\,dx \\ &= -\int_a^b \frac{\partial}{\partial x}\left[\,\tfrac{1}{2}\,u(t,x)^2\,\right]dx = -\tfrac{1}{2}\,u(t,x)^2\,\Big|_{x=a}^{b} = \tfrac{1}{2}\,u(t,a)^2 - \tfrac{1}{2}\,u(t,b)^2. \end{aligned} \tag{2.44}$$

The final expression represents the net *mass flux* through the endpoints of the interval. Thus, the only way in which the mass on the interval $[a,b]$ changes is through its endpoints; inside, mass can be neither created nor destroyed, which is the precise meaning of the mass conservation law in continuum mechanics. In particular, if there is zero net mass flux, then the total mass is constant, and hence conserved. For example, if the initial data (2.37) has finite total mass,

$$\left| \int_{-\infty}^{\infty} f(x)\,dx \right| < \infty, \tag{2.45}$$

which requires that $f(x) \to 0$ reasonably rapidly as $|\,x\,| \to \infty$, then the total mass of the solution — at least up to the formation of a shock discontinuity — remains constant and equal to its initial value:

$$\int_{-\infty}^{\infty} u(t,x)\,dx = \int_{-\infty}^{\infty} u(0,x)\,dx = \int_{-\infty}^{\infty} f(x)\,dx. \tag{2.46}$$

Similarly, if $u(t,x)$ represents the traffic density on a highway at time t and position x, then the integrated conservation law (2.44) tells us that the rate of change in the number of vehicles on the stretch of road between a and b equals the number of vehicles entering at point a minus the number leaving at point b — which assumes that there are no other exits or entrances on this part of the highway. Thus, in the traffic model, (2.44) represents the conservation of vehicles.

The preceding calculation relied on the fact that the integrand can be written as an x derivative. This is a common feature of physical conservation laws in continuum mechanics, and motivates the following general definition.

Definition 2.7. A *conservation law*, in one space dimension, is an equation of the form

$$\frac{\partial T}{\partial t} + \frac{\partial X}{\partial x} = 0. \tag{2.47}$$

The function T is known as the *conserved density*, while X is the associated *flux*.

In the simplest situations, the conserved density $T(t,x,u)$ and flux $X(t,x,u)$ depend on the time t, the position x, and the solution $u(t,x)$ to the physical system. (Higher-order conservation laws, which also depend on derivatives of u, arise in the analysis of integrable partial differential equations; see Section 8.5 and [**36**, **87**].) For example, the nonlinear transport equation (2.31) is itself a conservation law, since it can be written in the form

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}\left(\,\tfrac{1}{2}\,u^2\,\right) = 0, \tag{2.48}$$

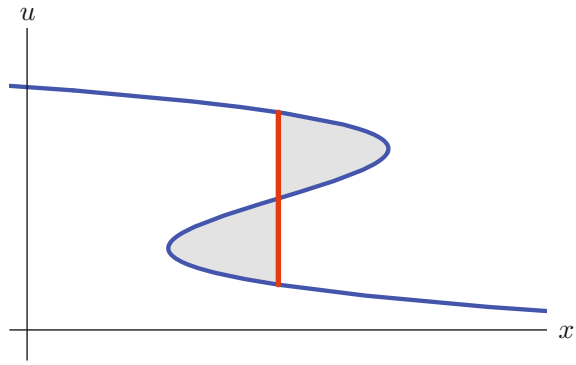


Figure 2.16. Equal Area Rule.

and so the conserved density is $T = u$ and the flux is $X = \frac{1}{2}u^2$. And indeed, it was this identity that made our computation (2.44) work. The general result, proved by an analogous computation, justifies calling (2.47) a conservation law.

Proposition 2.8. *Given a conservation law* (2.47), *then, on any closed interval* $a \le x \le b$,

$$\frac{d}{dt}\int_a^b T\,dx = \,-\,X\,\Big|_{x=a}^{b}. \tag{2.49}$$

Proof: The proof is an immediate consequence of the Fundamental Theorem of Calculus — assuming sufficient smoothness that allows one to bring the derivative inside the integral sign:

$$\frac{d}{dt}\int_a^b T\,dx = \int_a^b \frac{\partial T}{\partial t}\,dx = -\int_a^b \frac{\partial X}{\partial x}\,dx = \,-\,X\,\Big|_{x=a}^{b}. \qquad Q.E.D.$$

We will refer to (2.49) as the *integrated form* of the conservation law (2.47). It states that the rate of change of the total density, integrated over an interval, is equal to the amount of flux through its two endpoints. In particular, if there is no net flux into or out of the interval, then the integrated density is *conserved*, meaning that it remains constant over time. All physical conservation laws — mass, momentum, energy, and so on — for systems governed by partial differential equations are of this form or its multi-dimensional extensions, [**87**].

With this in hand, let us return to the physical context of the nonlinear transport equation. By definition, a *shock* is a discontinuity in the solution $u(t, x)$. We will make the physically plausible assumption that mass (or vehicle) conservation continues to hold even within the shock. Recall that the total mass, which at time t is the area[†] under the curve $u(t, x)$, must be conserved. This continues to hold even when the mathematical solution becomes multiply valued, in which case one employs a line integral $\int_C u\,dx$, where C represents the graph of the solution, to compute the mass/area. Thus, to construct a discontinuous shock solution with the *same* mass, one replaces part of the multiply valued

[†] We are implicitly assuming that the mass is finite, as in (2.45), although the overall construction does not rely on this restriction.

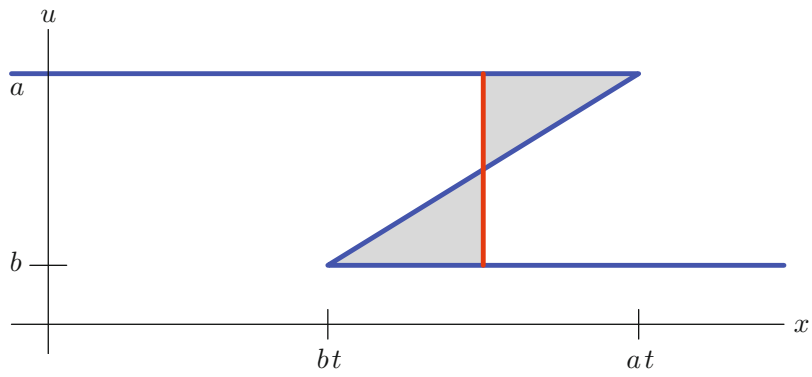


Figure 2.17. Multiply–valued step wave.

graph by a vertical shock line in such a way that the resulting function is single-valued and has the same area under its graph. Referring to Figure 2.16, observe that the region under the shock graph is obtained from that under the multi-valued solution graph by deleting the upper shaded lobe and appending the lower shaded lobe. Thus the resulting area will be the same, provided the shock line is drawn so that the areas of the two shaded lobes are equal. This construction is known as the *Equal Area Rule*; it ensures that the total mass of the shock solution matches that of the multiply valued solution, which in turn is equal to the initial mass, as required by the physical conservation law.

Example 2.9. An illuminating special case occurs when the initial data has the form of a *step function* with a single discontinuity at the origin:

$$u(0,x) = \begin{cases} a, & x < 0, \\ b, & x > 0. \end{cases} \tag{2.50}$$

If $a > b$, then the initial data is already in the form of a shock wave. For $t > 0$, the mathematical solution constructed by continuing along the characteristic lines is multiply valued in the region $bt < x < at$, where it assumes both values a and b; see Figure 2.17. Moreover, the initial vertical line of discontinuity has become a tilted line, because each point $(0, u)$ on it has moved along the associated characteristic a distance ut. The Equal Area Rule tells us to draw the shock line halfway along, at $x = \frac{1}{2}(a+b)t$, in order that the two triangles have the same area. We deduce that the shock moves with speed $c = \frac{1}{2}(a+b)$, equal to the average of the two speeds at the jump. The resulting shock-wave solution is

$$u(t,x) = \begin{cases} a, & x < ct, \\ b, & x > ct, \end{cases} \qquad \text{where} \qquad c = \frac{a+b}{2}. \tag{2.51}$$

A plot of its characteristic lines appears in Figure 2.18. Observe that colliding pairs of characteristic lines terminate at the shock line, whose slope is the average of their individual slopes.

The fact that the shock speed equals the *average* of the solution values on either side is, in fact, of general validity, and is known as the *Rankine–Hugoniot condition*, named after the nineteenth-century Scottish physicist William Rankine and French engineer Pierre Hugoniot, although historically these conditions first appeared in a 1849 paper by George Stokes, [**109**]. However, intimidated by criticism by his contemporary applied mathematicians Lords Kelvin and Rayleigh, Stokes thought he was mistaken, and even ended up

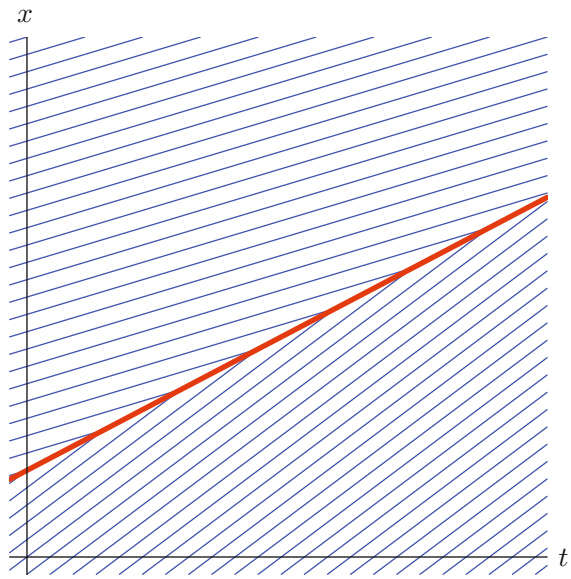


Figure 2.18. Characteristic lines for the step wave shock.

deleting the relevant part when his collected works were published in 1883, [**110**]. The missing section was restored in the 1966 reissue, [**111**].

Proposition 2.10. *Let $u(t,x)$ be a solution to the nonlinear transport equation that has a discontinuity at position $x = \sigma(t)$, with finite, unequal left- and right-hand limits*

$$u^-(t) = u\big(t,\sigma(t)^-\big) = \lim_{x \to \sigma(t)^-} u(t,x), \qquad u^+(t) = u\big(t,\sigma(t)^+\big) = \lim_{x \to \sigma(t)^+} u(t,x), \tag{2.52}$$

on either side of the shock discontinuity. Then, to maintain conservation of mass, the speed of the shock must equal the average of the solution values on either side:

$$\frac{d\sigma}{dt} = \frac{u^-(t) + u^+(t)}{2}\,. \tag{2.53}$$

Proof: Referring to Figure 2.19, consider a small time interval, from t to $t + \Delta t$, with $\Delta t > 0$. During this time, the shock moves from position $a = \sigma(t)$ to position $b = \sigma(t + \Delta t)$. The total mass contained in the interval $[a, b]$ at time t, before the shock has passed through, is

$$M(t) = \int_a^b u(t,x)\,dx \approx u^+(t)\,(b-a) = u^+(t)\,\big[\,\sigma(t+\Delta t) - \sigma(t)\,\big],$$

where we assume that $\Delta t \ll 1$ is very small, and so the integrand is well approximated by its limiting value (2.52). Similarly, after the shock has passed, the total mass remaining in the interval is

$$M(t+\Delta t) = \int_a^b u(t+\Delta t,x)\,dx \approx u^-(t+\Delta t)\,(b-a) = u^-(t+\Delta t)\,\big[\,\sigma(t+\Delta t) - \sigma(t)\,\big].$$

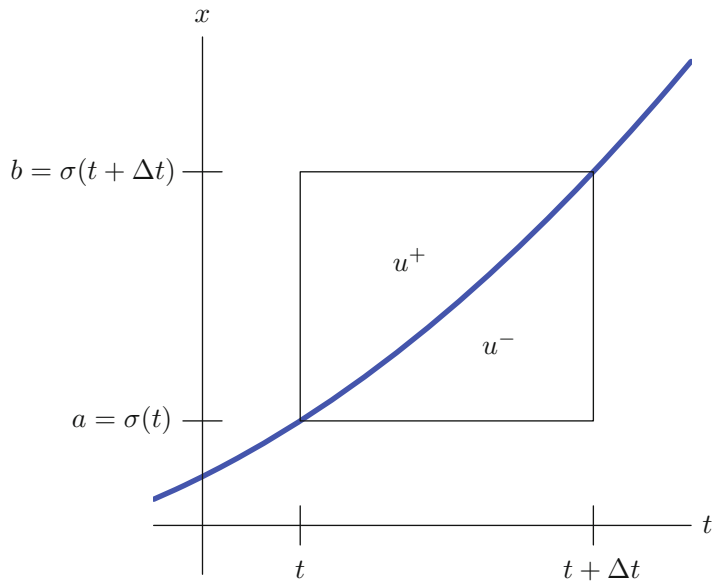


Figure 2.19. Conservation of mass near a shock.

Thus, the rate of change in mass across the shock at time t is given by

$$\frac{dM}{dt} = \lim_{\Delta t \to 0} \frac{M(t+\Delta t) - M(t)}{\Delta t}$$
$$= \lim_{\Delta t \to 0} \big[\, u^-(t+\Delta t) - u^+(t) \,\big] \, \frac{\sigma(t+\Delta t) - \sigma(t)}{\Delta t} = \big[\, u^-(t) - u^+(t) \,\big] \, \frac{d\sigma}{dt} \, .$$

On the other hand, at any $t < \tau < t + \Delta t$, the mass flux into the interval $[\, a, b \,]$ through the endpoints is given by the right-hand side of (2.44):

$$\tfrac{1}{2} \big[\, u(\tau, a)^2 - u(\tau, b)^2 \,\big] \;\longrightarrow\; \tfrac{1}{2} \big[\, u^-(t)^2 - u^+(t)^2 \,\big], \qquad \text{since } \; \tau \to t \;\; \text{as} \;\; \Delta t \to 0.$$

Conservation of mass requires that the rate of change in mass be equal to the mass flux:

$$\frac{dM}{dt} = \big[\, u^-(t) - u^+(t) \,\big] \, \frac{d\sigma}{dt} = \tfrac{1}{2} \big[\, u^-(t)^2 - u^+(t)^2 \,\big] \, .$$

Solving for $d\sigma/dt$ establishes (2.53). *Q.E.D.*

Example 2.11. By way of contrast, let us investigate the case when the initial data is a step function (2.50), but with $a < b$, so the jump goes upwards. In this case, the characteristic lines diverge from the initial discontinuity, and the mathematical solution is not specified at all in the wedge-shaped region $a\,t < x < b\,t$. Our task is to decide how to "fill in" the solution values between the two regions where the solution is well defined and constant.

One possible connection is by a straight line. Indeed, a simple modification of the rational solution (2.36) produces the *similarity solution*†

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

† See Section 8.2 for general techniques for constructing similarity (scale-invariant) solutions to partial differential equations.

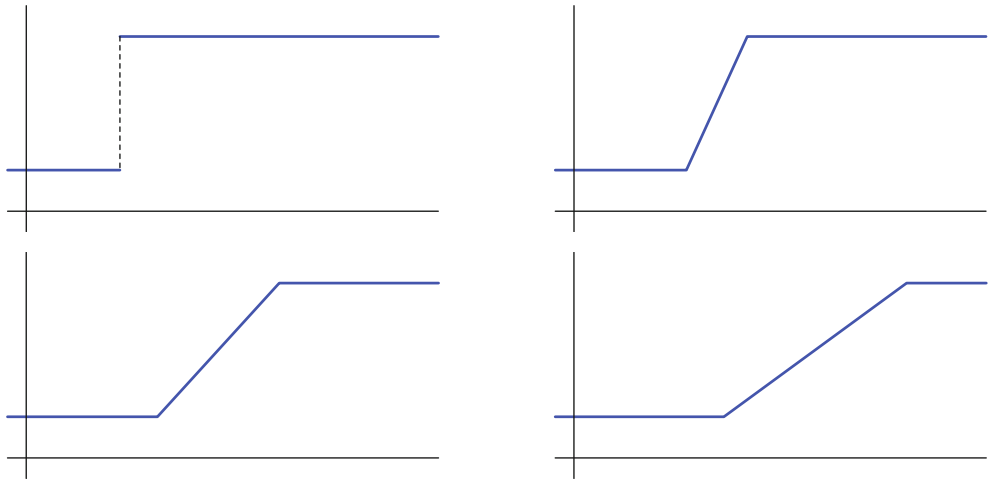


Figure 2.20. Rarefaction wave.

which not only solves the differential equation, but also has the required values $u(t, at) = a$ and $u(t, bt) = b$ at the two edges of the wedge. This can be used to construct the piecewise affine *rarefaction wave*

$$u(t,x) = \begin{cases} a, & x \le at, \\ x/t, & at \le x \le bt, \\ b, & x \ge bt, \end{cases} \tag{2.54}$$

which is graphed at four representative times in Figure 2.20.

A second possibility would be to continue the discontinuity as a shock wave, whose speed is governed by the Rankine-Hugoniot condition, leading to a discontinuous solution having the same formula as (2.51). Which of the two competing solutions should we use? The first, (2.54), makes better physical sense; indeed, if we were to smooth out the discontinuity, then the resulting solutions would converge to the rarefaction wave and not the reverse shock wave; see Exercise 2.3.13. Moreover, the discontinuous solution (2.51) has characteristic lines emanating from the discontinuity, which means that the shock is creating new values for the solution as it moves along, and this can, in fact, be done in a variety of ways. In other words, the discontinuous solution violates *causality*, meaning that the solution profile at any given time uniquely prescribes its subsequent motion. Causality requires that, while characteristics may terminate at a shock discontinuity, they cannot begin there, because their slopes will not be uniquely prescribed by the shock profile, and hence the characteristics to the left of the shock must have larger slope (or speed), while those to the right must have smaller slope. Since the shock speed is the average of the two characteristic slopes, this requires the *Entropy Condition*

$$u^-(t) > \frac{d\sigma}{dt} = \frac{u^-(t) + u^+(t)}{2} > u^+(t). \tag{2.55}$$

With further analysis, it can be shown, [**57**], that the rarefaction wave (2.54) is the unique solution† to the initial value problem satisfying the entropy condition (2.55).

† Albeit not a classical solution, but rather a weak solution, as per Section 10.4.

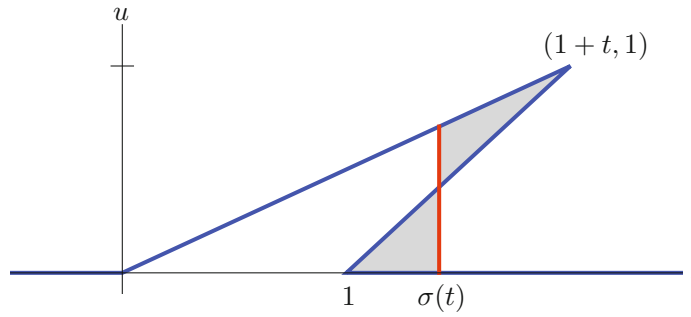


Figure 2.21. Equal Area Rule for the triangular wave.

These prototypical solutions epitomize the basic phenomena modeled by the nonlinear transport equation: *rarefaction waves*, which emanate from regions where the initial data satisfies $f'(x) > 0$, causing the solution to spread out as time progresses, and *compression waves*, emanting from regions where $f'(x) < 0$, causing the solution to progressively steepen and eventually break into a shock discontinuity. Anyone caught in a traffic jam recognizes the compression waves, where the vehicles are bunched together and almost stationary, while the interspersed rarefaction waves correspond to freely moving traffic. (An intelligent driver will take advantage of the rarefaction waves moving backwards through the jam to switch lanes!) The familiar, frustrating traffic jam phenomenon, even on accident- or construction-free stretches of highway, is, thus, an intrinsic effect of the nonlinear transport models that govern traffic flow, [**122**].

Example 2.12. *Triangular wave*: Suppose the initial data has the triangular profile

$$u(0,x) = f(x) = \begin{cases} x, & 0 \le x \le 1, \\ 0, & \text{otherwise}, \end{cases}$$

as in the first graph in Figure 2.22. The initial discontinuity at $x = 1$ will propagate as a shock wave, while the slanted line behaves as a rarefaction wave. To find the profile at time t, we first graph the multi-valued solution obtained by moving each point on the graph of f to the right an amount equal to t times its height. As noted above, this motion preserves straight lines. Thus, points on the x–axis remain fixed, and the diagonal line now goes from $(0,0)$ to $(1+t,1)$, which is where the uppermost point $(1,1)$ on the graph of f has moved to, and hence has slope $(1+t)^{-1}$, while the initial vertical shock line has become tilted, going from $(1,0)$ to $(0,1+t)$. We now need to find the position $\sigma(t)$ of the shock line in order to satisfy the Equal Area Rule, namely so that the areas of the two shaded regions in Figure 2.21 are identical. The reader is invited to determine this geometrically; instead, we invoke the Rankine–Hugoniot condition (2.53). At the shock line, $x = \sigma(t)$, the left- and right-hand limiting values are, respectively,

$$u^-(t) = u\big(t, \sigma(t)^-\big) = \frac{\sigma(t)}{1+t}, \qquad u^+(t) = u\big(t, \sigma(t)^+\big) = 0,$$

and hence (2.53) prescribes the shock speed to be

$$\frac{d\sigma}{dt} = \frac{1}{2}\left(\frac{\sigma(t)}{1+t} + 0\right) = \frac{\sigma(t)}{2\,(1+t)}\,.$$

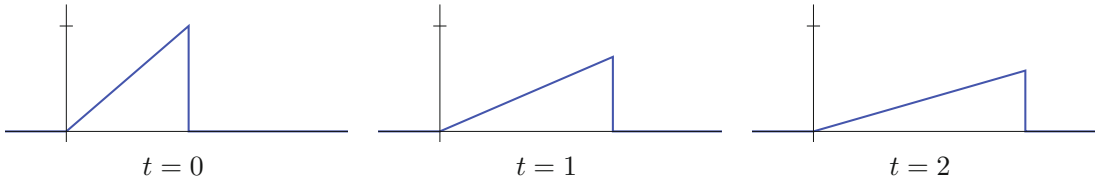


Figure 2.22. Triangular-wave solution.

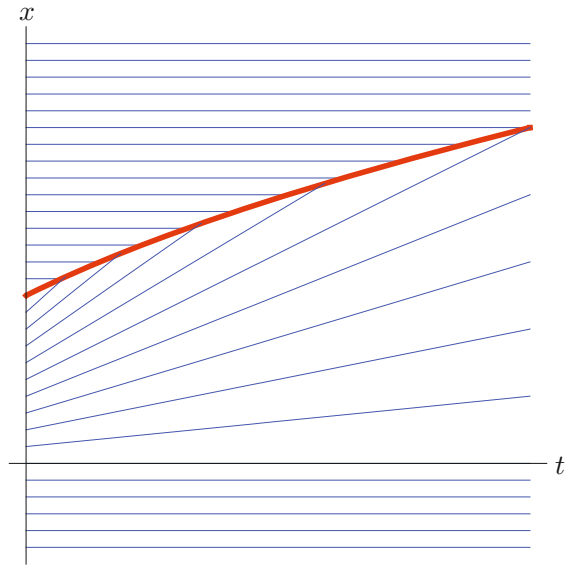


Figure 2.23. Characteristic lines for the triangular-wave shock.

The solution to the resulting separable ordinary differential equation is easily found. Since the shock starts out at $\sigma(0) = 1$, we deduce that

$$\sigma(t) = \sqrt{1+t}\,, \qquad \text{with} \qquad \frac{d\sigma}{dt} = \frac{1}{2\sqrt{1+t}}\,.$$

Further, the strength of the shock, namely its height, is

$$u^-(t) = \frac{\sigma(t)}{1+t} = \frac{1}{\sqrt{1+t}}\,.$$

We conclude that, as t increases, the solution remains a triangular wave, of steadily decreasing slope, while the shock moves off to $x = +\infty$ at a progressively slower speed and smaller height. Its position follows a parabolic trajectory in the (t, x)–plane. See Figure 2.22 for representative plots of the triangular-wave solution, while Figure 2.23 illustrates the characteristic lines and shock-wave trajectory.

In more general situations, continuing on after the initial shock formation, other characteristic lines may start to cross, thereby producing new shocks. The shocks themselves continue to propagate, often at different velocities. When a fast-moving shock catches up

with a slow-moving shock, one must then decide how to merge the shocks so as to retain a physically meaningful solution. The Rankine–Hugoniot (Equal Area) and Entropy Conditions continue to uniquely specify the dynamics. However, at this point, the mathematical details have become too intricate for us to pursue any further, and we refer the interested reader to Whitham's book, [**122**]. See also [**57**] for a proof of the following existence theorem for shock-wave solutions to the nonlinear transport equation.

Theorem 2.13. *If the initial data $u(0,x) = f(x)$ is piecewise*† C^1 *with finitely many jump discontinuities, then, for $t > 0$, there exists a unique (weak) solution to the nonlinear transport equation* (2.31) *that also satisfies the Rankine–Hugoniot condition* (2.53) *and the entropy condition* (2.55).

Remark: Our derivation of the Rankine–Hugoniot shock speed condition (2.53) relied on the fact that we can write the original partial differential equation in the form of a conservation law. But there are, in fact, other ways to do this. For instance, multiplying the nonlinear transport equation (2.31) by u allows us write it in the alternative conservative form

$$u\,\frac{\partial u}{\partial t} + u^2\,\frac{\partial u}{\partial x} = \frac{\partial}{\partial t}\big(\tfrac{1}{2}\,u^2\big) + \frac{\partial}{\partial x}\big(\tfrac{1}{3}\,u^3\big) = 0. \tag{2.56}$$

In this formulation, the conserved density is $T = \frac{1}{2}\,u^2$, and the associated flux is $X = \frac{1}{3}\,u^3$. The integrated form (2.49) of the conservation law (2.56) is

$$\frac{d}{dt}\int_a^b \tfrac{1}{2}\,u(t,x)^2\,dx = \tfrac{1}{3}\big[\,u(t,a)^3 - u(t,b)^3\,\big]. \tag{2.57}$$

In some physical models, the integral on the left-hand side represents the energy within the interval $[a,b]$, and the conservation law tells us that energy can enter the interval as a flux only through its ends. If we assume that energy is conserved at a shock, then, repeating our previous argument, we are led to the alternative equation

$$\frac{d\sigma}{dt} = \frac{\frac{1}{3}\big[\,u^-(t)^3 - u^+(t)^3\,\big]}{\frac{1}{2}\big[\,u^-(t)^2 - u^+(t)^2\,\big]} = \frac{2}{3}\,\frac{u^-(t)^2 + u^-(t)\,u^+(t) + u^+(t)^2}{u^-(t) + u^+(t)} \tag{2.58}$$

for the shock speed. Thus, a shock that conserves energy moves at a different speed from one that conserves mass! The evolution of a shock wave depends not just on the underlying differential equation, but also on the physical assumptions governing the selection of a suitable conservation law.

More General Wave Speeds

Let us finish this section by considering a nonlinear transport equation

$$u_t + c(u)\,u_x = 0, \tag{2.59}$$

whose wave speed is a more general function of the disturbance u. (Further extensions, allowing c to depend also on t and x, are discussed in Exercise 2.3.20.) Most of the

† Meaning continuous everywhere, and continuously differentiable except at a discrete set of points; see Definition 3.7 below for the precise definition.

development is directly parallel to the special case (2.31) discussed above, and so the details are left for the reader to fill in, although the shock dynamics does require some care.

In this case, the *characteristic curve* equation is

$$\frac{dx}{dt} = c\big(u(t,x)\big). \tag{2.60}$$

As before, the solution u is constant on characteristics, and hence the characteristics are straight lines, now with slope $c(u)$. Thus, to solve the initial value problem

$$u(0,x) = f(x), \tag{2.61}$$

through each point $(0,y)$ on the x–axis, one draws the characteristic line of slope $c(u(0,y)) = c(f(y))$. Until the onset of a shock discontinuity, the solution maintains its initial value $u(0,y) = f(y)$ along the characteristic line.

A shock forms whenever two characteristic lines cross. As before, the mathematical equation no longer uniquely specifies the subsequent dynamics, and we need to appeal to an appropriate conservation law. We write the transport equation in the form

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} C(u) = 0, \qquad \text{where} \qquad C(u) = \int c(u)\, du \tag{2.62}$$

is any convenient anti-derivative of the wave speed. Thus, following the same computation as in (2.44), we discover that conservation of mass now takes the integrated form

$$\frac{d}{dt} \int_a^b u(t,x)\, dx = C(u(t,a)) - C(u(t,b)), \tag{2.63}$$

with $C(u)$ playing the role of the mass flux. Requiring the conservation of mass, i.e., of the area under the graph of the solution, means that the Equal Area Rule remains valid. However, the Rankine–Hugoniot shock-speed condition must be modified in accordance with the new dynamics. Mimicking the preceding argument, but with the modified mass flux, we find that the shock speed is now given by

$$\frac{d\sigma}{dt} = \frac{C(u^-(t)) - C(u^+(t))}{u^-(t) - u^+(t)}\,. \tag{2.64}$$

Note that if

$$c(u) = u, \qquad \text{then} \qquad C(u) = \int u\, du = \tfrac{1}{2}\, u^2,$$

and so (2.64) reduces to our earlier formula (2.53). Moreover, in the limit as the shock magnitude approaches zero, $u^-(t) - u^+(t) \to 0$, the right-hand side of (2.64) converges to the derivative $C'(u) = c(u)$ and hence recovers the wave speed, as it should.

Exercises

2.3.1. Discuss the behavior of the solution to the nonlinear transport equation (2.31) for the following initial data:

(a) $u(0,x) = \begin{cases} 2, & x < -1, \\ 1, & x > -1; \end{cases}$ (b) $u(0,x) = \begin{cases} -2, & x < -1, \\ 1, & x > -1; \end{cases}$ (c) $u(0,x) = \begin{cases} 1, & x < 1, \\ -2, & x > 1. \end{cases}$

2.3.2. Solve the following initial value problems: (a) $u_t + 3u u_x = 0$, $u(0,x) = \begin{cases} 2, & x < 1, \\ 0, & x > 1; \end{cases}$

(b) $u_t - u u_x = 0$, $u(1,x) = \begin{cases} -1, & x < 0, \\ 3, & x > 0; \end{cases}$ (c) $u_t - 2u u_x = 0$, $u(0,x) = \begin{cases} 1, & x < 1, \\ 0, & x > 1. \end{cases}$

2.3.3. Let $u(0,x) = (x^2+1)^{-1}$. Does the resulting solution to the nonlinear transport equation (2.31) produce a shock wave? If so, find the time of onset of the shock, and sketch a graph of the solution just before and soon after the shock wave. If not, explain what happens to the solution as t increases.

2.3.4. Solve Exercise 2.3.3 when $u(0,x) =$ (a) $-(x^2+1)^{-1}$, (b) $x(x^2+1)^{-1}$.

2.3.5. Consider the initial value problem $u_t - 2u u_x = 0$, $u(0,x) = e^{-x^2}$. Does the resulting solution produce a shock wave? If so, find the time of onset of the shock and the position at which it first forms. If not, explain what happens to the solution as t increases.

2.3.6. (a) For what values of $\alpha, \beta, \gamma, \delta$ is $u(t,x) = \dfrac{\alpha x + \beta}{\gamma t + \delta}$ a solution to (2.31)?

(b) For what values of $\alpha, \beta, \gamma, \delta, \lambda, \mu$ is $u(t,x) = \dfrac{\lambda t + \alpha x + \beta}{\gamma t + \mu x + \delta}$ a solution to (2.31)?

2.3.7. A *triangular wave* is a shock-wave solution to the initial value problem for (2.31) that has initial data $u(0,x) = \begin{cases} mx, & 0 \le x \le \ell, \\ 0, & \text{otherwise.} \end{cases}$ Assuming $m > 0$, write down a formula for the triangular-wave solution at times $t > 0$. Discuss what happens to the triangular wave as time progresses.

2.3.8. Solve Exercise 2.3.7 when $m < 0$.

2.3.9. Solve (2.31) for $t > 0$ subject to the following initial conditions, and graph your solution at some representative times. In what sense does your solution conserve mass?

(a) $u(0,x) = \begin{cases} 1, & 0 < x < 1, \\ 0, & \text{otherwise,} \end{cases}$ (b) $u(0,x) = \begin{cases} x, & -1 < x < 1, \\ 0, & \text{otherwise,} \end{cases}$

(c) $u(0,x) = \begin{cases} -x, & -1 < x < 1, \\ 0, & \text{otherwise,} \end{cases}$ (d) $u(0,x) = \begin{cases} 1 - |x|, & -1 < x < 1, \\ 0, & \text{otherwise.} \end{cases}$

2.3.10. An *N–wave* is a solution to the nonlinear transport equation (2.31) that has initial conditions $u(0,x) = \begin{cases} mx, & -\ell \le x \le \ell, \\ 0, & \text{otherwise,} \end{cases}$ where $m > 0$. (a) Write down a formula for the N–wave solution at times $t > 0$. (b) What about when $m < 0$?

♢ 2.3.11. Suppose $u(t,x)$ and $\widetilde{u}(t,x)$ are two solutions to the nonlinear transport equation (2.31) such that, for some $t_\star > 0$, they agree: $u(t_\star, x) = \widetilde{u}(t_\star, x)$ for all x. Do the solutions necessarily have the same initial conditions: $u(0,x) = \widetilde{u}(0,x)$? Use your answer to discuss the uniqueness of solutions to the nonlinear transport equation.

2.3.12. Suppose that $x_1 < x_2$ are such that the characteristic lines of (2.31) through $(0, x_1)$ and $(0, x_2)$ cross at a shock at $(t, \sigma(t))$ and, moreover, the left- and right-hand shock values (2.52) are $f(x_1) = u^-(t)$, $f(x_1) = u^+(t)$. Explain why the signed area of the region between the graph of $f(x)$ and the secant line connecting $(x_1, f(x_1))$ to $(x_2, f(x_2))$ is zero.

♢ 2.3.13. Consider the initial value problem $u^\varepsilon(0,x) = 2 + \tan^{-1}(x/\varepsilon)$ for the nonlinear transport equation (2.31). (a) Show that, as $\varepsilon \to 0^+$, the initial condition converges to a step function (2.51). What are the values of a, b? (b) Show that, moreover, the resulting solution $u^\varepsilon(0,x)$ to the nonlinear transport equation converges to the corresponding rarefaction wave (2.54) resulting from the limiting initial condition.

♢ 2.3.14. (*a*) Under what conditions can equation (2.35) be solved for a single-valued function $u(t,x)$? *Hint*: Use the Implicit Function Theorem. (*b*) Use implicit differentiation to prove that the resulting function $u(t,x)$ is a solution to the nonlinear transport equation.

2.3.15. For what values of $\alpha, \beta, \gamma, \delta, k$ is $u(t,x) = \left(\dfrac{\alpha\, x + \beta}{\gamma\, t + \delta} \right)^k$ a solution to the transport equation $u_t + u^2\, u_x = 0$?

2.3.16. (*a*) Solve the initial value problem $u_t + u^2\, u_x = 0$, $u(0,x) = f(x)$, by the method of characteristics. (*b*) Discuss the behavior of solutions and compare/contrast with (2.31).

2.3.17. (*a*) Determine the Rankine–Hugoniot condition, based on conservation of mass, for the speed of a shock for the equation $u_t + u^2\, u_x = 0$. (*b*) Solve the initial value problem $u(0,x) = \begin{cases} a, & x < 0, \\ b, & x > 0, \end{cases}$ when (*i*) $|\,a\,| > |\,b\,|$, (*ii*) $|\,a\,| < |\,b\,|$. *Hint*: Use Exercise 2.3.15 to determine the shape of a rarefaction wave.

2.3.18. Solve Exercise 2.3.17 when the wave speed $c(u) =$ (*i*) $1 - 2\,u$, (*ii*) u^3, (*iii*) $\sin u$.

♢ 2.3.19. Justify the shock-speed formula (2.58).

♢ 2.3.20. Consider the general quasilinear first-order partial differential equation

$$\frac{\partial u}{\partial t} + c(t,x,u)\, \frac{\partial u}{\partial x} = h(t,x,u).$$

Let us define a *lifted characteristic curve* to be a solution $(t, x(t), u(t))$ to the system of ordinary differential equations $\dfrac{dx}{dt} = c(t,x,u)$, $\dfrac{du}{dt} = h(t,x,u)$. The corresponding *characteristic curve* $\big(t, x(t)\big)$ is obtained by projecting to the (t,x)–plane. Prove that if $u(t,x)$ is a solution to the partial differential equation, and $u(t_0,x_0) = u_0$, then the lifted characteristic curve passing through (t_0, x_0, u_0) lies on the graph of $u(t,x)$. Conclude that the graph of the solution to the initial value problem $u(t_0,x) = f(x)$ is the union of all lifted characteristic curves passing through the initial data points $\big(t_0, x_0, f(x_0)\big)$.

2.3.21. Let $a > 0$. (*a*) Apply the method of Exercise 2.3.20 to solve the initial value problem for the *damped transport equation*: $u_t + u\, u_x + a\, u = 0$, $u(0,x) = f(x)$.
(*b*) Does the damping eliminate shocks?

2.3.22. Apply the method of Exercise 2.3.20 to solve the initial value problem

$$u_t + t\, u_x = u^2, \qquad u(0,x) = \frac{1}{1+x^2}\,.$$

2.4 The Wave Equation: d'Alembert's Formula

Newton's Second Law states that force equals mass times acceleration. It forms the bedrock underlying the derivation of mathematical models describing all of classical dynamics. When applied to a one-dimensional medium, such as the transverse displacements of a violin string or the longitudinal motions of an elastic bar, the resulting model governing small vibrations is the second-order partial differential equation

$$\rho(x)\, \frac{\partial^2 u}{\partial t^2} = \frac{\partial}{\partial x} \left(\kappa(x)\, \frac{\partial u}{\partial x} \right). \tag{2.65}$$

Here $u(t,x)$ represents the displacement of the string or bar at time t and position x, while $\rho(x) > 0$ denotes its density and $\kappa(x) > 0$ its stiffness or tension, both of which are

assumed not to vary with t. The right-hand side of the equation represents the restoring force due to a (small) displacement of the medium from its equilibrium, whereas the left-hand side is the product of mass per unit length and acceleration. A correct derivation of the model from first principles would require a significant detour, and we refer the reader to [**120**, **124**] for the details.

We will simplify the general model by assuming that the underlying medium is *uniform*, and so both its density ρ and stiffness κ are constant. Then (2.65) reduces to the one-dimensional *wave equation*

$$\frac{\partial^2 u}{\partial t^2} = c^2 \, \frac{\partial^2 u}{\partial x^2} \, , \qquad \text{where the constant} \qquad c = \sqrt{\frac{\kappa}{\rho}} \, > \, 0 \tag{2.66}$$

is known as the *wave speed*, for reasons that will soon become apparent.

In general, to uniquely specify the solution to any dynamical system arising from Newton's Second Law, including the wave equation (2.66) and the more general vibration equation (2.65), one must fix both its initial position and initial velocity. Thus, the initial conditions take the form

$$u(0, x) = f(x), \qquad \frac{\partial u}{\partial t}\,(0, x) = g(x), \tag{2.67}$$

where, for simplicity, we set the initial time $t_0 = 0$. (See also Exercise 2.4.6.) The *initial value problem* seeks the corresponding C^2 function $u(t, x)$ that solves the wave equation (2.66) and has the required initial values (2.67). In this section, we will learn how to solve the initial value problem on the entire line $-\infty < x < \infty$. The analysis of the wave equation on bounded intervals will be deferred until Chapters 4 and 7. The two- and three-dimensional versions of the wave equation are treated in Chapters 11 and 12, respectively.

d'Alembert's Solution

Let us now derive the explicit solution formula for the second-order wave equation (2.66) first found by d'Alembert. The starting point is to write the partial differential equation in the suggestive form

$$\Box u = (\partial_t^2 - c^2 \, \partial_x^2) \, u = u_{tt} - c^2 \, u_{xx} = 0. \tag{2.68}$$

Here

$$\Box = \partial_t^2 - c^2 \, \partial_x^2$$

is a common mathematical notation for the *wave operator*, which is a linear second-order partial differential operator. In analogy with the elementary polynomial factorization

$$t^2 - c^2 \, x^2 = (t - c \, x)(t + c \, x),$$

we can factor the wave operator into a product of two first-order partial differential operators:[†]

$$\Box = \partial_t^2 - c^2 \, \partial_x^2 = (\partial_t - c \, \partial_x) \, (\partial_t + c \, \partial_x). \tag{2.69}$$

[†] The cross terms cancel, thanks to the equality of mixed partial derivatives: $\partial_t \partial_x u = \partial_x \partial_t u$. Constancy of the wave speed c is essential here.

Now, if the second factor annihilates the function $u(t,x)$, meaning

$$(\partial_t + c\,\partial_x)\,u = u_t + c\,u_x = 0, \tag{2.70}$$

then u is automatically a solution to the wave equation, since

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

We recognize (2.70) as the first-order transport equation (2.4) with constant wave speed c. Proposition 2.1 tells us that its solutions are traveling waves with wave speed c:

$$u(t,x) = p(\xi) = p(x - ct), \tag{2.71}$$

where p is an arbitrary function of the characteristic variable $\xi = x - ct$. As long as $p \in \mathrm{C}^2$ (i.e., is twice continuously differentiable), the resulting function $u(t,x)$ is a classical solution to the wave equation (2.66), as you can easily check.

Now, the factorization (2.69) can equally well be written in the reverse order:

$$\Box = \partial_t^2 - c^2\,\partial_x^2 = (\partial_t + c\,\partial_x)\,(\partial_t - c\,\partial_x). \tag{2.72}$$

The same argument tells us that any solution to the "backwards" transport equation

$$u_t - c\,u_x = 0, \tag{2.73}$$

with constant wave speed $-c$, also provides a solution to the wave equation. Again, by Proposition 2.1, with c replaced by $-c$, the general solution to (2.73) has the form

$$u(t,x) = q(\eta) = q(x + ct), \tag{2.74}$$

where q is an arbitrary function of the alternative characteristic variable $\eta = x + ct$. The solutions (2.74) represent traveling waves moving to the *left* with constant speed $c > 0$. Provided $q \in \mathrm{C}^2$, the functions (2.74) will provide a second family of solutions to the wave equation.

We conclude that, unlike first-order transport equations, the wave equation (2.68) is *bidirectional* in that it admits both left and right traveling-wave solutions. Moreover, by linearity the sum of any two solutions is again a solution, and so we can immediately construct solutions that are superpositions of left and right traveling waves. The remarkable fact is that *every* solution to the wave equation can be so represented.

Theorem 2.14. *Every solution to the wave equation* (2.66) *can be written as a superposition,*

$$u(t,x) = p(\xi) + q(\eta) = p(x - ct) + q(x + ct), \tag{2.75}$$

of right and left traveling waves. Here $p(\xi)$ and $q(\eta)$ are arbitrary C^2 functions, each depending on its respective characteristic variable

$$\xi = x - ct, \qquad \eta = x + ct. \tag{2.76}$$

Proof: As in our treatment of the transport equation, we will simplify the wave equation through an inspired change of variables. In this case, the new independent variables are the characteristic variables ξ, η defined by (2.76). We set

$$u(t,x) = v(x - ct, x + ct) = v(\xi,\eta), \qquad \text{whereby} \qquad v(\xi,\eta) = u\left(\frac{\eta - \xi}{2\,c}\,,\,\frac{\eta + \xi}{2}\right). \tag{2.77}$$

Then, employing the chain rule to compute the partial derivatives,

$$\frac{\partial u}{\partial t} = c\left(-\frac{\partial v}{\partial \xi} + \frac{\partial v}{\partial \eta} \right), \qquad \frac{\partial u}{\partial x} = \frac{\partial v}{\partial \xi} + \frac{\partial v}{\partial \eta}, \tag{2.78}$$

and, further,

$$\frac{\partial^2 u}{\partial t^2} = c^2 \left(\frac{\partial^2 v}{\partial \xi^2} - 2\,\frac{\partial^2 v}{\partial \xi\,\partial \eta} + \frac{\partial^2 v}{\partial \eta^2} \right), \qquad \frac{\partial^2 u}{\partial x^2} = \frac{\partial^2 v}{\partial \xi^2} + 2\,\frac{\partial^2 v}{\partial \xi\,\partial \eta} + \frac{\partial^2 v}{\partial \eta^2}.$$

Therefore

$$\Box\, u = \frac{\partial^2 u}{\partial t^2} - c^2\,\frac{\partial^2 u}{\partial x^2} = -\,4\,c^2\,\frac{\partial^2 v}{\partial \xi\,\partial \eta}. \tag{2.79}$$

We conclude that $u(t,x)$ solves the wave equation $\Box\, u = 0$ if and only if $v(\xi,\eta)$ solves the second-order partial differential equation

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

which we write in the form

$$\frac{\partial}{\partial \xi}\left(\frac{\partial v}{\partial \eta} \right) = \frac{\partial w}{\partial \xi} = 0, \qquad \text{where} \qquad w = \frac{\partial v}{\partial \eta}.$$

Thus, applying the methods of Section 2.1 (and making the appropriate assumptions on the domain of definition of w), we deduce that

$$w = \frac{\partial v}{\partial \eta} = r(\eta),$$

where r is an arbitrary function of the characteristic variable η. Integrating both sides of the latter partial differential equation with respect to η, we find

$$v(\xi,\eta) = p(\xi) + q(\eta), \qquad \text{where} \qquad q(\eta) = \int r(\eta)\,d\eta,$$

while $p(\xi)$ represents the η integration "constant". Replacing the characteristic variables by their formulas in terms of t and x completes the proof. *Q.E.D.*

Let us see how the solution formula (2.75) can be used to solve the initial value problem (2.67). Substituting into the initial conditions, we deduce that

$$u(0,x) = p(x) + q(x) = f(x), \qquad \frac{\partial u}{\partial t}\,(0,x) = -\,c\,p'(x) + c\,q'(x) = g(x). \tag{2.80}$$

To solve this pair of equations for the functions p and q, we differentiate the first,

$$p'(x) + q'(x) = f'(x),$$

and then subtract off the second equation divided by c; the result is

$$2\,p'(x) = f'(x) - \frac{1}{c}\,g(x).$$

Therefore,

$$p(x) = \frac{1}{2}\,f(x) - \frac{1}{2\,c}\int_0^x g(z)\,dz + a,$$

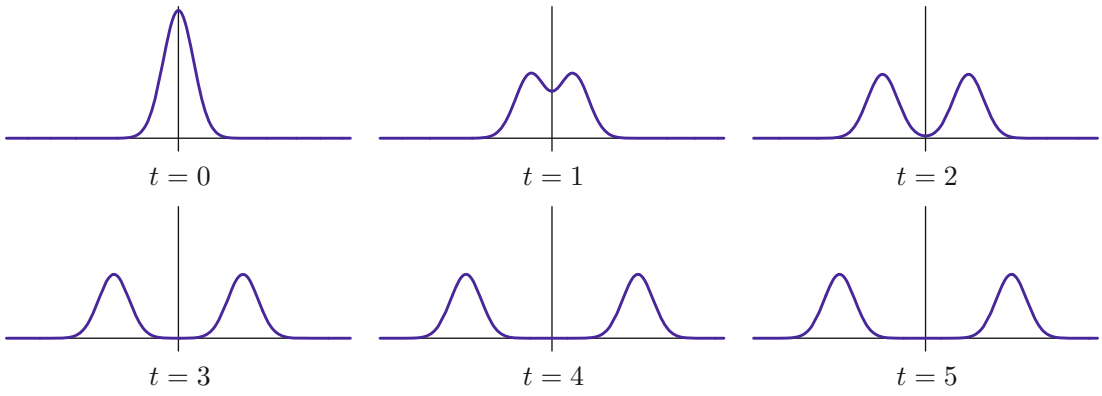


Figure 2.24. Splitting of waves.

where a is an integration constant. The first equation in (2.80) then yields

$$q(x) = f(x) - p(x) = \frac{1}{2} f(x) + \frac{1}{2\,c} \int_0^x g(z)\, dz - a.$$

Substituting these two expressions back into our solution formula (2.75), we obtain

$$\begin{aligned} u(t,x) = p(\xi) + q(\eta) &= \frac{f(\xi) + f(\eta)}{2} - \frac{1}{2\,c} \int_0^{\xi} g(z)\, dz + \frac{1}{2\,c} \int_0^{\eta} g(z)\, dz \\ &= \frac{f(\xi) + f(\eta)}{2} + \frac{1}{2\,c} \int_{\xi}^{\eta} g(z)\, dz, \end{aligned}$$

where ξ, η are the characteristic variables (2.76). In this manner, we have arrived at *d'Alembert's solution* to the initial value problem for the wave equation on the real line.

Theorem 2.15. *The solution to the initial value problem*

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}, \qquad u(0,x) = f(x), \qquad \frac{\partial u}{\partial t}(0,x) = g(x), \qquad -\infty < x < \infty, \tag{2.81}$$

is given by

$$u(t,x) = \frac{f(x - c\,t) + f(x + c\,t)}{2} + \frac{1}{2\,c} \int_{x - c\,t}^{x + c\,t} g(z)\, dz. \tag{2.82}$$

Remark: In order that (2.82) define a classical solution to the wave equation, we need $f \in \mathrm{C}^2$ and $g \in \mathrm{C}^1$. However, the formula itself makes sense for more general initial conditions. We will continue to treat the resulting functions as solutions, albeit nonclassical, since they fit under the more general rubric of "weak solution", to be developed in Section 10.4.

Example 2.16. Suppose there is no initial velocity, so $g(x) \equiv 0$, and hence the motion is purely the result of the initial displacement $u(0,x) = f(x)$. In this case, (2.82) reduces to

$$u(t,x) = \tfrac{1}{2} f(x - c\,t) + \tfrac{1}{2} f(x + c\,t). \tag{2.83}$$

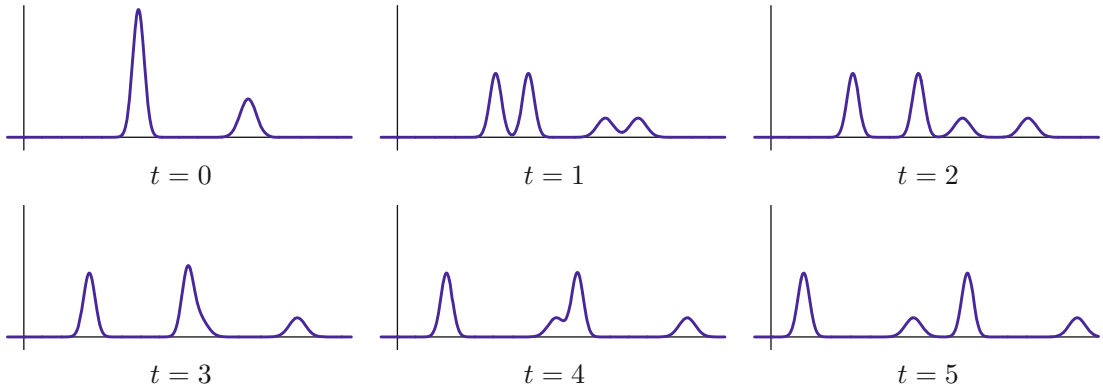


Figure 2.25. Interaction of waves.

The effect is that the initial displacement splits into two waves, one moving to the right and the other moving to the left, each of constant speed c, and each of exactly the same shape as $f(x)$, but only half as tall. For example, if the initial displacement is a localized pulse centered at the origin, say

$$u(0,x) = e^{-x^2}, \qquad \frac{\partial u}{\partial t}(0,x) = 0,$$

then the solution

$$u(t,x) = \tfrac{1}{2}\, e^{-(x-ct)^2} + \tfrac{1}{2}\, e^{-(x+ct)^2}$$

consists of two half size pulses running away from the origin with the same speed c, but in opposite directions. A graph of the solution at several successive times can be seen in Figure 2.24.

If we take two initially separated pulses, say

$$u(0,x) = e^{-x^2} + 2\, e^{-(x-1)^2}, \qquad \frac{\partial u}{\partial t}(0,x) = 0,$$

centered at $x = 0$ and $x = 1$, then the solution

$$u(t,x) = \tfrac{1}{2}\, e^{-(x-ct)^2} + e^{-(x-1-ct)^2} + \tfrac{1}{2}\, e^{-(x+ct)^2} + e^{-(x-1+ct)^2}$$

will consist of four pulses, two moving to the right and two to the left, all with the same speed. An important observation is that when a right-moving pulse collides with a left-moving pulse, they emerge from the collision unchanged, which is a consequence of the inherent linearity of the wave equation. In Figure 2.25, the first picture plots the initial displacement. In the second and third pictures, the two localized bumps have each split into two copies moving in opposite directions. In the fourth and fifth, the larger right-moving bump is in the process of interacting with the smaller left-moving bump. Finally, in the last picture the interaction is complete, and the individual pairs of left- and right-moving waves move off in tandem in opposing directions, experiencing no further collisions.

In general, if the initial displacement is localized, so that $|\, f(x)\,| \ll 1$ for $|\, x\,| \gg 0$, then, after a finite time, the left- and right-moving waves will separate, and the observer will see two half-size replicas running away, with speed c, in opposite directions. If the displacement

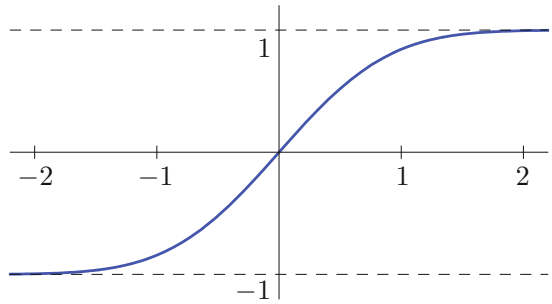


Figure 2.26. The error function erf x.

is not localized, then the left and right traveling waves will never fully disengage, and one might be hard pressed to recognize that a complicated solution pattern is, in reality, just the superposition of two simple traveling waves. For example, consider the elementary trigonometric solution

$$\cos ct\ \cos x = \tfrac{1}{2}\cos(x - ct) + \tfrac{1}{2}\cos(x + ct). \qquad (2.84)$$

In accordance with the left-hand expression, an observer will see a standing cosinusoidal wave that vibrates up and down with frequency c. However, the d'Alembert form of the solution on the right-hand side says that this is just the sum of left- and right-traveling cosine waves! The interactions of their peaks and troughs reproduce the standing wave. Thus, the same solution can be interpreted in two seemingly incompatible ways. And, in fact, this paradox lies at the heart of the perplexing wave-particle duality of quantum physics.

Example 2.17. By way of contrast, suppose there is no initial displacement, so $f(x) \equiv 0$, and the motion is purely the result of the initial velocity $u_t(0, x) = g(x)$. Physically, this models a violin string at rest being struck by a "hammer blow" at the initial time. In this case, the d'Alembert formula (2.82) reduces to

$$u(t, x) = \frac{1}{2\,c}\int_{x-ct}^{x+ct} g(z)\,dz. \qquad (2.85)$$

For example, when $u(0, x) = 0$, $u_t(0, x) = e^{-x^2}$, the resulting solution (2.85) is

$$u(t, x) = \frac{1}{2\,c}\int_{x-ct}^{x+ct} e^{-x^2}\,dz = \frac{\sqrt{\pi}}{4\,c}\big[\,\mathrm{erf}(x + ct) - \mathrm{erf}(x - ct)\,\big], \qquad (2.86)$$

where

$$\mathrm{erf}\, x = \frac{2}{\sqrt{\pi}}\int_0^x e^{-z^2}\,dz \qquad (2.87)$$

is known as the *error function* due to its many applications throughout probability and statistics, [**39**]. The error function integral cannot be written in terms of elementary functions; nevertheless, its properties have been well studied and its values tabulated, [**86**]. A graph appears in Figure 2.26. The constant in front of the integral (2.87) has been chosen so that the error function has asymptotic values

$$\lim_{x\to\infty} \mathrm{erf}\, x = 1, \qquad \lim_{x\to-\infty} \mathrm{erf}\, x = -1, \qquad (2.88)$$

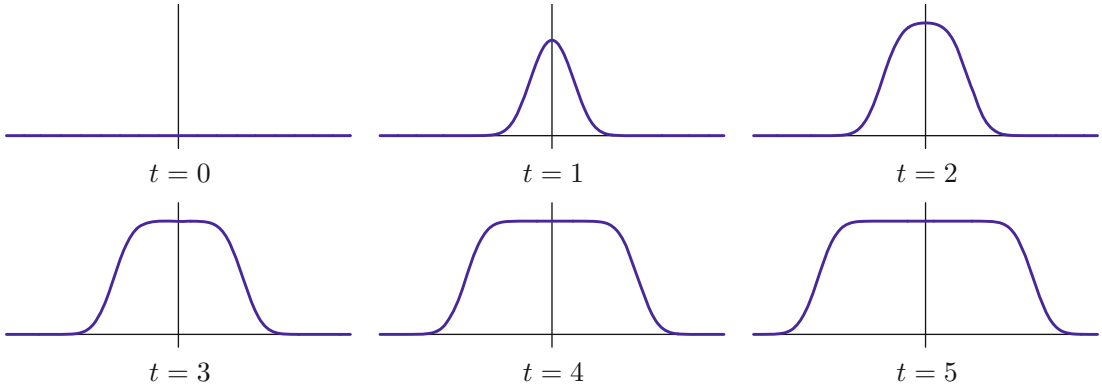


Figure 2.27. Error function solution to the wave equation.

which follow from a well-known integration formula to be derived in Exercise 2.4.21.

A graph of the solution (2.86) at successive times is displayed in Figure 2.27. The first graph shows the zero initial displacement. Gradually, the effect of the initial hammer blow is felt further and further away along the string, as the two wave fronts propagate away from the origin, both with speed c, but in opposite directions. Thus, unlike the case of a nonzero initial displacement in Figure 2.24, where the solution eventually returns to its equilibrium position $u = 0$ after the wave passes by, a nonzero initial velocity leaves the string permanently deformed.

In general, the lines of slope $\pm c$, where the respective characteristic variables are constant,

$$\xi = x - ct = a, \qquad \eta = x + ct = b, \tag{2.89}$$

are known as the *characteristics* of the wave equation. Thus, the second-order wave equation has *two* distinct characteristic lines passing through each point in the (t, x)–plane.

Remark: The characteristic lines are the one-dimensional counterparts of the light cone in Minkowski space-time, which plays a starring role in special relativity, [**70**, **75**]. See Section 12.5 for further details.

In Figure 2.28, we plot the two characteristics going through a point $(0, y)$ on the x axis. The wedge-shaped region $\{\, y - ct \le x \le y + ct,\ t \ge 0 \,\}$ lying between them is known as the *domain of influence* of the point $(0, y)$, since, in general, the value of the initial data at a point will affect the subsequent solution values only in its domain of influence. Indeed, the effect of an initial displacement at the point y propagates along the two characteristic lines, while the effect of an initial velocity there will be felt at every point in the triangular wedge.

External Forcing and Resonance

When a homogeneous vibrating medium is subjected to external forcing, the wave equation acquires an additional, inhomogeneous term:

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} + F(t, x), \tag{2.90}$$

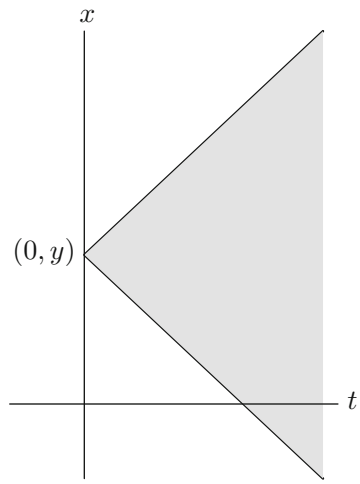


Figure 2.28. Characteristic lines and domain of influence.

in which $F(t,x)$ represents a force imposed at time t and spatial position x. With a bit more work, d'Alembert's solution technique can be readily adapted to incorporate the forcing term.

Let us, for simplicity, assume that the differential equation is supplemented by homogeneous initial conditions,

$$u(0,x) = 0, \qquad u_t(0,x) = 0, \tag{2.91}$$

meaning that there is no initial displacement or velocity. To solve the initial value problem (2.90–91), we switch to the same characteristic coordinates (2.76), setting

$$v(\xi,\eta) = u\left(\frac{\eta-\xi}{2\,c}, \frac{\eta+\xi}{2}\right).$$

Invoking the chain rule formulas (2.79), we find that the forced equation (2.90) becomes

$$\frac{\partial^2 v}{\partial\xi\,\partial\eta} = -\frac{1}{4\,c^2}\,F\left(\frac{\eta-\xi}{2\,c}, \frac{\eta+\xi}{2}\right). \tag{2.92}$$

Let us integrate both sides of the equation with respect to η, on the interval $\xi \le \zeta \le \eta$:

$$\frac{\partial v}{\partial\xi}(\xi,\eta) - \frac{\partial v}{\partial\xi}(\xi,\xi) = -\frac{1}{4\,c^2}\int_\xi^\eta F\left(\frac{\zeta-\xi}{2\,c}, \frac{\zeta+\xi}{2}\right) d\zeta. \tag{2.93}$$

But, recalling (2.78),

$$\frac{\partial v}{\partial\xi}(\xi,\eta) = \frac{1}{2\,c}\,\frac{\partial u}{\partial t}\left(\frac{\eta-\xi}{2\,c}, \frac{\eta+\xi}{2}\right) + \frac{1}{2}\,\frac{\partial u}{\partial x}\left(\frac{\eta-\xi}{2\,c}, \frac{\eta+\xi}{2}\right),$$

and so, in particular,

$$\frac{\partial v}{\partial\xi}(\xi,\xi) = \frac{1}{2\,c}\,\frac{\partial u}{\partial t}(0,\xi) + \frac{1}{2}\,\frac{\partial u}{\partial x}(0,\xi) = 0,$$

which vanishes owing to our choice of homogeneous initial conditions (2.91). Indeed, the initial velocity condition says that $u_t(0,x) = 0$, while differentiating the initial displacement

condition $u(0,x) = 0$ with respect to x implies that $u_x(0,x) = 0$ for all x, including $x = \xi$. As a result, (2.93) simplifies to

$$\frac{\partial v}{\partial \xi}(\xi,\eta) = -\frac{1}{4c^2}\int_\xi^\eta F\left(\frac{\zeta-\xi}{2c}, \frac{\zeta+\xi}{2}\right) d\zeta.$$

We now integrate the latter equation with respect to ξ on the interval $\xi \le \chi \le \eta$, producing

$$-v(\xi,\eta) = v(\eta,\eta) - v(\xi,\eta) = -\frac{1}{4c^2}\int_\xi^\eta \int_\chi^\eta F\left(\frac{\zeta-\chi}{2c}, \frac{\zeta+\chi}{2}\right) d\zeta\, d\chi,$$

since $v(\eta,\eta) = u(0,\eta) = 0$, thanks again to the initial conditions. In this manner, we have produced an explicit formula for the solution to the characteristic variable version of the forced wave equation subject to the homogeneous initial conditions. Reverting to the original physical coordinates, the left-hand side of this equation becomes $-u(t,x)$. As for the double integral on the right-hand side, it takes place over the triangular region

$$T(\xi,\eta) = \{\, (\chi,\zeta) \mid \xi \le \chi \le \zeta \le \eta \,\}. \tag{2.94}$$

Let us introduce "physical" integration variables by setting

$$\chi = y - c\,s, \qquad \zeta = y + c\,s.$$

The defining inequalities of the triangle (2.94) become

$$x - ct \le y - c\,s \le y + c\,s \le x + ct,$$

and so, in the physical coordinates, the triangular integration domain assumes the form

$$D(t,x) = \{\, (s,y) \mid x - c\,(t-s) \le y \le x + c\,(t-s),\ 0 \le s \le t \,\}, \tag{2.95}$$

which is graphed in Figure 2.29. The change of variables formula for double integrals requires that we compute the Jacobian determinant

$$\det\begin{pmatrix} \partial\chi/\partial y & \partial\chi/\partial s \\ \partial\zeta/\partial y & \partial\zeta/\partial s \end{pmatrix} = \det\begin{pmatrix} 1 & -c \\ 1 & c \end{pmatrix} = 2c,$$

and so $d\chi\, d\zeta = 2c\, ds\, dy$. Therefore,

$$u(t,x) = \frac{1}{2c}\iint_{D(t,x)} F(s,y)\, ds\, dy = \frac{1}{2c}\int_0^t \int_{x-c\,(t-s)}^{x+c\,(t-s)} F(s,y)\, dy\, ds, \tag{2.96}$$

which gives the solution formula for the forced wave equation when subject to homogeneous initial conditions.

To solve the general initial value problem, we appeal to linear superposition, writing its solution as a sum of the solution (2.96) to the forced wave equation subject to homogeneous initial conditions plus the d'Alembert solution (2.82) to the unforced equation subject to inhomogeneous boundary conditions.

Theorem 2.18. *The solution to the general initial value problem*

$$u_{tt} = c^2 u_{xx} + F(t,x), \qquad u(0,x) = f(x), \qquad u_t(0,x) = g(x), \qquad -\infty < x < \infty, \qquad t > 0,$$

for the wave equation subject to an external forcing is given by

$$u(t,x) = \frac{f(x-ct) + f(x+ct)}{2} + \frac{1}{2c}\int_{x-ct}^{x+ct} g(y)\, dy + \frac{1}{2c}\int_0^t \int_{x-c\,(t-s)}^{x+c\,(t-s)} F(s,y)\, dy\, ds. \tag{2.97}$$

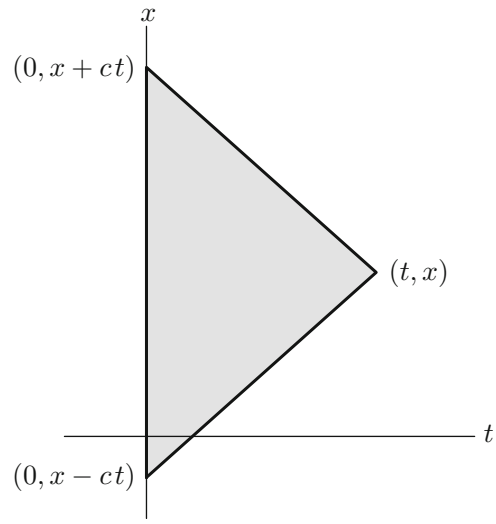


Figure 2.29. Domain of dependence.

Observe that the solution is a linear superposition of the respective effects of the initial displacement, the initial velocity, and the external forcing. The triangular integration region (2.95), lying between the x–axis and the characteristic lines going backwards from (t, x), is known as the *domain of dependence* of the point (t, x). This is because, for any $t > 0$, the solution value $u(t, x)$ depends only on the values of the initial data and the forcing function at points lying within the domain of dependence $D(t, x)$. Indeed, the first term in the solution formula (2.97) requires only the initial displacement at the corners $(0, x + ct)$, $(0, x - ct)$; the second term requires only the initial velocity at points on the x–axis lying on the vertical side of $D(t, x)$; while the final term requires the value of the external force on the entire triangular region.

Example 2.19. Let us solve the initial value problem

$$u_{tt} = u_{xx} + \sin \omega t \, \sin x, \qquad u(0, x) = 0, \qquad u_t(0, x) = 0,$$

for the wave equation with unit wave speed subject to a sinusoidal forcing function whose amplitude varies periodically in time with frequency $\omega > 0$. According to formula (2.96), the solution is

$$\begin{aligned} u(t, x) &= \frac{1}{2} \int_0^t \int_{x-t+s}^{x+t-s} \sin \omega s \sin y \, dy \, ds \\ &= \frac{1}{2} \int_0^t \sin \omega s \, \big[\cos(x - t + s) - \cos(x + t - s) \big] \, ds \\ &= \begin{cases} \dfrac{\sin \omega t - \omega \sin t}{1 - \omega^2} \, \sin x, & 0 < \omega \neq 1, \\ \dfrac{\sin t - t \cos t}{2} \, \sin x, & \omega = 1. \end{cases} \end{aligned}$$

Notice that, when $\omega \neq 1$, the solution is bounded, being a combination of two vibrational modes: an externally induced mode at frequency ω along with an internal mode, at frequency 1. If $\omega = p/q \neq 1$ is a rational number, then the solution varies periodically in

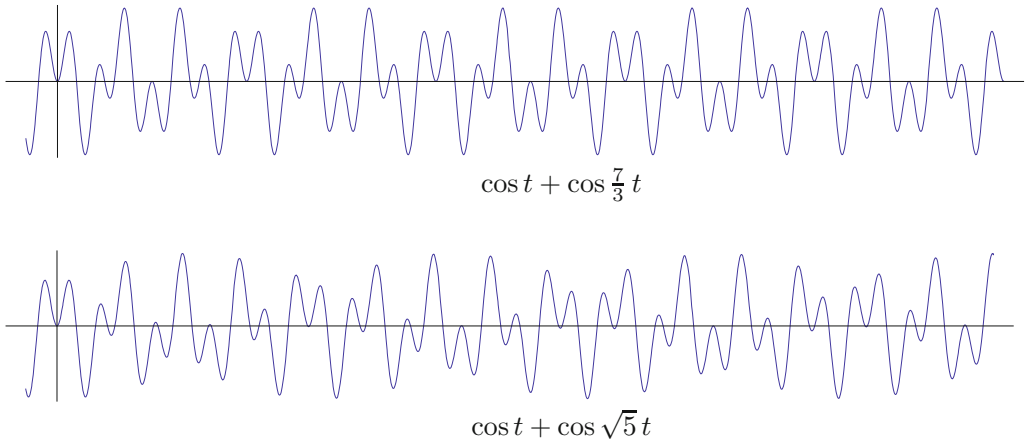


Figure 2.30. Periodic and quasiperiodic functions.

time. On the other hand, if ω is irrational, then the solution is only *quasiperiodic*, and never exactly repeats itself. Finally, if $\omega = 1$, the solution grows without limit as t increases, indicating that this is a *resonant frequency*. We will investigate external forcing and the mechanisms leading to resonance in dynamical partial differential equations in more detail in Chapters 4 and 6.

Example 2.20. To appreciate the difference between periodic and quasiperiodic vibrations, consider the elementary trigonometric function

$$u(t) = \cos t + \cos \omega\, t,$$

which is a linear combination of two simple periodic vibrations, of frequencies 1 and ω. If $\omega = p/q$ is a rational number, then $u(t)$ is a periodic function of period $2\,\pi q$, so $u(t+2\,\pi q) = u(t)$. However, if ω is an irrational number, then $u(t)$ is not periodic, and never repeats. You are encouraged to inspect the graphs in Figure 2.30. The first is periodic — can you spot where it begins to repeat? — whereas the second is only quasiperiodic. The only quasiperiodic functions we will encounter in this text are linear combinations of periodic trigonometric functions whose frequencies are *not* all rational multiples of each other. To the uninitiated, such quasiperiodic motions may appear to be random, even though they are built from a few simple periodic constituents. While ostensibly complicated, quasiperiodic motion is *not* true chaos, which is is an inherently nonlinear phenomenon, [**77**].

Exercises

2.4.1. Solve the initial value problem $u_{tt} = c^2 u_{xx}$, $u(0,x) = e^{-x^2}$, $u_t(0,x) = \sin x$.

2.4.2. (a) Solve the wave equation $u_{tt} = u_{xx}$ when the initial displacement is the box function $u(0,x) = \begin{cases} 1, & 1 < x < 2, \\ 0, & \text{otherwise}, \end{cases}$ while the initial velocity is 0.

(b) Sketch the resulting solution at several representative times.

2.4.3. Answer Exercise 2.4.2 when the initial velocity is the box function, while the initial displacement is zero.

2.4.4. Write the following solutions to the wave equation $u_{tt} = u_{xx}$ in d'Alembert form (2.82). *Hint*: What is the appropriate initial data?
(*a*) $\cos x \cos t$, (*b*) $\cos 2x \, \sin 2t$, (*c*) e^{x+t}, (*d*) $t^2 + x^2$, (*e*) $t^3 + 3tx^2$.

♡ 2.4.5. (*a*) Solve the *dam break problem*, that is, the wave equation when the initial displacement is a step function $\sigma(x) = \begin{cases} 1, & x > 0, \\ 0, & x < 0, \end{cases}$ and there is no initial velocity. (*b*) Analyze the case in which there is no initial displacement, while the initial velocity is a step function. (*c*) Are your solutions classical solutions? Explain your answer. (*d*) Prove that the step function is the limit, as $n \to \infty$, of the functions $f_n(x) = \dfrac{1}{\pi} \tan^{-1} n\,x + \dfrac{1}{2}$. (*e*) Show that, in both cases, the step function solution can be realized as the limit, as $n \to \infty$, of solutions to the initial value problems with the functions $f_n(x)$ as initial displacement or velocity.

♢ 2.4.6. Suppose $u(t, x)$ solves the initial value problem $u(0, x) = f(x)$, $u_t(0, x) = g(x)$, for the wave equation (2.66). Prove that the solution to the initial value problem $u(t_0, x) = f(x)$, $u_t(t_0, x) = g(x)$, is $u(t - t_0, x)$.

2.4.7. Find all resonant frequencies for the wave equation with wave speed c when subject to the external forcing function $F(t, x) = \sin \omega t \, \sin k\,x$ for fixed $\omega, k > 0$.

2.4.8. Consider the initial value problem $u_{tt} = 4\,u_{xx} + F(t, x)$, $u(0, x) = f(x)$, $u_t(0, x) = g(x)$. Determine (*a*) the domain of influence of the point $(0, 2)$; (*b*) the domain of dependence of the point $(3, -1)$; (*c*) the domain of influence of the point $(3, -1)$.

2.4.9. (*a*) A solution to the wave equation $u_{tt} = 2\,u_{xx}$ is generated by a displacement concentrated at position $x_0 = 1$ and time $t_0 = 0$, but no initial velocity. At what time will an observer at position $x_1 = 5$ feel the effect of this displacement? Will the observer continue to feel an effect in the future? (*b*) Answer part (*a*) when there is an initial velocity concentrated at position $x_0 = 1$ and time $t_0 = 0$, but no initial displacement.

2.4.10. Suppose $u(t, x)$ solves the initial value problem $u_{tt} = 4\,u_{xx} + \sin \omega t \, \cos x$, $u(0, x) = 0$, $u_t(0, x) = 0$. Is $h(t) = u(t, 0)$ a periodic function?

♡ 2.4.11. (*a*) Write down an explicit formula for the solution to the initial value problem

$$\frac{\partial^2 u}{\partial t^2} - 4\,\frac{\partial^2 u}{\partial x^2} = 0, \qquad u(0, x) = \sin x, \qquad \frac{\partial u}{\partial t}\,(0, x) = \cos x, \qquad -\infty < x < \infty, \quad t \geq 0.$$

(*b*) *True or false*: The solution is a periodic function of t.
(*c*) Now solve the forced initial value problem

$$\frac{\partial^2 u}{\partial t^2} - 4\,\frac{\partial^2 u}{\partial x^2} = \cos 2t, \qquad u(0, x) = \sin x, \qquad \frac{\partial u}{\partial t}\,(0, x) = \cos x, \qquad -\infty < x < \infty, \quad t \geq 0.$$

(*d*) *True or false*: The forced equation exhibits resonance. Explain.
(*e*) Does the answer to part (*d*) change if the forcing function is $\sin 2t$?

2.4.12. Given a classical solution $u(t, x)$ of the wave equation, let $E = \frac{1}{2}(u_t^2 + c^2 u_x^2)$ be the associated *energy density* and $P = u_t u_x$ the *momentum density*.
(*a*) Show that both E and P are conserved densities for the wave equation.
(*b*) Show that $E(t, x)$ and $P(t, x)$ both satisfy the wave equation.

♢ 2.4.13. Let $u(t, x)$ be a classical solution to the wave equation $u_{tt} = c^2 u_{xx}$. The *total energy*

$$E(t) = \int_{-\infty}^{\infty} \frac{1}{2} \left[\left(\frac{\partial u}{\partial t} \right)^2 + c^2 \left(\frac{\partial u}{\partial x} \right)^2 \right] dx \tag{2.98}$$

represents the sum of kinetic and potential energies of the displacement $u(t, x)$ at time t. Suppose that $\nabla u \to \mathbf{0}$ sufficiently rapidly as $x \to \pm\infty$; more precisely, one can find $\alpha > \frac{1}{2}$ and $C(t) > 0$ such that $|\, u_t(t, x)\,|$, $|\, u_x(t, x)\,| \leq C(t)/|\, x\,|^{\alpha}$ for each fixed t and all sufficiently large $|\, x\,| \gg 0$. For such solutions, establish the *Law of Conservation of Energy* by showing that $E(t)$ is finite and constant. *Hint*: You do not need the formula for the solution.

♢ 2.4.14. (*a*) Use Exercise 2.4.13 to prove that the only classical solution to the initial-boundary value problem $u_{tt} = c^2 u_{xx}$, $u(0,x) = 0$, $u_t(0,x) = 0$, satisfying the indicated decay assumptions is the trivial solution $u(t,x) \equiv 0$. (*b*) Establish the following *Uniqueness Theorem* for the wave equation: there is at most one such solution to the initial-boundary value problem $u_{tt} = c^2 u_{xx}$, $u(0,x) = f(x)$, $u_t(0,x) = g(x)$.

2.4.15. The *telegrapher's equation* $u_{tt} + a\,u_t = c^2 u_{xx}$, with $a > 0$, models the vibration of a string under frictional damping. (*a*) Show that, under the decay assumptions of Exercise 2.4.13, the wave energy (2.98) of a classical solution is a nonincreasing function of t. (*b*) Prove uniqueness of such solutions to the initial value problem for the telegrapher's equation.

2.4.16. What happens to the proof of Theorem 2.14 if $c = 0$?

2.4.17. (*a*) Explain why the d'Alembert factorization method doesn't work when the wave speed $c(x)$ depends on the spatial variable x.
(*b*) Does it work when $c(t)$ depends only on the time t?

2.4.18. The *Poisson–Darboux equation* is $\dfrac{\partial^2 u}{\partial t^2} - \dfrac{\partial^2 u}{\partial x^2} - \dfrac{2}{x}\dfrac{\partial u}{\partial x} = 0$. Solve the initial value problem $u(0,x) = 0$, $u_t(0,x) = g(x)$, where $g(x) = g(-x)$ is an even function. *Hint*: Set $w = x\,u$.

♡ 2.4.19. (*a*) Solve the initial value problem $u_{tt} - 2\,u_{tx} - 3\,u_{xx} = 0$, $u(0,x) = x^2$, $u_t(0,x) = e^x$. *Hint*: Factor the associated linear differential operator. (*b*) Determine the domain of influence of a point $(0,x)$. (*c*) Determine the domain of dependence of a point (t,x) with $t > 0$.

♢ 2.4.20. (*a*) Use polar coordinates to prove that, for any $a > 0$,

$$\iint_{\mathbb{R}^2} e^{-a\,(x^2+y^2)}\,dx\,dy = \frac{\pi}{a}\,. \tag{2.99}$$

(*b*) Explain why

$$\int_{-\infty}^{\infty} e^{-a\,x^2}\,dx = \sqrt{\frac{\pi}{a}}\,. \tag{2.100}$$

♢ 2.4.21. Use Exercise 2.4.20 to prove the error function formulae (2.88).

Chapter 4
Separation of Variables

Three cardinal linear second-order partial differential equations have collectively driven the development of the entire subject. The first two we have already encountered: The *wave equation* describes vibrations and waves in continuous media, including sound waves, water waves, elastic waves, electromagnetic waves, and so on. The *heat equation* models diffusion processes, including thermal energy in solids, solutes in liquids, and biological populations. Third, and in many ways the most important of all, is the *Laplace equation* and its inhomogeneous counterpart, the *Poisson equation*, which govern equilibrium mechanics. The latter two equations arise in an astonishing variety of mathematical and physical contexts, ranging through elasticity and solid mechanics, fluid mechanics, electromagnetism, potential theory, thermomechanics, geometry, probability, number theory, and many other fields. The solutions to the Laplace equation are known as harmonic functions, and the discovery of their many remarkable properties forms one of the most celebrated chapters in the history of mathematics. All three equations, along with their multi-dimensional kin, will appear repeatedly throughout this text.

The aim of the current chapter is to develop the method of separation of variables for solving these key partial differential equations in their two-independent-variable incarnations. For the wave and heat equations, the variables are time, t, and a single space coordinate, x, leading to initial-boundary value problems modeling the dynamical behavior of a one-dimensional medium. For the Laplace and Poisson equations, both variables represent space coordinates, x and y, and the associated boundary value problems model the equilibrium configuration of a planar body, e.g., the deformations of a membrane. Separation of variables seeks special solutions that can be written as the product of functions of the individual variables, thereby reducing the partial differential equation to a pair of ordinary differential equations. More-general solutions can then be expressed as infinite series in the appropriate separable solutions. For the two-variable equations considered here, this results in a Fourier series representation of the solution. In the case of the wave equation, separation of variables serves to focus attention on the vibrational character of the solution, whereas the earlier d'Alembert approach emphasizes its particle-like aspects. Unfortunately, for the Laplace equation, separation of variables applies only to boundary value problems in very special geometries, e.g., rectangles and disks. Further development of the separation of variables method for solving partial differential equations in three or more variables can be found in Chapters 11 and 12.

In the final section, we take the opportunity to summarize the fundamental tripartite classification of planar second-order partial differential equations. Each of the three paradigmatic equations epitomizes one of the classes: *hyperbolic*, such as the wave equa-

P.J. Olver, *Introduction to Partial Differential Equations*, Undergraduate Texts in Mathematics,
DOI 10.1007/978-3-319-02099-0_4, © Springer International Publishing Switzerland 2014

tion; *parabolic*, such as the heat equation; and *elliptic*, such as the Laplace and Poisson equations. Each category enjoys its own distinctive properties and features, both analytic and numeric, and, in effect, forms a separate mathematical subdiscipline.

4.1 The Diffusion and Heat Equations

Let us begin with a brief physical derivation of the heat equation from first principles. We consider a *bar* — meaning a thin, heat-conducting body. "Thin" means that we can regard the bar as a one-dimensional continuum with no significant transverse temperature variation. We will assume that the bar is fully insulated along its length, and so heat can enter (or leave) only through its uninsulated endpoints. We use t to represent time, and $a \le x \le b$ to denote spatial position along the bar, which occupies the interval $[a, b]$. Our goal is to find the temperature $u(t, x)$ of the bar at position x and time t.

The dynamical equations governing the temperature are based on three fundamental physical principles. First is the Law of Conservation of Heat Energy. Recalling the general Definition 2.7, this particular conservation law takes the form

$$\frac{\partial \varepsilon}{\partial t} + \frac{\partial w}{\partial x} = 0, \tag{4.1}$$

in which $\varepsilon(t, x)$ represents the thermal *energy density* at time t and position x, while $w(t, x)$ denotes the *heat flux*, i.e., the rate of flow of thermal energy along the bar. Our sign convention is that $w(t, x) > 0$ at points where the energy flows in the direction of increasing x (left to right). The integrated form (2.49) of the conservation law, namely

$$\frac{d}{dt} \int_a^b \varepsilon(t, x)\, dx = w(t, a) - w(t, b), \tag{4.2}$$

states that the rate of change in the thermal energy within the bar is equal to the total heat flux passing through its uninsulated ends. The signs of the boundary terms confirm that heat flux *into* the bar results in an increase in temperature.

The second ingredient is a *constitutive assumption* concerning the bar's material properties. It has been observed that, under reasonable conditions, thermal energy is proportional to temperature:

$$\varepsilon(t, x) = \sigma(x)\, u(t, x). \tag{4.3}$$

The factor

$$\sigma(x) = \rho(x)\, \chi(x) > 0 \tag{4.4}$$

is the product of the *density* ρ of the material and its *specific heat capacity* χ, which is the amount of heat energy required to raise the temperature of a unit mass of the material by one degree. Note that we are assuming that the medium is not changing in time, and so physical quantities such as density and specific heat depend only on position x. We also assume, perhaps with less physical justification, that its material properties do not depend upon the temperature; otherwise, we would be forced to deal with a much thornier nonlinear diffusion equation, [**70**, **99**].

The third physical principle relates heat flux and temperature. Physical experiments show that the thermal energy moves from hot to cold at a rate that is in direct proportion to

the temperature gradient, which, in the one-dimensional case, means its derivative $\partial u/\partial x$. The resulting relation

$$w(t,x) = -\kappa(x)\,\frac{\partial u}{\partial x} \tag{4.5}$$

is known as *Fourier's Law of Cooling*. The proportionality factor $\kappa(x) > 0$ is the *thermal conductivity* of the bar at position x, and the minus sign reflects the everyday observation that heat energy moves from hot to cold. A good heat conductor, e.g., silver, will have high conductivity, while a poor conductor, e.g., glass, will have low conductivity.

Combining the three laws (4.1, 3, 5) produces the *linear diffusion equation*

$$\frac{\partial}{\partial t}\big(\,\sigma(x)\,u\,\big) = \frac{\partial}{\partial x}\left(\kappa(x)\,\frac{\partial u}{\partial x}\right), \qquad a < x < b, \tag{4.6}$$

governing the thermodynamics of a one-dimensional medium. It is also used to model a wide variety of diffusive processes, including chemical diffusion, diffusion of contaminants in liquids and gases, population dispersion, and the spread of infectious diseases. If there is an external heat source along the length of the bar, then the diffusion equation acquires an additional prescribed inhomogeneous term:

$$\frac{\partial}{\partial t}\big(\,\sigma(x)\,u\,\big) = \frac{\partial}{\partial x}\left(\kappa(x)\,\frac{\partial u}{\partial x}\right) + h(t,x), \qquad a < x < b. \tag{4.7}$$

In order to uniquely prescribe the solution $u(t,x)$, we need to specify an initial temperature distribution

$$u(t_0,x) = f(x), \qquad a \le x \le b. \tag{4.8}$$

In addition, we must impose a suitable boundary condition at each end of the bar. There are three common types. The first is a *Dirichlet boundary condition*, where the end is held at a prescribed temperature. For example,

$$u(t,a) = \alpha(t) \tag{4.9}$$

fixes the temperature (possibly time-varying) at the left end. Alternatively, the *Neumann boundary condition*

$$\frac{\partial u}{\partial x}\,(t,a) = \mu(t) \tag{4.10}$$

prescribes the heat flux $w(t,a) = -\kappa(a)u_x(t,a)$ there. In particular, a homogeneous Neumann condition, $u_x(t,a) \equiv 0$, models an insulated end that prevents thermal energy flowing in or out. The *Robin*[†] *boundary condition*,

$$\frac{\partial u}{\partial x}\,(t,a) + \beta(t)\,u(t,a) = \tau(t), \tag{4.11}$$

models the heat exchange resulting from the end of the bar being placed in a heat bath (thermal reservoir) at temperature $\tau(t)$.

Each end of the bar is required to satisfy one of these boundary conditions. For example, a bar with both ends having prescribed temperatures is governed by the pair of Dirichlet boundary conditions

$$u(t,a) = \alpha(t), \qquad u(t,b) = \beta(t), \tag{4.12}$$

[†] Since it is named after the nineteenth-century French analyst Victor Gustave Robin, the pronunciation should be with a French accent.

whereas a bar with two insulated ends requires two homogeneous Neumann boundary conditions

$$\frac{\partial u}{\partial x}(t,a) = 0, \qquad \frac{\partial u}{\partial x}(t,b) = 0. \tag{4.13}$$

Mixed boundary conditions, with one end at a fixed temperature and the other insulated, are similarly formulated, e.g.,

$$u(t,a) = \alpha(t), \qquad \frac{\partial u}{\partial x}(t,b) = 0. \tag{4.14}$$

Finally, the *periodic boundary conditions*

$$u(t,a) = u(t,b), \qquad \frac{\partial u}{\partial x}(t,a) = \frac{\partial u}{\partial x}(t,b), \tag{4.15}$$

correspond to a circular *ring* obtained by joining the two ends of the bar. As before, we are assuming that the heat is allowed to flow only around the ring — insulation prevents the radiation of heat from one side of the ring affecting the other side.

The Heat Equation

In this book, we will retain the term "heat equation" to refer to the case in which the bar is composed of a uniform material, and so its density ρ, conductivity κ, and specific heat χ are all positive constants. We also exclude external heat sources (other than at the endpoints), meaning that the bar remains insulated along its entire length. Under these assumptions, the general diffusion equation (4.6) reduces to the homogeneous *heat equation*

$$\frac{\partial u}{\partial t} = \gamma \, \frac{\partial^2 u}{\partial x^2} \tag{4.16}$$

for the temperature $u(t,x)$ at time t and position x. The constant

$$\gamma = \frac{\kappa}{\sigma} = \frac{\kappa}{\rho \, \chi} \tag{4.17}$$

is called the *thermal diffusivity*; it incorporates all of the bar's relevant physical properties. The solution $u(t,x)$ will be uniquely prescribed once we specify initial conditions (4.8) and a suitable boundary condition at both of its endpoints.

As we learned in Section 3.1, the separable solutions to the heat equation are based on the exponential ansatz[†]

$$u(t,x) = e^{-\lambda t} \, v(x), \tag{4.18}$$

where $v(x)$ depends only on the spatial variable. Functions of this form, which "separate" into a product of a function of t times a function of x, are known as *separable solutions*. Substituting (4.18) into (4.16) and canceling the common exponential factors, we find that $v(x)$ must solve the second-order linear ordinary differential equation

$$-\gamma \, \frac{d^2 v}{dx^2} = \lambda \, v.$$

[†] Anticipating the eventual signs of the eigenvalues, and to facilitate later discussions, we now include a minus sign in the exponential term.

Each nontrivial solution $v(x) \not\equiv 0$ is an *eigenfunction*, with associated *eigenvalue* λ, for the linear differential operator $L[v] = -\gamma\, v''(x)$. With the separable eigensolutions (4.18) in hand, we will then be able to reconstruct the desired solution $u(t, x)$ as a linear combination, or rather infinite series, thereof.

Let us concentrate on the simplest case: a uniform, insulated bar of length ℓ that is held at zero temperature at both ends. We specify its initial temperature $f(x)$ at time $t_0 = 0$, and so the relevant initial and boundary conditions are

$$\begin{aligned} u(t,0) = 0, \qquad u(t,\ell) = 0, \qquad & t \geq 0, \\ u(0,x) = f(x), \qquad & 0 \leq x \leq \ell. \end{aligned} \tag{4.19}$$

The eigensolutions (4.18) are found by solving the Dirichlet boundary value problem

$$\gamma\, \frac{d^2 v}{dx^2} + \lambda\, v = 0, \qquad v(0) = 0, \qquad v(\ell) = 0. \tag{4.20}$$

By direct calculation (as you are asked to do in Exercises 4.1.19–20), one finds that if λ is either complex, or real and nonpositive, then the only solution to the boundary value problem (4.20) is the trivial solution $v(x) \equiv 0$. This means that all the eigenvalues must necessarily be real and positive. In fact, the reality and positivity of the eigenvalues need not be explicitly checked. Rather, they follow from very general properties of positive definite boundary value problems, of which (4.20) is a particular case. See Section 9.5 for the underlying theory and Theorem 9.34 for the relevant result.

When $\lambda > 0$, the general solution to the differential equation is a trigonometric function

$$v(x) = a \cos \omega x + b \sin \omega x, \qquad \text{where} \qquad \omega = \sqrt{\lambda/\gamma}\,,$$

and a and b are arbitrary constants. The first boundary condition requires $v(0) = a = 0$. This serves to eliminate the cosine term, and then the second boundary condition requires

$$v(\ell) = b \sin \omega \ell = 0.$$

Therefore, since we require $b \neq 0$ — otherwise, the solution is trivial and does not qualify as an eigenfunction — $\omega \ell$ must be an integer multiple of π, and so

$$\omega = \frac{\pi}{\ell}, \qquad \frac{2\pi}{\ell}, \qquad \frac{3\pi}{\ell}, \qquad \dots .$$

We conclude that the eigenvalues and eigenfunctions of the boundary value problem (4.20) are

$$\lambda_n = \gamma \left(\frac{n\pi}{\ell}\right)^2, \qquad v_n(x) = \sin \frac{n\pi x}{\ell}, \qquad n = 1, 2, 3, \dots . \tag{4.21}$$

The corresponding eigensolutions (4.18) are

$$u_n(t,x) = \exp\left(-\frac{\gamma\, n^2 \pi^2 t}{\ell^2}\right) \sin \frac{n\pi x}{\ell}, \qquad n = 1, 2, 3, \dots . \tag{4.22}$$

Each represents a trigonometrically oscillating temperature profile that maintains its form while decaying to zero at an exponentially fast rate.

To solve the general initial value problem, we assemble the eigensolutions into an infinite series,

$$u(t,x) = \sum_{n=1}^{\infty} b_n\, u_n(t,x) = \sum_{n=1}^{\infty} b_n \exp\left(-\frac{\gamma\, n^2 \pi^2 t}{\ell^2}\right) \sin \frac{n\pi x}{\ell}, \tag{4.23}$$

whose coefficients b_n are to be fixed by the initial conditions. Indeed, assuming that the series converges, the initial temperature profile is

$$u(0,x) = \sum_{n=1}^{\infty} b_n \sin \frac{n\pi x}{\ell} = f(x). \tag{4.24}$$

This has the form of a Fourier sine series (3.52) on the interval $[0, \ell]$. Thus, the coefficients are determined by the Fourier formulae (3.53), and so

$$b_n = \frac{2}{\ell} \int_0^{\ell} f(x) \sin \frac{n\pi x}{\ell}\, dx, \qquad n = 1, 2, 3, \dots . \tag{4.25}$$

The resulting formula (4.23) describes the Fourier sine series for the temperature $u(t,x)$ of the bar at each later time $t \geq 0$.

Example 4.1. Consider the initial temperature profile

$$u(0,x) = f(x) = \begin{cases} -x, & 0 \leq x \leq \frac{1}{5}, \\ x - \frac{2}{5}, & \frac{1}{5} \leq x \leq \frac{7}{10}, \\ 1 - x, & \frac{7}{10} \leq x \leq 1, \end{cases} \tag{4.26}$$

on a bar of length 1, plotted in the first graph in Figure 4.1. Using (4.25), the first few Fourier coefficients of $f(x)$ are computed (by either exact or numerical integration) to be

$$\begin{array}{llll} b_1 \approx .0897, & b_2 \approx -.1927, & b_3 \approx -.0289, & b_4 = 0, \\ b_5 \approx -.0162, & b_6 \approx .0132, & b_7 \approx .0104, & b_8 = 0, \end{array} \qquad \dots .$$

The resulting Fourier series solution to the heat equation is

$$\begin{aligned} u(t,x) &= \sum_{n=1}^{\infty} b_n\, u_n(t,x) = \sum_{n=1}^{\infty} b_n\, e^{-\gamma n^2 \pi^2 t} \sin n\pi x \\ &\approx .0897\, e^{-\gamma \pi^2 t} \sin \pi x - .1927\, e^{-4\gamma \pi^2 t} \sin 2\pi x - .0289\, e^{-9\gamma \pi^2 t} \sin 3\pi x - \cdots . \end{aligned}$$

In Figure 4.1, the solution, for $\gamma = 1$, is plotted at some representative times. Observe that the corners in the initial profile are immediately smoothed out. As time progresses, the solution decays, at a fast exponential rate of $e^{-\pi^2 t} \approx e^{-9.87 t}$, to a uniform, zero temperature, which is the equilibrium temperature distribution for the homogeneous Dirichlet boundary conditions. As the solution decays to thermal equilibrium, the higher Fourier modes rapidly disappear, and the solution assumes the progressively more symmetric shape of a single sine arc, of rapidly decreasing amplitude.

Smoothing and Long–Time Behavior

The fact that we can write the solution to an initial-boundary value problem in the form of an infinite series (4.23) is progress of a sort. However, because we are unable to sum the series in closed form, this "solution" is much less satisfying than a direct, explicit formula. Nevertheless, there are important qualitative and quantitative features of the solution that can be easily gleaned from such series expansions.

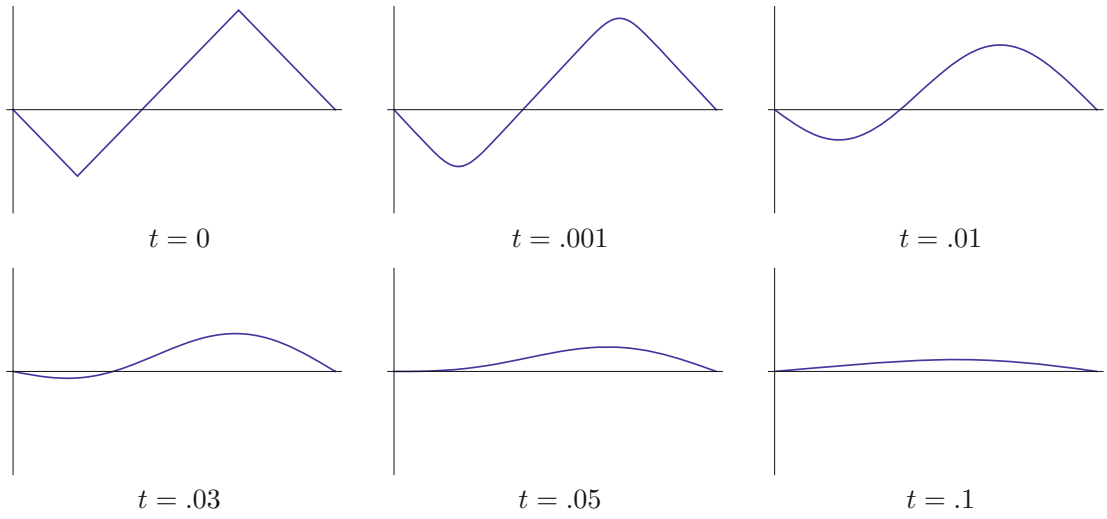


Figure 4.1. A solution to the heat equation.

If the initial data $f(x)$ is integrable (e.g., piecewise continuous), then its Fourier coefficients are uniformly bounded; indeed, for any $n \geq 1$,

$$|\, b_n \,| \;\leq\; \frac{2}{\ell} \int_0^\ell \left| f(x) \sin \frac{n\pi x}{\ell} \right| dx \;\leq\; \frac{2}{\ell} \int_0^\ell |\, f(x) \,|\, dx \equiv M. \tag{4.27}$$

This property holds even for quite irregular data. Under these conditions, each term in the series solution (4.23) is bounded by an exponentially decaying function

$$\left| \; b_n \exp\left(-\frac{\gamma\, n^2\, \pi^2}{\ell^2}\, t \right) \sin \frac{n\pi x}{\ell} \; \right| \;\leq\; M \exp\left(-\frac{\gamma\, n^2\, \pi^2}{\ell^2}\, t \right).$$

This means that, as soon as $t > 0$, most of the high-frequency terms, $n \gg 0$, will be extremely small. Only the first few terms will be at all noticeable, and so the solution essentially degenerates into a finite sum over the first few Fourier modes. As time increases, more and more of the Fourier modes will become negligible, and the sum further degenerates into fewer and fewer significant terms. Eventually, as $t \to \infty$, *all* of the Fourier modes will decay to zero. Therefore, the solution will converge exponentially fast to a zero temperature profile: $u(t, x) \to 0$ as $t \to \infty$, representing the bar in its final uniform thermal equilibrium. The fact that its equilibrium temperature is zero is the result of holding both ends of the bar fixed at zero temperature, whereby any initial thermal energy is eventually dissipated away through the ends. The small-scale temperature fluctuations tend to rapidly cancel out through diffusion of thermal energy, and the last term to disappear is the one with the slowest decay, namely

$$u(t, x) \;\approx\; b_1 \exp\left(-\frac{\gamma\, \pi^2}{\ell^2}\, t \right) \sin \frac{\pi x}{\ell}\,, \qquad \text{where} \qquad b_1 = \frac{1}{\pi} \int_0^\pi f(x) \sin x \, dx. \tag{4.28}$$

For generic initial data, the coefficient $b_1 \neq 0$, and the solution approaches thermal equilibrium at an exponential rate prescribed by the smallest eigenvalue, $\lambda_1 = \gamma\, \pi^2/\ell^2$, which is proportional to the thermal diffusivity divided by the square of the length of the bar. The

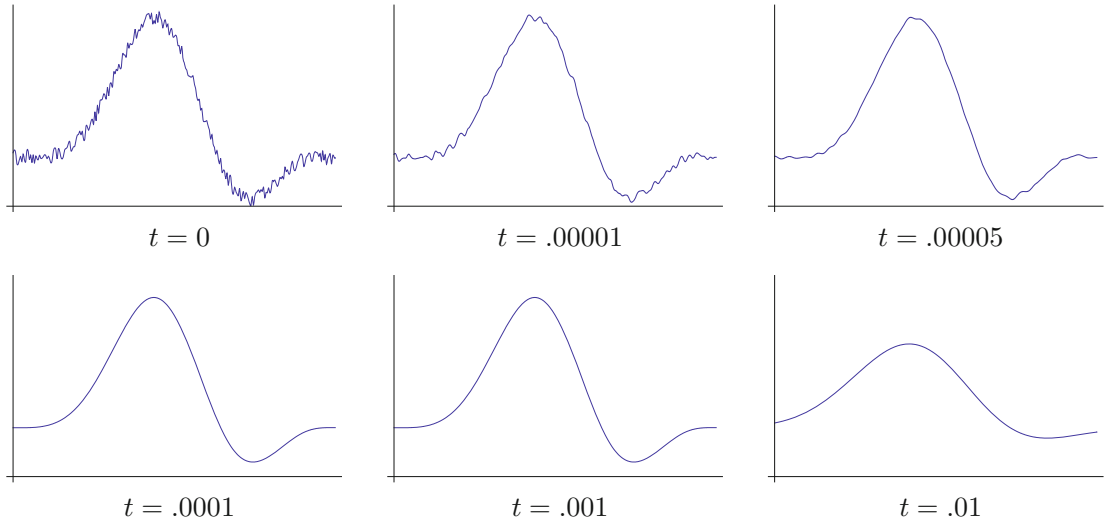


Figure 4.2. Denoising a signal with the heat equation.

longer the bar, or the smaller the diffusivity, the longer it takes for the effect of holding the ends at zero temperature to propagate along its entire length. Also, again provided $b_1 \neq 0$, the asymptotic shape of the temperature profile is a small, exponentially decaying sine arc, just as we observed in Example 4.1. In exceptional situations, namely when $b_1 = 0$, the solution decays even faster, at a rate equal to the eigenvalue $\lambda_k = \gamma\, k^2\, \pi^2/\ell^2$ corresponding to the first nonzero term, $b_k \neq 0$, in the Fourier series; its asymptotic shape now oscillates k times over the interval.

Another, closely related, observation is that, for any fixed time $t > 0$ after the initial moment, the coefficients in the Fourier sine series (4.23) decay exponentially fast as $n \to \infty$. According to the discussion at the end of Section 3.3, this implies that the Fourier series converges to an infinitely differentiable function of x at each positive time t, *no matter how unsmooth the initial temperature profile*. We have discovered the basic smoothing property of heat flow, which we state for a general initial time t_0.

Theorem 4.2. *If $u(t, x)$ is a solution to the heat equation with piecewise continuous initial data $f(x) = u(t_0, x)$, or, more generally, initial data satisfying* (4.27)*, then, for any $t > t_0$, the solution $u(t, x)$ is an infinitely differentiable function of x.*

In other words, the heat equation *instantaneously* smoothes out any discontinuities and corners in the initial temperature profile by fast damping of the high-frequency modes. The heat equation's effect on irregular initial data underlies its effectiveness for smoothing and denoising signals. We take the initial data $u(0, x) = f(x)$ to be a noisy signal, and then evolve the heat equation forward to a prescribed time $t^\star > 0$. The resulting function $g(x) = u(t^\star, x)$ will be a smoothed version of the original signal $f(x)$ in which most of the high-frequency noise has been eliminated. Of course, if we run the heat flow for too long, all of the low-frequency features will also be smoothed out and the result will be a uniform, constant signal. Thus, the choice of *stopping time* $t^\star$ is crucial to the success

of this method. Figure 4.2 shows the effect of running the heat equation,† with $\gamma = 1$, on a signal that has been contaminated by random noise. Observe how quickly the noise is removed. By the final time, the overall smoothing effect of the heat flow has caused significant degradation (blurring) of the original signal. The heat equation approach to denoising has the advantage that no Fourier coefficients need be explicitly computed, nor does one need to reconstruct the smoothed signal. Basic numerical solution schemes for the heat equation are to be discussed in Chapter 5.

An important theoretical consequence of the smoothing property is that diffusion is a one-way process — one cannot run time backwards and accurately infer what a temperature distribution looked like in the past. In particular, if the initial data $u(0, x) = f(x)$ is not smooth, then the value of $u(t, x)$ for any $t < 0$ cannot be defined, because if $u(t_0, x)$ were defined and integrable at some $t_0 < 0$ then, by Theorem 4.2, $u(t, x)$ would be smooth at all subsequent times $t > t_0$, including $t = 0$, in contradiction to our assumption. Moreover, for most initial data, the Fourier coefficients in the solution formula (4.23) are, at any $t < 0$, exponentially *growing* as $n \to \infty$, indicating that high-frequency noise has completely overwhelmed the solution, thereby precluding any kind of convergence of the Fourier series.

Mathematically, we can reverse future and past by changing t to $-t$. In the differential equation, this merely reverses the sign of the time-derivative term; the x derivatives are unaffected. Thus, by the above reasoning, the *backwards heat equation*

$$\frac{\partial u}{\partial t} = -\gamma \frac{\partial^2 u}{\partial x^2}, \qquad \text{with a negative diffusion coefficient} \qquad -\gamma < 0, \tag{4.29}$$

is an *ill-posed problem* in the sense that small changes in the initial data — e.g., a small perturbation of a high-frequency mode — can produce arbitrarily large changes in the solution arbitrarily close to the initial time. In other words, the solution does not depend continuously on the initial data. Even worse, for nonsmooth initial data, the solution is not even well defined in forwards time $t > 0$ (although it is well-posed if we run t backwards). The same holds for more general diffusion processes, e.g., (4.6). If, as in all physically relevant cases, the coefficient of u_{xx} is everywhere positive, then the initial value problem is well-posed for $t > 0$, but ill-posed for $t < 0$. On the other hand, if the coefficient is everywhere negative, the reverse holds. A coefficient that changes signs would cause the differential equation to be ill-posed in both directions.

While theoretically undesirable, the unsmoothing effect of the backwards heat equation has potential benefits in certain contexts. For example, in image processing, diffusion will gradually blur an image by damping out the high-frequency modes. Image enhancement is the reverse process, and can be based on running the heat flow backwards in some stable manner. In forensics, determining the time of death based on the current temperature of a corpse also requires running the equations governing the dissipation of body heat backwards in time. One option would be to restrict the backwards evolution to the first few Fourier modes, which prevents the small-scale fluctuations from overwhelming the computation. Ill-posed problems also arise in the reconstruction of subterranean profiles from seismic data, a central problem of the oil and gas industry. These and other applications are driving contemporary research into how to cleverly circumvent the ill-posedness of backwards diffusion processes.

† To avoid artifacts at the ends of the interval, we are, in fact, using periodic boundary conditions in the plots. Away from the ends, running the equation with Dirichlet boundary conditions leads to almost identical results.

Remark: The irreversibility of the heat equation, along with the irreversibility of nonlinear transport in the presence of shock waves discussed in Section 2.3, highlight a crucial distinction between partial differential equations and ordinary differential equations. Ordinary differential equations are always reversible — the existence, uniqueness, and continuous dependence properties of solutions are all equally valid in reverse time (although their detailed qualitative and quantitative properties will, of course, depend upon whether time is running forwards or backwards). The irreversibility and ill-posedness of partial differential equations modeling thermodynamical, biological, and other diffusive processes in our universe may explain why Time's Arrow points exclusively to the future.

The Heated Ring Redux

Let us next consider the periodic boundary value problem modeling heat flow in an insulated circular ring. We fix the length of the ring to be $\ell = 2\pi$, with $-\pi \le x \le \pi$ representing the "angular" coordinate around the ring. For simplicity, we also choose units in which the thermal diffusivity is $\gamma = 1$. Thus, we seek to solve the heat equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}\,, \qquad -\pi < x < \pi, \qquad t > 0, \tag{4.30}$$

subject to periodic boundary conditions

$$u(t,-\pi) = u(t,\pi), \qquad \frac{\partial u}{\partial x}\,(t,-\pi) = \frac{\partial u}{\partial x}\,(t,\pi), \qquad t \ge 0, \tag{4.31}$$

that ensure continuity of the solution when the angular coordinate switches from $-\pi$ to π. The initial temperature distribution is

$$u(0,x) = f(x), \qquad -\pi < x \le \pi. \tag{4.32}$$

The resulting temperature $u(t,x)$ will be a periodic function in x of period 2π.

Substituting the separable solution ansatz (3.15) into the heat equation and the boundary conditions results in the periodic eigenvalue problem

$$\frac{d^2 v}{dx^2} + \lambda\, v = 0, \qquad v(-\pi) = v(\pi), \qquad v'(-\pi) = v'(\pi). \tag{4.33}$$

As we already noted in Section 3.1, the eigenvalues of this particular boundary value problem are $\lambda_n = n^2$, where $n = 0, 1, 2, \dots$ is a nonnegative integer; the corresponding eigenfunctions are the trigonometric functions

$$v_n(x) = \cos n\,x, \qquad \widetilde{v}_n(x) = \sin n\,x, \qquad n = 0, 1, 2, \dots\,.$$

Note that $\lambda_0 = 0$ is a simple eigenvalue, with constant eigenfunction $\cos 0\,x = 1$ — the sine solution $\sin 0\,x \equiv 0$ is trivial — while the positive eigenvalues are, in fact, double, each possessing two linearly independent eigenfunctions. The corresponding eigensolutions to the heated ring equation (4.30–31) are

$$u_n(t,x) = e^{-n^2 t}\cos n\,x, \qquad \widetilde{u}_n(t,x) = e^{-n^2 t}\sin n\,x, \qquad n = 0, 1, 2, 3, \dots\,.$$

The resulting infinite series solution is

$$u(t,x) = \tfrac{1}{2}\,a_0 + \sum_{n=1}^{\infty} \big(\, a_n\, e^{-n^2 t}\cos n\,x + b_n\, e^{-n^2 t}\sin n\,x \,\big), \tag{4.34}$$

with as yet unspecified coefficients a_n, b_n. The initial conditions require

$$u(0,x) = \tfrac{1}{2}\,a_0 + \sum_{n=1}^{\infty} (a_n \cos n\,x + b_n \sin n\,x) = f(x), \tag{4.35}$$

which is precisely the complete Fourier series (3.34) of the initial temperature profile $f(x)$. Consequently,

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos n\,x\,dx, \qquad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin n\,x\,dx, \tag{4.36}$$

are its usual Fourier coefficients (3.35).

As in the Dirichlet problem, after the initial instant, the high-frequency terms in the series (4.34) become extremely small, since $e^{-n^2 t} \ll 1$ for $n \gg 0$. Therefore, as soon as $t > 0$, the solution instantaneously becomes smooth, and quickly degenerates into what is in essence a finite sum over the first few Fourier modes. Moreover, as $t \to \infty$, *all* of the Fourier modes will decay to zero with the exception of the constant mode, associated with the null eigenvalue $\lambda_0 = 0$. Consequently, the solution will converge, at an exponential rate, to a constant-temperature profile,

$$u(t,x) \;\longrightarrow\; \tfrac{1}{2}\,a_0 \;=\; \frac{1}{2\,\pi}\int_{-\pi}^{\pi} f(x)\,dx,$$

which equals the *average* of the initial temperature profile. In physical terms, since the insulation prevents any thermal energy from escaping the ring, it rapidly redistributes itself so that the ring achieves a uniform constant temperature — its eventual equilibrium state.

Prior to attaining equilibrium, only the very lowest frequency Fourier modes will still be noticeable, and so the solution will asymptotically look like

$$u(t,x) \approx \tfrac{1}{2}\,a_0 + e^{-t}\,(a_1 \cos x + b_1 \sin x) = \tfrac{1}{2}\,a_0 + r_1\,e^{-t}\cos(x + \delta_1), \tag{4.37}$$

where

$$a_1 = r_1 \cos\delta_1 = \frac{1}{2\,\pi}\int_{-\pi}^{\pi} f(x)\cos x\,dx, \qquad b_1 = r_1 \sin\delta_1 = \frac{1}{2\,\pi}\int_{-\pi}^{\pi} f(x)\sin x\,dx.$$

Thus, for most initial data, the solution approaches thermal equilibrium at an exponential rate of e^{-t}. The exceptions are when $a_1 = b_1 = 0$, for which the rate of convergence is even faster, namely at a rate $e^{-k^2 t}$, where k is the smallest integer such that at least one of the k^{th} order Fourier coefficients a_k, b_k is nonzero.

In fact, once we are convinced that the bar must tend to thermal equilibrium as $t \to \infty$, we can predict the final temperature without knowing the explicit solution formula. Our derivation in Section 4.1 implies that the heat equation has the form of a conservation law (4.1), with the conserved density being the temperature $u(t,x)$. As in (4.2), the integrated form of the conservation law reads

$$\begin{aligned}\frac{d}{dt}\int_{-\pi}^{\pi} u(t,x)\,dx = \int_{-\pi}^{\pi} \frac{\partial u}{\partial t}\,(t,x)\,dx &= \gamma \int_{-\pi}^{\pi} \frac{\partial^2 u}{\partial x^2}\,(t,x)\,dx \\ &= \gamma\left[\,\frac{\partial u}{\partial x}\,(t,\pi) - \frac{\partial u}{\partial x}\,(t,-\pi)\,\right] = 0,\end{aligned}$$

where the flux terms cancel thanks to the periodic boundary conditions (4.31). Physically, any flux out of one end of the circular bar is immediately fed into the other, abutting end,

and so there is no net loss of thermal energy. We conclude that, for the periodic boundary value problem, the total *thermal energy*

$$E(t) = \int_{-\pi}^{\pi} u(t,x)\,dx = \text{constant} \tag{4.38}$$

remains constant for all time. (In contrast, the thermal energy does *not* remain constant for the Dirichlet boundary value problem, decaying steadily to 0 due to the out-flux of heat through the ends of the bar; see Exercise 4.1.13 for further details.)

Remark: More correctly, according to (4.3), the thermal energy is obtained by multiplying the temperature by the product, $\sigma = \rho\,\chi$, of the density and the specific heat of the body. For the heat equation, both are constant, and so the physical thermal energy equals $\sigma\,E(t)$. Mathematically, we can safely ignore this extra constant factor, or, equivalently, work in physical units in which $\sigma = 1$. This does not extend to nonuniform bodies, whose *thermal energy* is given by $E(t) = \int_{-\pi}^{\pi} \sigma(x)\,u(t,x)\,dx$, and whose constancy, under suitable boundary conditions, follows from the conservation-law form (4.6) of the linear diffusion equation.

In general, a system is in (static) *equilibrium* if it remains unaltered as time progresses. Thus, any equilibrium configuration has the form $u = u^\star(x)$, and hence satisfies $\partial u^\star/\partial t = 0$. If, in addition, $u^\star(x)$ is an equilibrium solution to the periodic heat equation (4.30–33), then it must satisfy

$$\frac{\partial u^\star}{\partial t} = 0 = \frac{\partial^2 u^\star}{\partial x^2}\,, \qquad u^\star(-\pi) = u^\star(\pi), \qquad \frac{\partial u^\star}{\partial x}\,(-\pi) = \frac{\partial u^\star}{\partial x}\,(\pi). \tag{4.39}$$

In other words, $u^\star$ is a solution to the periodic boundary value problem (4.33) for the null eigenvalue $\lambda = 0$. Thus, *the null eigenfunctions* (*including the zero solution*) *are all the possible equilibrium solutions*. In particular, for the periodic boundary value problem, the null eigenfunctions are constant, and therefore solutions to the periodic heat equation will tend to a constant equilibrium temperature.

Now, once we know that the solution tends to a constant, $u(t,x) \to a$ as $t \to \infty$, then its thermal energy tends to

$$E(t) = \int_{-\pi}^{\pi} u(t,x)\,dx \;\longrightarrow\; \int_{-\pi}^{\pi} a\,dx = 2\,\pi\,a \qquad \text{as} \qquad t \longrightarrow \infty.$$

On the other hand, as we just demonstrated, the thermal energy is constant, so

$$E(t) = E(0) = \int_{-\pi}^{\pi} u(0,x)\,dx = \int_{-\pi}^{\pi} f(x)\,dx.$$

Combining these two, we conclude that

$$\int_{-\pi}^{\pi} f(x)\,dx = 2\,\pi\,a, \quad \text{and so the equilibrium temperature} \quad a = \frac{1}{2\,\pi}\int_{-\pi}^{\pi} f(x)\,dx$$

equals the average initial temperature. This reconfirms our earlier result, but avoids having to know an explicit series solution formula. As a result, the latter method can be applied to a much wider range of situations.

Inhomogeneous Boundary Conditions

So far, we have concentrated our attention on homogeneous boundary conditions. There is a simple trick that will convert a boundary value problem with inhomogeneous but constant Dirichlet boundary conditions,

$$\frac{\partial u}{\partial t} = \gamma\, \frac{\partial^2 u}{\partial x^2}\,, \qquad u(t,0) = \alpha, \qquad u(t,\ell) = \beta, \qquad t \geq 0, \tag{4.40}$$

into a homogeneous Dirichlet problem. We begin by solving for the equilibrium temperature profile. As in (4.39), the equilibrium does not depend on t and hence satisfies the boundary value problem

$$\frac{\partial u^\star}{\partial t} = 0 = \gamma\, \frac{\partial^2 u^\star}{\partial x^2}\,, \qquad u^\star(0) = \alpha, \qquad u^\star(\ell) = \beta.$$

Solving the ordinary differential equation yields $u^\star(x) = a + b\,x$, where the constants a, b are fixed by the boundary conditions. We conclude that the equilibrium solution is a straight line connecting the boundary values:

$$u^\star(x) = \alpha + \frac{\beta - \alpha}{\ell}\, x. \tag{4.41}$$

The difference

$$\widetilde{u}(t,x) = u(t,x) - u^\star(x) = u(t,x) - \alpha - \frac{\beta - \alpha}{\ell}\, x \tag{4.42}$$

measures the deviation of the solution from equilibrium. It clearly satisfies the homogeneous boundary conditions at both ends:

$$\widetilde{u}(t,0) = 0 = \widetilde{u}(t,\ell).$$

Moreover, by linearity, since both $u(t,x)$ and $u^\star(x)$ are solutions to the heat equation, so is $\widetilde{u}(t,x)$. The initial data must be similarly adapted:

$$\widetilde{u}(0,x) = u(t,x) - u^\star(x) = f(x) - \alpha - \frac{\beta - \alpha}{\ell}\, x \equiv \widetilde{f}(x). \tag{4.43}$$

Solving the resulting homogeneous initial-boundary value problem, we write $\widetilde{u}(t,x)$ in Fourier series form (4.23), where the Fourier coefficients are specified by the modified initial data $\widetilde{f}(x)$ in (4.43). The solution to the inhomogeneous boundary value problem thus has the series form

$$u(t,x) = \alpha + \frac{\beta - \alpha}{\ell}\, x \;+\; \sum_{n=1}^{\infty} \widetilde{b}_n \exp\left(-\frac{\gamma\, n^2\, \pi^2}{\ell^2}\, t \right) \sin \frac{n\,\pi\, x}{\ell}\,, \tag{4.44}$$

where

$$\widetilde{b}_n = \frac{2}{\ell} \int_0^\ell \widetilde{f}(x)\, \sin \frac{n\,\pi\, x}{\ell}\, dx, \qquad n = 1, 2, 3, \ldots\,. \tag{4.45}$$

Since $\widetilde{u}(t,0)$ decays to zero at an exponential rate as $t \to \infty$, the actual temperature profile (4.44) will asymptotically decay to the equilibrium profile,

$$u(t,x) \quad \longrightarrow \quad u^\star(x) = \alpha \;+\; \frac{\beta - \alpha}{\ell}\, x,$$

at the same exponentially fast rate, governed by the first eigenvalue $\lambda_1 = \pi^2/\ell^2$ — unless $\widetilde{b}_1 = 0$, in which case the decay rate is even faster.

This method does not work as well when the boundary conditions are time-dependent:

$$u(t,0) = \alpha(t), \qquad u(t,\ell) = \beta(t).$$

Attempting to mimic the preceding technique, we discover that the deviation[†]

$$\widetilde{u}(t,x) = u(t,x) - u^\star(t,x), \qquad \text{where} \qquad u^\star(t,x) = \alpha(t) + \frac{\beta(t) - \alpha(t)}{\ell}\, x, \tag{4.46}$$

satisfies the homogeneous boundary conditions, but now solves an inhomogeneous or forced version of the heat equation:

$$\frac{\partial \widetilde{u}}{\partial t} = \frac{\partial^2 \widetilde{u}}{\partial x^2} + h(t,x), \quad \text{where} \quad h(t,x) = -\,\frac{\partial u^\star}{\partial t}\,(t,x) = -\,\alpha'(t) - \frac{\beta'(t) - \alpha'(t)}{\ell}\, x. \tag{4.47}$$

Solution techniques for the latter partial differential equation will be discussed in Section 8.1 below.

Robin Boundary Conditions

Consider a bar of unit length and unit thermal diffusivity, insulated along its length, which has one of its ends held at 0° and the other put in a heat bath. The resulting thermodynamics are modeled by the heat equation subject to Dirichlet boundary conditions at $x = 0$ and Robin boundary conditions at $x = 1$:

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}\,, \qquad u(t,0) = 0, \qquad \frac{\partial u}{\partial x}\,(t,1) + \beta\, u(t,1) = 0, \tag{4.48}$$

where $\beta \neq 0$ is a constant[‡] that measures the rate of transfer of thermal energy, with $\beta > 0$ when the bath is cold and so the energy is being extracted from the bar. As before, the general solution to the resulting initial-boundary value problem can be assembled from the separable eigensolutions based on our usual exponential ansatz $u(t,x) = e^{-\lambda t}\, v(x)$. Substituting this expression into (4.48), we find that the eigenfunction $v(x)$ must satisfy the boundary value problem

$$-\,\frac{d^2 v}{dx^2} = \lambda\, v, \qquad v(0) = 0, \qquad v'(1) + \beta\, v(1) = 0. \tag{4.49}$$

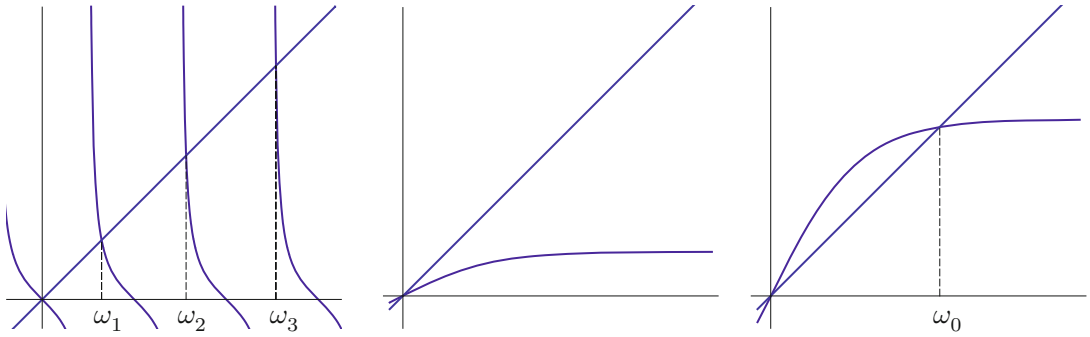
In order to find nontrivial solutions $v(x) \not\equiv 0$ to (4.49), let us first assume $\lambda = \omega^2 > 0$, where, without loss of generality, $\omega > 0$. The solution to the ordinary differential equation that satisfies the Dirichlet boundary condition at $x = 0$ is a constant multiple of $v(x) = \sin \omega\, x$. Substituting this function into the Robin boundary condition at $x = 1$, we find

$$\omega \cos \omega + \beta \sin \omega = 0, \qquad \text{or, equivalently,} \qquad \omega = -\,\beta \tan \omega. \tag{4.50}$$

It is not hard to see that there is an infinite number of real, positive solutions $0 < \omega_1 < \omega_2 < \omega_3 < \cdots \to \infty$ to the latter transcendental equation. Indeed, they can be characterized as the abscissas $\omega_n > 0$ of the intersection points of the graphs of the two functions $f(\omega) = \omega$

[†] In this case, $u^\star(t,x)$ is not an equilibrium solution. Indeed, we do not expect the bar to go to equilibrium if the temperature of its endpoints is constantly changing.

[‡] The case $\beta = 0$ reduces to the mixed boundary value problem, whose analysis is left to the reader.



$\lambda > 0, \ \beta = 1.$ $\lambda < 0, \ \beta = -.5.$ $\lambda < 0, \ \beta = -2.$

Figure 4.3. Eigenvalue equation for Robin boundary conditions.

and $g(\omega) = -\beta \tan \omega$, as shown in the first plot in Figure 4.3. Each root ω_n defines a positive eigenvalue $\lambda_n = \omega_n^2 > 0$ to the boundary value problem (4.49) and hence an exponentially decaying eigensolution

$$u_n(t, x) = e^{-\lambda_n t} \sin \omega_n x \tag{4.51}$$

to the Robin boundary value problem (4.48). While there is no explicit formula, numerical approximations to the eigenvalues are easily found via a numerical root finder, e.g., Newton's Method, [**24**, **94**]. In particular, for $\beta = 1$, the first three eigenvalues are $\lambda_1 = \omega_1^2 \approx 4.1159$, $\lambda_2 = \omega_2^2 \approx 24.1393$, $\lambda_3 = \omega_3^2 \approx 63.6591$.

What about a zero eigenvalue? If $\lambda = 0$ in (4.49), then the solution to the ordinary differential equation that satisfies the Dirichlet boundary condition is a constant multiple of $v(x) = x$. This function satisfies the Robin boundary condition $v'(1) + \beta v(1) = 0$ if and only if $\beta = -1$. In this special configuration, the heat equation admits a time-independent eigensolution $u_0(t, x) = x$ with eigenvalue $\lambda_0 = 0$. Physically, the rate of transfer of thermal energy into the bar through its end in the heat bath is exactly enough to cancel the heat loss through the Dirichlet end, resulting in a steady-state solution. All other eigenmodes correspond to positive eigenvalues, and hence are exponentially decaying. The general solution decays to the steady state, which is a constant multiple of the null eigensolution: $u(t, x) \to c\, x$ as $t \to \infty$, at an exponential rate prescribed, generically, by the first positive eigenvalue $\lambda_1 > 0$.

However, in contrast to the more common types of boundary conditions (Dirichlet, Neumann, mixed, periodic), we cannot automatically rule out the existence of negative eigenvalues in the Robin case. Suppose $\lambda = -\omega^2 < 0$ with $\omega > 0$. Now the solution to (4.49) that satisfies the Dirichlet boundary condition at $x = 0$ is a constant multiple of the hyperbolic sine function $v(x) = \sinh \omega\, x$. Substituting this expression into the Robin boundary condition at $x = 1$ produces

$$\omega \cosh \omega + \beta \sinh \omega = 0, \qquad \text{or, equivalently,} \qquad \omega = -\beta \tanh \omega, \tag{4.52}$$

where

$$\tanh \omega = \frac{\sinh \omega}{\cosh \omega} = \frac{e^{\omega} - e^{-\omega}}{e^{\omega} + e^{-\omega}} \tag{4.53}$$

is the hyperbolic tangent. If $\beta > -1$, there are no solutions $\omega > 0$ to this transcendental equation, and in this case all the eigenvalues are strictly positive and all solutions to the

heat equation are exponentially decaying. On the other hand, if $\beta < -1$, there is a single solution $\omega_0 > 0$, which produces a single negative eigenvalue $\lambda_0 = -\omega_0^2$. Representative graphs illustrating the two possibilities appear in Figure 4.3; in the first, the graph of $f(\omega) = \omega$ does not intersect the graph of $g(\omega) = \frac{1}{2}\tanh\omega$ when $\omega > 0$, whereas it intersects the graph of $\widehat{g}(\omega) = 2\tanh\omega$ at a single point, with abscissa $\omega_0 \approx 1.9150$, producing the negative eigenvalue $\lambda_0 \approx -\omega_0^2 \approx -3.6673$. Thus, when $\beta < -1$, there is, in addition to all the exponentially decaying eigenmodes associated with the positive eigenvalues, a single unstable exponentially *growing* eigenmode

$$u_0(t,x) = e^{\lambda_0 t}\,\sinh\omega_0\, x. \tag{4.54}$$

Physically, $\beta < -1$ implies that thermal energy is entering the Robin end of the bar at a faster rate than can be removed through the Dirichlet end, and hence the bar experiences an exponential increase in its overall temperature.

Remark: Even though some Robin boundary conditions admit exponentially growing solutions, and hence lead to *unstable* dynamics, the initial-boundary value problem remains *well-posed* because the solution exists and is uniquely determined by the initial data, and, moreover, small changes in the initial conditions induce relatively small changes in the resulting solution on bounded time intervals.

The Root Cellar Problem

As a final example, we discuss a problem that involves analysis of the heat equation on a semi-infinite interval. The question is this: how deep should you dig a root cellar? In the prerefrigeration era, a root cellar was used to keep food cool in the summer, but not freeze in the winter. We assume that the temperature inside the Earth depends only on the depth and the time of year. Let $u(t,x)$ denote the deviation in the temperature from its annual mean at depth $x > 0$ and time t. We shall assume that the temperature at the Earth's surface, $x = 0$, fluctuates in a periodic manner; specifically, we set

$$u(t,0) = a\,\cos\omega\, t, \tag{4.55}$$

where the oscillatory frequency

$$\omega = \frac{2\,\pi}{365.25\text{ days}} = 2.0 \times 10^{-7}\text{sec}^{-1} \tag{4.56}$$

refers to yearly temperature variations. In this model, we shall ignore daily temperature fluctuations, since their effect is not significant below a very thin surface layer. At large depths the temperature is assumed to be unvarying:

$$u(t,x) \;\longrightarrow\; 0 \qquad \text{as} \qquad x \;\longrightarrow\; \infty, \tag{4.57}$$

where 0 refers to the mean temperature.

Thus, we must solve the heat equation on a semi-infinite bar $0 < x < \infty$, with time-dependent boundary conditions (4.55, 57) at the ends. The analysis will be simplified a little if we replace the cosine by a complex exponential, and so we look for a complex solution with boundary conditions

$$u(t,0) = a\,e^{\,\mathrm{i}\,\omega\, t}, \qquad \lim_{x\to\infty} u(t,x) = 0. \tag{4.58}$$

Let us try a separable solution of the form

$$u(t,x) = v(x)\, e^{\,\mathrm{i}\,\omega\, t}. \tag{4.59}$$

Substituting this expression into the heat equation $u_t = \gamma\, u_{xx}$ leads to

$$\mathrm{i}\,\omega\, v(x)\, e^{\,\mathrm{i}\,\omega\, t} = \gamma\, v''(x)\, e^{\,\mathrm{i}\,\omega\, t}.$$

Canceling the common exponential factors, we conclude that $v(x)$ should solve the boundary value problem

$$\gamma\, v''(x) = \mathrm{i}\,\omega\, v, \qquad\qquad v(0) = a, \qquad\qquad \lim_{x\to\infty} v(x) = 0.$$

The solutions to the ordinary differential equation are

$$v_1(x) = e^{\sqrt{\mathrm{i}\,\omega/\gamma}\,x} = e^{\sqrt{\omega/(2\gamma)}\,(1+\mathrm{i}\,)\,x}, \qquad\qquad v_2(x) = e^{-\sqrt{\mathrm{i}\,\omega/\gamma}\,x} = e^{-\sqrt{\omega/(2\gamma)}\,(1+\mathrm{i}\,)\,x}.$$

The first solution is exponentially growing as $x \to \infty$, and so not germane to our problem. The solution to the boundary value problem must therefore be a multiple of the exponentially decaying solution:

$$v(x) = a\, e^{-\sqrt{\omega/(2\gamma)}\,(1+\mathrm{i}\,)\,x}.$$

Substituting back into (4.59), we find the (complex) solution to the root cellar problem to be

$$u(t,x) = a\, e^{-x\sqrt{\omega/(2\gamma)}}\, e^{\,\mathrm{i}\,(\omega\, t - \sqrt{\omega/(2\gamma)}\,x)}. \tag{4.60}$$

The corresponding real solution is obtained by taking the real part,

$$u(t,x) = a\, e^{-x\sqrt{\omega/(2\gamma)}}\, \cos\left(\omega\, t - \sqrt{\frac{\omega}{2\gamma}}\, x\right). \tag{4.61}$$

The first factor in (4.61) is exponentially decaying as a function of the depth. Thus, the further underground one is, the less noticeable is the effect of the surface temperature fluctuations. The second factor is periodic in time, with the same annual frequency ω. The interesting feature is that the temperature variations (4.61) are typically out of phase with respect to the surface temperature fluctuations, having an overall *phase lag* of

$$\delta = \sqrt{\frac{\omega}{2\gamma}}\, x$$

that depends linearly on the depth x. In particular, a cellar built at a depth where δ is an odd multiple of π will be completely out of phase, being hottest in the winter, and coldest in the summer. Thus, the (shallowest) ideal depth at which to build a root cellar would take $\delta = \pi$, corresponding to a depth of

$$x = \pi \sqrt{\frac{2\gamma}{\omega}}\,. \tag{4.62}$$

For typical soils in the Earth, $\gamma \approx 10^{-6}$ meters^2 sec^{-1}, and so, with ω given by (4.56), $x \approx 9.9$ meters. However, at this depth, the relative amplitude of the oscillations is

$$e^{-x\sqrt{\omega/2\gamma}} = e^{-\pi} = .04,$$

and hence there is only a 4% temperature fluctuation. In Minneapolis, the temperature varies, roughly, from $-40°$C to $+40°$C, and hence our 10-meter-deep root cellar would

experience only a 3.2°C annual temperature deviation from the winter, when it is the warmest, to the summer, when it is the coldest. Building the cellar twice as deep would lead to a temperature fluctuation of .2%, now in phase with the surface variations, which means that the cellar would be, for all practical purposes, at constant temperature year round.

Exercises

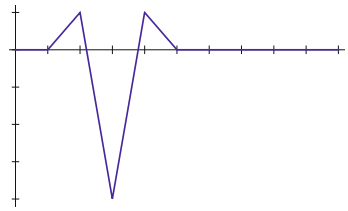
4.1.1. Suppose the ends of a bar of length 1 and thermal diffusivity $\gamma = 1$ are held fixed at respective temperatures 0° and 10°. (*a*) Determine the equilibrium temperature profile. (*b*) Determine the rate at which the equilibrium temperature profile is approached. (*c*) What does the temperature profile look like as it nears equilibrium?

4.1.2. A uniform insulated bar 1 meter long is stored at room temperature of 20° Celsius. An experimenter places one end of the bar in boiling water and the other end in ice water. (*a*) Set up an initial-boundary value problem that models the temperature in the bar. (*b*) Find the equilibrium temperature distribution. (*c*) Discuss how your answer depends on the material properties of the bar.

4.1.3. Consider the initial-boundary value problem

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \qquad \begin{aligned} &u(t,0) = 0 = u(t,10), && t > 0, \\ &u(0,x) = f(x), && 0 < x < 10, \end{aligned}$$

for the heat equation where the initial data has the following form:



$$f(x) = \begin{cases} x - 1, & 1 \le x \le 2, \\ 11 - 5x, & 2 \le x \le 3, \\ 5x - 19, & 3 \le x \le 4, \\ 5 - x, & 4 \le x \le 5, \\ 0, & \text{otherwise.} \end{cases}$$

Discuss what happens to the solution as t increases. You do *not* need to write down an explicit formula, but for full credit you must explain (sketches can help) at least three or four interesting things that happen to the solution as time progresses.

4.1.4. Find a series solution to the initial-boundary value problem for the heat equation $u_t = u_{xx}$ for $0 < x < 1$ when one the end of the bar is held at 0° and the other is insulated. Discuss the asymptotic behavior of the solution as $t \to \infty$.

4.1.5. Answer Exercise 4.1.4 when both ends of the bar are insulated.

4.1.6. A metal bar, of length $\ell = 1$ meter and thermal diffusivity $\gamma = 2$, is taken out of a 100° oven and then fully insulated except for *one* end, which is fixed to a large ice cube at 0°. (*a*) Write down an initial-boundary value problem that describes the temperature $u(t,x)$ of the bar at all subsequent times. (*b*) Write a series formula for the temperature distribution $u(t,x)$ at time $t > 0$. (*c*) What is the equilibrium temperature distribution in the bar, i.e., for $t \gg 0$? How fast does the solution go to equilibrium? (*d*) Just before the temperature distribution reaches equilibrium, what does it look like? Sketch a picture and discuss.

4.1.7. A metal bar of length $\ell = 1$ and thermal diffusivity $\gamma = 1$ is fully insulated, including its ends. Suppose the initial temperature distribution is $u(0,x) = \begin{cases} x, & 0 \le x \le \frac{1}{2}, \\ 1 - x, & \frac{1}{2} \le x \le 1. \end{cases}$

(*a*) Use Fourier series to write down the temperature distribution at time $t > 0$.

(b) What is the equilibrium temperature distribution in the bar, i.e., for $t \gg 0$?
(c) How fast does the solution go to equilibrium? (d) Just before the temperature distribution reaches equilibrium, what does it look like? Sketch a picture and discuss.

4.1.8. (a) Find the series solution to the heat equation $u_t = u_{xx}$ on $-2 < x < 2$, $t > 0$, when subject to the Dirichlet boundary conditions $u(t, -2) = u(t, 2) = 0$ and the initial condition $u(0, x) = \begin{cases} x, & |\, x\,| < 1, \\ 0, & \text{otherwise.} \end{cases}$ (b) Sketch a graph of the solution at some representative times. (c) At what rate does the temperature approach thermal equilibrium?

4.1.9. Solve the heat equation when the right-hand end of a bar of unit length is held at a fixed constant temperature α while the left-hand end is insulated. Discuss the asymptotic behavior of the solution.

4.1.10. For each of the following initial temperature distributions, (i) write out the Fourier series solution to the heated ring (4.30–32), and (ii) find the resulting equilibrium temperature as $t \to \infty$: (a) $\cos x$, (b) $\sin^3 x$, (c) $|\, x\,|$, (d) $\begin{cases} 1, & -\pi < x < 0, \\ 0, & 0 < x < \pi. \end{cases}$

♢ 4.1.11. Suppose that the temperature $u(t, x)$ of a homogeneous bar satisfies the heat equation. Show that the associated heat flux $w(t, x)$ is also a solution to the same heat equation.

♢ 4.1.12. Show that the time derivative $v = u_t$ of any solution to the heat equation is also a solution. If $u(t, x)$ satisfies the initial condition $u(0, x) = f(x)$, what initial condition does $v(t, x)$ inherit?

♢ 4.1.13. Explain why the thermal energy $E(t) = \int_0^\ell u(t, x)\, dx$ is not constant for the Dirichlet initial-boundary value problem for the heat equation on the interval $[0, \ell\,]$.

♢ 4.1.14. (a) Show that the thermal energy $E(t) = \int_0^\ell u(t, x)\, dx$ is constant for the Neumann boundary value problem on the interval $[0, \ell\,]$. (b) Use part (a) to prove that the constant equilibrium solution for the homogeneous Neumann boundary value problem is equal to the mean initial temperature $u(0, x)$.

4.1.15. Let $u(t, x)$ be any nonconstant solution to the periodic heat equation (4.30–31). Prove that the squared L^2 norm of the solution, $N(t) = \int_{-\pi}^{\pi} u(t, x)^2\, dx$, is a strictly decreasing function of t. *Remark*: Interestingly, comparing this result with formula (4.38), we find that, for the periodic boundary value problem, the integral of u is constant, but the integral of u^2 is strictly decreasing. How is this possible?

♡ 4.1.16. The *cable equation* $v_t = \gamma\, v_{xx} - \alpha\, v$, with $\gamma, \alpha > 0$, also known as the *lossy heat equation*,was derived by the nineteenth-century Scottish physicist William Thomson to model propagation of signals in a transatlantic cable. Later, in honor of his work on thermodynamics, including determining the value of absolute zero temperature, he was named Lord Kelvin by Queen Victoria. The cable equation was later used to model the electrical activity of neurons. (a) Show that the general solution to the cable equation is given by $v(t, x) = e^{-\alpha\, t}\, u(t, x)$, where $u(t, x)$ solves the heat equation $u_t = \gamma\, u_{xx}$.
(b) Find a Fourier series solution to the Dirichlet initial-boundary value problem
$$v_t = \gamma\, v_{xx} - \alpha\, v, \qquad v(0, x) = f(x), \qquad v(t, 0) = 0 = v(t, 1), \qquad 0 \le x \le 1, \qquad t > 0.$$
Does your solution approach an equilibrium value? If so, how fast?
(c) Answer part (b) for the Neumann problem
$$v_t = \gamma\, v_{xx} - \alpha\, v, \qquad v(0, x) = f(x), \qquad v_x(t, 0) = 0 = v_x(t, 1), \qquad 0 \le x \le 1, \qquad t > 0.$$

♢ 4.1.17. The *convection-diffusion equation* $u_t + c\, u_x = \gamma\, u_{xx}$ is a simple model for the diffusion of a pollutant in a fluid flow moving with constant speed c. Show that $v(t, x) = u(t, x + c\, t)$ solves the heat equation. What is the physical interpretation of this change of variables?

4.1.18. Combine Exercises 4.1.16–17 to solve the *lossy convection-diffusion equation*
$$u_t = \gamma\, u_{xx} + c\, u_x - \alpha\, u.$$

♢ 4.1.19. Let $\gamma > 0$ and $\lambda \leq 0$. (*a*) Find all solutions to the differential equation $\gamma\, v'' + \lambda\, v = 0$. (*b*) Prove that the only solution that satisfies the boundary conditions $v(0) = 0$, $v(\ell) = 0$, is the zero solution $v(x) \equiv 0$.

♢ 4.1.20. Answer Exercise 4.1.19 when λ is a non-real complex number.

4.2 The Wave Equation

Let us return to the one-dimensional *wave equation*

$$\frac{\partial^2 u}{\partial t^2} = c^2\, \frac{\partial^2 u}{\partial x^2}\,, \tag{4.63}$$

with constant wave speed $c > 0$, used to model the vibrations of bars and strings. In Chapter 2, we learned how to explicitly solve the wave equation by the method of d'Alembert. Unfortunately, d'Alembert's approach does not extend to other equations of interest to us, and so alternative solution techniques, particularly those based on Fourier methods, are worth developing. Indeed, the resulting series solutions provide valuable insight into wave dynamics on bounded intervals.

Separation of Variables and Fourier Series Solutions

One of the oldest — and still one of the most widely used — techniques for constructing explicit analytic solutions to a wide range of linear partial differential equations is the method of *separation of variables*. We have, in fact, already employed a simplified version of the method when constructing each eigensolution to the heat equation as an exponential function of t times a function of x. In general, the separation of variables method seeks solutions to the partial differential equation that can be written as the product of functions of the individual independent variables. For the wave equation, we seek solutions

$$u(t, x) = w(t)\, v(x) \tag{4.64}$$

that can be written as the product of a function of t alone and a function of x alone. When the method succeeds (which is not guaranteed in advance), both factors are found as solutions to certain ordinary differential equations.

Let us see whether such an expression can possibly solve the wave equation. First of all,

$$\frac{\partial^2 u}{\partial t^2} = w''(t)\, v(x), \qquad\qquad \frac{\partial^2 u}{\partial x^2} = w(t)\, v''(x),$$

where the primes indicate ordinary derivatives. Substituting these expressions into the wave equation (4.63), we obtain

$$w''(t)\, v(x) = c^2\, w(t)\, v''(x).$$

Dividing both sides by $w(t)\, v(x)$ (which we assume is not identically zero, since otherwise, the solution would be trivial) yields

$$\frac{w''(t)}{w(t)} = c^2\, \frac{v''(x)}{v(x)}\,,$$

which effectively "separates" the t and x variables on each side of the equation, whence the name "separation of variables".

Now, how could a function of t alone be equal to a function of x alone? A moment's reflection should convince the reader that this can happen if and only if the two functions are constant,† so

$$\frac{w''(t)}{w(t)} = c^2 \, \frac{v''(x)}{v(x)} = \lambda, \tag{4.65}$$

where we use λ to indicate the common *separation constant*. Thus, the individual factors $w(t)$ and $v(x)$ must satisfy ordinary differential equations

$$\frac{d^2w}{dt^2} - \lambda\, w = 0, \qquad\qquad \frac{d^2v}{dx^2} - \frac{\lambda}{c^2}\, v = 0,$$

as promised. We already know how to solve both of these ordinary differential equations by elementary techniques. There are three different cases, depending on the sign of the separation constant λ. As a result, each value of λ leads to four independent separable solutions to the wave equation, as listed in the accompanying table.

Separable Solutions to the Wave Equation

λ	$w(t)$	$v(x)$	$u(t,x) = w(t)\, v(x)$
$\lambda = -\omega^2 < 0$	$\cos \omega t,\ \sin \omega t$	$\cos \dfrac{\omega x}{c}\,,\ \sin \dfrac{\omega x}{c}$	$\cos \omega t\ \cos \dfrac{\omega x}{c}\,,\ \cos \omega t\ \sin \dfrac{\omega x}{c}\,,$ $\sin \omega t\ \cos \dfrac{\omega x}{c}\,,\ \sin \omega t\ \sin \dfrac{\omega x}{c}$
$\lambda = 0$	$1,\ t$	$1,\ x$	$1,\ x,\ t,\ tx$
$\lambda = \omega^2 > 0$	$e^{-\omega t},\ e^{\omega t}$	$e^{-\omega x/c},\ e^{\omega x/c}$	$e^{-\omega\,(t+x/c)},\ e^{\omega\,(t-x/c)},$ $e^{-\omega\,(t-x/c)},\ e^{\omega\,(t+x/c)}$

So far, we have not taken the boundary conditions into account. Consider first the case of a string of length ℓ with two fixed ends, and thus subject to homogeneous Dirichlet boundary conditions

$$u(t,0) = 0 = u(t,\ell).$$

Substituting the separable ansatz (4.65), we find that $v(x)$ must satisfy

$$\frac{d^2v}{dx^2} - \frac{\lambda}{c^2}\, v = 0, \qquad\qquad v(0) = 0 = v(\ell). \tag{4.66}$$

The complete system of (nontrivial) solutions to this boundary value problem were found in (4.21):

$$v_n(x) = \sin \frac{n\,\pi\, x}{\ell}\,, \qquad\qquad \lambda_n = -\left(\frac{n\,\pi\, c}{\ell}\right)^2, \qquad\qquad n = 1, 2, 3, \ldots\,.$$

† *Technical detail*: one should assume that the underlying domain is connected for this to be valid as stated. In practice, this technicality can be safely ignored.

Hence, according to the table, the corresponding separable solutions are

$$u_n(t,x) = \cos\frac{n\pi c t}{\ell}\,\sin\frac{n\pi x}{\ell}\,, \qquad \widetilde{u}_n(t,x) = \sin\frac{n\pi c t}{\ell}\,\sin\frac{n\pi x}{\ell}\,. \tag{4.67}$$

We will now employ these solutions to construct a candidate series solution to the wave equation subject to the prescribed boundary conditions:

$$u(t,x) = \sum_{n=1}^{\infty}\left[\,b_n\cos\frac{n\pi c t}{\ell}\,\sin\frac{n\pi x}{\ell} + d_n\sin\frac{n\pi c t}{\ell}\,\sin\frac{n\pi x}{\ell}\,\right]. \tag{4.68}$$

The solution is thus a linear combination of the natural Fourier modes vibrating with frequencies

$$\omega_n = \frac{n\pi c}{\ell} = \frac{n\pi}{\ell}\sqrt{\frac{\kappa}{\rho}}\,, \qquad n = 1,2,3,\dots\,, \tag{4.69}$$

where the second expression follows from (2.66). Observe that, the longer the length ℓ of the string, or the higher its density ρ, the slower the vibrations, whereas increasing its stiffness or tension κ speeds them up — in exact accordance with our physical intuition.

The Fourier coefficients b_n and d_n in (4.68) will be uniquely determined by the initial conditions

$$u(0,x) = f(x), \qquad \frac{\partial u}{\partial t}\,(0,x) = g(x), \qquad 0 < x < \ell.$$

Differentiating the series term by term, we discover that we must represent the initial displacement and velocity as Fourier sine series

$$u(0,x) = \sum_{n=1}^{\infty} b_n \sin\frac{n\pi x}{\ell} = f(x), \qquad \frac{\partial u}{\partial t}\,(0,x) = \sum_{n=1}^{\infty} d_n\frac{n\pi c}{\ell}\,\sin\frac{n\pi x}{\ell} = g(x).$$

Therefore,

$$b_n = \frac{2}{\ell}\int_0^{\ell} f(x)\,\sin\frac{n\pi x}{\ell}\,dx, \qquad n = 1,2,3,\dots\,, \tag{4.70}$$

are the Fourier sine coefficients (3.85) of the initial displacement $f(x)$, while

$$d_n = \frac{2}{n\pi c}\int_0^{\ell} g(x)\,\sin\frac{n\pi x}{\ell}\,dx, \qquad n = 1,2,3,\dots\,. \tag{4.71}$$

are rescaled versions of the Fourier sine coefficients of the initial velocity $g(x)$.

Example 4.3. A string of unit length fixed at both ends is held taut at its center and then released. Our task is to describe the ensuing vibrations. Let us assume that the physical units are chosen so that $c^2 = 1$, and so we are asked to solve the initial-boundary value problem

$$u_{tt} = u_{xx}, \qquad u(0,x) = f(x), \qquad u_t(0,x) = 0, \qquad u(t,0) = u(t,1) = 0. \tag{4.72}$$

To be specific, we assume that the center of the string has been moved by half a unit, and so the initial displacement is

$$f(x) = \begin{cases} x, & 0 \le x \le \frac{1}{2}, \\ 1 - x, & \frac{1}{2} \le x \le 1. \end{cases}$$

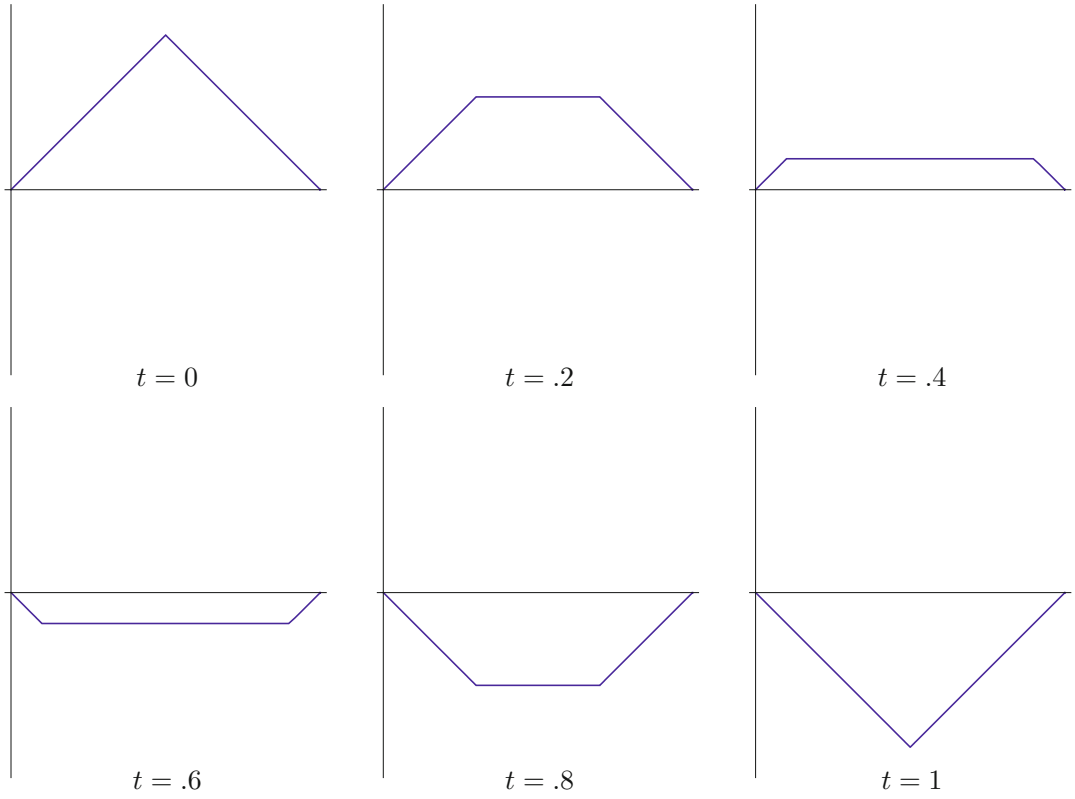


Figure 4.4. Plucked string solution of the wave equation.

The vibrational frequencies $\omega_n = n\pi$ are the integral multiples of π, and so the natural modes of vibration are

$$\cos n\pi t \, \sin n\pi x \qquad \text{and} \qquad \sin n\pi t \, \sin n\pi x \qquad \text{for} \qquad n = 1, 2, \dots .$$

Consequently, the general solution to the boundary value problem is

$$u(t,x) = \sum_{n=1}^{\infty} \big(b_n \cos n\pi t \, \sin n\pi x + d_n \sin n\pi t \, \sin n\pi x \big),$$

where

$$b_n = 2\int_0^1 f(x) \sin n\pi x \, dx = \begin{cases} 4\displaystyle\int_0^{1/2} x \sin n\pi x \, dx = \frac{4\,(-1)^k}{(2k+1)^2\pi^2}, & n = 2k+1, \\ 0, & n = 2k, \end{cases}$$

while $d_n = 0$. Therefore, the solution is the Fourier sine series

$$u(t,x) = \frac{4}{\pi^2} \sum_{k=0}^{\infty} (-1)^k \frac{\cos(2k+1)\pi t \, \sin(2k+1)\pi x}{(2k+1)^2}, \tag{4.73}$$

whose profile is depicted in Figure 4.4. At time $t = 1$, the original displacement is reproduced exactly, but upside down. The subsequent dynamics proceeds as before, but in mirror-image form. The original displacement reappears at time $t = 2$, after which time

the motion is periodically repeated. Interestingly, at times $t_k = .5, 1.5, 2.5, \dots$, the displacement is identically zero, $u(t_k, x) \equiv 0$, although the velocity is not, $u_t(t_k, x) \not\equiv 0$. The solution appears to be piecewise affine, i.e., its graph is a collection of straight lines. This can, in fact, be proved as a consequence of the d'Alembert formula; see Exercise 4.2.13. Observe that, unlike the heat equation, the wave equation does *not* smooth out discontinuities and corners in the initial data. And, although we will loosely refer to such piecewise C^2 functions as "solutions", they are not, in fact, classical solutions. (Their status as weak solutions, though, can be established using the methods of Section 10.4.)

While the series form (4.68) of the solution is perhaps less satisfying than a d'Alembert-style formula, we can still use it to deduce important qualitative properties. First of all, since each term is periodic in t with period $2\,\ell/c$, the entire solution is time periodic with that period: $u(t + 2\,\ell/c, x) = u(t, x)$. In fact, after half a period, the solution reduces to

$$u\left(\frac{\ell}{c}\,, x\right) = \sum_{n=1}^{\infty} (-1)^n\, b_n\, \sin\frac{n\,\pi x}{\ell} = -\sum_{n=1}^{\infty} b_n\, \sin\frac{n\,\pi(\ell - x)}{\ell} = -\,u(0, \ell - x) = -\,f(\ell - x).$$

In general,

$$u\left(t + \frac{\ell}{c}\,,\ x\right) = -\,u(t, \ell - x), \qquad\qquad u\left(t + \frac{2\,\ell}{c}\,,\ x\right) = u(t, x). \tag{4.74}$$

Therefore, the initial wave form is reproduced, first as an upside down mirror image of itself at time $t = \ell/c$, and then in its original form at time $t = 2\,\ell/c$. This has the important consequence that vibrations of (homogeneous) one-dimensional media are inherently periodic, because the fundamental frequencies (4.69) are all integer multiples of the lowest one: $\omega_n = n\omega_1$.

Remark: The immediately preceding remark has important musical consequences. To the human ear, sonic vibrations that are integral multiples of a single frequency, and thus periodic in time, sound harmonious, whereas those with irrationally related frequencies, and hence experiencing aperiodic vibrations, sound dissonant. This is why most tonal instruments rely on vibrations in one dimension, be it a violin or piano string, a column of air in a wind instrument (flute, clarinet, trumpet, or saxophone), a xylophone bar, or a triangle. On the other hand, most percussion instruments rely on the vibrations of two-dimensional media, e.g., drums and cymbals, or three-dimensional solid bodies, e.g., blocks. As we shall see in Chapters 11 and 12, the frequency ratios of the latter are irrationally related, and hence their motion is only quasiperiodic, as in Example 2.20. For some reason, our appreciation of music is psychologically attuned to the differences between rationally related/periodic and irrationally related/quasiperiodic vibrations, [**105**].

Consider next a string with both ends left free, and so subject to the Neumann boundary conditions

$$\frac{\partial u}{\partial x}\,(t, 0) = 0 = \frac{\partial u}{\partial x}\,(t, \ell). \tag{4.75}$$

The solutions of (4.66) satisfying $v'(0) = 0 = v'(\ell)$ are now

$$v_n(x) = \cos\frac{n\,\pi x}{\ell} \qquad \text{with} \qquad \omega_n = \frac{n\,\pi c}{\ell}, \qquad n = 0, 1, 2, 3, \dots\,.$$

The resulting solution takes the form of a Fourier cosine series

$$u(t, x) = a_0 + c_0\, t + \sum_{n=1}^{\infty} \left(a_n \cos\frac{n\,\pi c\, t}{\ell}\, \cos\frac{n\,\pi x}{\ell} + c_n \sin\frac{n\,\pi c\, t}{\ell}\, \cos\frac{n\,\pi x}{\ell} \right). \tag{4.76}$$

The first two terms come from the null eigenfunction $v_0(x) = 1$ with $\omega_0 = 0$. The string vibrates with the same fundamental frequencies (4.69) as in the fixed-end case, but there is now an additional *unstable mode* $c_0\, t$ that is no longer periodic, but grows linearly in time. In general, the presence of null eigenfunctions implies that the wave equation admits unstable modes.

Substituting (4.76) into the initial conditions

$$u(0,x) = f(x), \qquad \frac{\partial u}{\partial t}\,(0,x) = g(x), \qquad 0 < x < \ell,$$

we find that the Fourier coefficients are prescribed, as before, by the initial displacement and velocity:

$$a_n = \frac{2}{\ell}\int_0^\ell f(x)\,\cos\frac{n\,\pi\,x}{\ell}\,dx, \qquad c_n = \frac{2}{n\,\pi\,c}\int_0^\ell g(x)\,\cos\frac{n\,\pi\,x}{\ell}\,dx, \qquad n = 1, 2, 3, \ldots\,.$$

The order-zero coefficients[†]

$$a_0 = \frac{1}{\ell}\int_0^\ell f(x)\,dx, \qquad\qquad c_0 = \frac{1}{\ell}\int_0^\ell g(x)\,dx,$$

are equal to the average initial displacement and average initial velocity of the string. In particular, when $c_0 = 0$, there is no net initial velocity, and the unstable mode is not excited. In this case, the solution is time-periodic, oscillating around the position given by the average initial displacement. On the other hand, if $c_0 \neq 0$, the string will move off with constant average speed c_0, all the while vibrating at the same fundamental frequencies.

Similar considerations apply to the periodic boundary value problem for the wave equation on a circular ring. The details are left as Exercise 4.2.6 for the reader.

Exercises

4.2.1. In music, an octave corresponds to doubling the frequency of the sound waves. On my piano, the middle C string has length .7 meter, while the string for the C an octave higher has length .6 meter. Assuming that they have the same density, how much tighter does the shorter string need to be tuned?

4.2.2. How much longer would a piano string have to be to make the same sound when it is pulled twice as tight?

4.2.3. Write down the solutions to the following initial-boundary value problems for the wave equation in the form of a Fourier series:
(a) $u_{tt} = u_{xx}$, $u(t,0) = u(t,\pi) = 0$, $u(0,x) = 1$, $u_t(0,x) = 0$;
(b) $u_{tt} = 2\,u_{xx}$, $u(t,0) = u(t,\pi) = 0$, $u(0,x) = 0$, $u_t(0,x) = 1$;
(c) $u_{tt} = 3\,u_{xx}$, $u(t,0) = u(t,\pi) = 0$, $u(0,x) = \sin^3 x$, $u_t(0,x) = 0$;
(d) $u_{tt} = 4\,u_{xx}$, $u(t,0) = u(t,1) = 0$, $u(0,x) = x$, $u_t(0,x) = -x$;
(e) $u_{tt} = u_{xx}$, $u(t,0) = u_x(t,1) = 0$, $u(0,x) = 1$, $u_t(0,x) = 0$;
(f) $u_{tt} = 2\,u_{xx}$, $u_x(t,0) = u_x(t,2\pi) = 0$, $u(0,x) = -1$, $u_t(0,x) = 1$;
(g) $u_{tt} = u_{xx}$, $u_x(t,0) = u_x(t,1) = 0$, $u(0,x) = x(1-x)$, $u_t(0,x) = 0$.

[†] Note that we have not included the usual $\frac{1}{2}$ factor in the constant terms in the Fourier series (4.76).

4.2.4. Find all separable solutions to the wave equation $u_{tt} = u_{xx}$ on the interval $0 \le x \le \pi$ subject to (*a*) mixed boundary conditions $u(t,0) = 0$, $u_x(t,\pi) = 0$;
(*b*) Neumann boundary conditions $u_x(t,0) = 0$, $u_x(t,\pi) = 0$.

4.2.5. (*a*) Under what conditions is the solution to the Neumann boundary value problem (4.75) a periodic function of t? What is the period? (*b*) Establish explicit periodicity formulas of the form (4.74). (*c*) Under what conditions is the velocity $\partial u/\partial t$ periodic in t?

♡ 4.2.6. (*a*) Formulate the periodic initial-boundary value problem for the wave equation on the interval $-\pi \le x \le \pi$, modeling the vibrations of a circular ring. (*b*) Write out a formula for the solution to your problem in the form of a Fourier series. (*c*) Is the solution a periodic function of t? If so, what is the period? (*d*) Suppose the initial displacement coincides with that in Figure 4.6, while the initial velocity is zero. Describe what happens to the solution as time evolves.

4.2.7. Show that the time derivative, $v = \partial u/\partial t$, of any solution to the wave equation is also a solution. If you know the initial conditions of u, what initial conditions does v satisfy?

4.2.8. Find all the separable real solutions to the wave equation subject to a restoring force: $u_{tt} = u_{xx} - u$. Discuss their long-term behavior.

♡ 4.2.9. Let $a, c > 0$ be positive constants. The *telegrapher's equation* $u_{tt} + a\,u_t = c^2\,u_{xx}$ represents a damped version of the wave equation. Consider the Dirichlet boundary value problem $u(t,0) = u(t,1) = 0$, on the interval $0 \le x \le 1$, with initial conditions $u(0,x) = f(x)$, $u_t(0,x) = 0$. (*a*) Find all separable solutions to the telegrapher's equation that satisfy the boundary conditions. (*b*) Write down a series solution for the initial boundary value problem. (*c*) Discuss the long term behavior of your solution. (*d*) State a criterion that distinguishes overdamped from underdamped versions of the equation.

4.2.10. The fourth-order partial differential equation $u_{tt} = -u_{xxxx}$ is a simple model for a vibrating elastic beam. (*a*) Find all separable real solutions to the beam equation. (*b*) Show that any (complex) solution to the Schrödinger equation $\mathrm{i}\,u_t = u_{xx}$ solves the beam equation.

4.2.11. The initial-boundary value problem

$$u_{tt} = -u_{xxxx}, \qquad \begin{array}{c} u(t,0) = u_{xx}(t,0) = u(t,1) = u_{xx}(t,1) = 0, \qquad 0 < x < 1, \\ u(0,x) = f(x), \qquad u_t(0,x) = 0, \qquad t > 0, \end{array}$$

models the vibrations of an elastic beam of unit length with simply supported ends, subject to a nonzero initial displacement $f(x)$ and zero initial velocity. (*a*) What are the vibrational frequencies for the beam? (*b*) Write down the solution to the initial-boundary value problem as a Fourier series. (*c*) Does the beam vibrate periodically
(*i*) for all initial conditions? (*ii*) for some initial conditions? (*iii*) for no initial conditions?

4.2.12. *Multiple choice*: The initial-boundary value problem

$$u_{tt} = u_{xxxx}, \qquad \begin{array}{c} u(t,0) = u_{xx}(t,0) = u(t,1) = u_{xx}(t,1) = 0, \qquad 0 < x < 1, \\ u(0,x) = f(x), \qquad u_t(0,x) = g(x), \qquad t > 0, \end{array}$$

is well-posed for (*a*) $t > 0$; (*b*) $t < 0$; (*c*) all t; (*d*) no t. Explain your answer.

The d'Alembert Formula for Bounded Intervals

In Theorem 2.15, we derived the explicit d'Alembert formula

$$u(t,x) = \frac{f(x - c\,t) + f(x + c\,t)}{2} + \frac{1}{2\,c}\int_{x-c\,t}^{x+c\,t} g(z)\,dz, \tag{4.77}$$

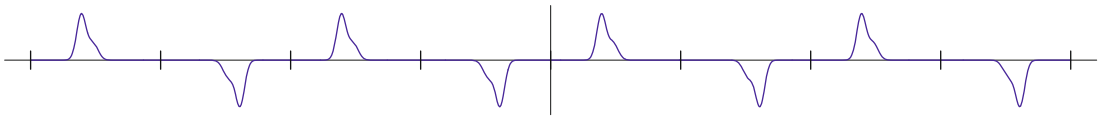


Figure 4.5. Odd periodic extension of a concentrated pulse.

for solving the basic initial value problem for the wave equation on an infinite interval:

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}, \qquad u(0,x) = f(x), \qquad \frac{\partial u}{\partial t}(0,x) = g(x), \qquad -\infty < x < \infty.$$

In this section we explain how to adapt the formula in order to solve initial-boundary value problems on bounded intervals, thereby effectively summing the Fourier series solution.

The easiest case to deal with is the periodic problem on $0 \le x \le \ell$, with boundary conditions

$$u(t,0) = u(t,\ell), \qquad u_x(t,0) = u_x(t,\ell). \tag{4.78}$$

If we extend the initial displacement $f(x)$ and velocity $g(x)$ to be periodic functions of period ℓ, so $f(x+\ell) = f(x)$ and $g(x+\ell) = g(x)$ for all $x \in \mathbb{R}$, then the resulting d'Alembert solution (4.77) will also be periodic in x, so $u(t, x+\ell) = u(t,x)$. In particular, it satisfies the boundary conditions (4.78) and so coincides with the desired solution. Details are to be supplied in Exercises 4.2.27–28.

Next, suppose we have fixed (Dirichlet) boundary conditions

$$u(t,0) = 0, \qquad u(t,\ell) = 0. \tag{4.79}$$

The resulting solution can be written as a Fourier sine series (4.68), and hence is both odd and 2ℓ–periodic in x. Therefore, to write the solution in d'Alembert form (4.77), we extend the initial displacement $f(x)$ and velocity $g(x)$ to be odd, periodic functions of period 2ℓ:

$$f(-x) = -f(x), \qquad f(x+2\ell) = f(x), \qquad g(-x) = -g(x), \qquad g(x+2\ell) = g(x).$$

This will ensure that the d'Alembert solution also remains odd and periodic. As a result, it satisfies the homogeneous Dirichlet boundary conditions (4.79) for all t, cf. Exercise 4.2.31. Keep in mind that, while the solution $u(t,x)$ is defined for all x, the only physically relevant values occur on the interval $0 \le x \le \ell$. Nevertheless, the effects of displacements in the unphysical regime will eventually be felt as the propagating waves pass through the physical interval.

For example, consider an initial displacement that is concentrated near $x = \xi$ for some $0 < \xi < \ell$. Its odd 2ℓ–periodic extension consists of two sets of replicas: those of the same form occurring at positions $\xi \pm 2\ell$, $\xi \pm 4\ell$, $\dots$, and their upside-down mirror images at the intermediate positions $-\xi$, $-\xi \pm 2\ell$, $-\xi \pm 4\ell$, $\dots$; Figure 4.5 shows a representative example. The resulting solution begins with each of the pulses, both positive and negative, splitting into two half-size replicas that propagate with speed c in opposite directions. When a left and right moving pulse meet, they emerge from the interaction unaltered. The process repeats periodically, with an infinite row of half-size pulses moving to the right kaleidoscopically interacting with an infinite row moving to the left.

However, only the part of this solution that lies on $0 \le x \le \ell$ is actually observed on the physical string. The effect is as if one were watching the full solution as it passes by a window of length ℓ. Such observers will interpret what they see a bit differently. To

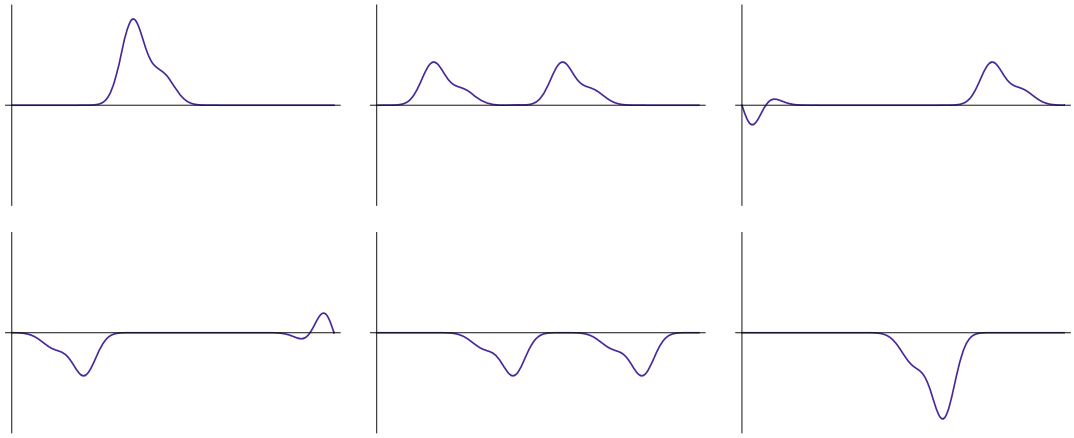


Figure 4.6. Solution to wave equation with fixed ends.

wit, the original pulse starting at position $0 < \xi < \ell$ splits up into two half-size replicas that move off in opposite directions. As each half-size pulse reaches an end of the string, it meets a mirror-image pulse that has been propagating in the opposite direction from the nonphysical regime. The pulse is reflected at the end of the interval and becomes an upside-down mirror image moving in the opposite direction. The original positive pulse has moved off the end of the string just as its mirror image has moved into the physical regime. (A common physical realization is a pulse propagating down a jump rope that is held fixed at its end; the reflected pulse returns upside down.) A similar reflection occurs as the other half-size pulse hits the other end of the physical interval, after which the solution consists of two upside-down half-size pulses moving back towards each other. At time $t = \ell/c$ they recombine at the point $\ell - \xi$ to instantaneously form a full-sized, but upside-down mirror image of the original disturbance — in accordance with (4.74). The recombined pulse in turn splits apart into two upside-down half-size pulses that, when each collides with the end, reflect and return to their original upright form. At time $t = 2\ell/c$, the pulses recombine to exactly reproduce the original displacement. The process then repeats, and the solution is periodic in time with period $2\ell/c$.

In Figure 4.6, the first picture displays the initial displacement. In the second, it has split into left- and right-moving half-size clones. In the third picture, the left-moving bump is in the process of colliding with the left end of the string. In the fourth picture, it has emerged from the collision, and is now upside down, reflected, and moving to the right. Meanwhile, the right-moving pulse is starting to collide with the right end. In the fifth picture, both pulses have completed their collisions and are now moving back towards each other, where, in the last picture, they recombine into an upside-down mirror image of the original pulse. The process then repeats itself, in mirror image, finally recombining to the original pulse, at which point the entire process starts over.

The Neumann (free) boundary value problem

$$\frac{\partial u}{\partial x}(t,0) = 0, \qquad \frac{\partial u}{\partial x}(t,\ell) = 0, \tag{4.80}$$

is handled similarly. Since the solution has the form of a Fourier cosine series in x, we

extend the initial conditions to be *even* 2ℓ–periodic functions

$$f(-x) = f(x), \qquad f(x + 2\ell) = f(x), \qquad g(-x) = g(x), \qquad g(x + 2\ell) = g(x).$$

The resulting d'Alembert solution (4.77) is also even and 2ℓ–periodic in x, and hence satisfies the boundary conditions, cf. Exercise 4.2.31(*b*). In this case, when a pulse hits one of the ends, its reflection remains upright, but becomes a mirror image of the original; a familiar physical illustration is a water wave that reflects off a solid wall. Further details are left to the reader in Exercise 4.2.22

In summary, we have now studied two very different ways to solve the one-dimensional wave equation. The first, based on the d'Alembert formula, emphasizes their particle-like aspects, where individual wave packets collide with each other, or reflect at the boundary, all the while maintaining their overall form, while the second, based on Fourier analysis, emphasizes the vibrational or wave-like character of the solutions. Some solutions look like vibrating waves, while others appear much more like interacting particles. But, like the proverbial blind men describing an elephant, these are merely two facets of the *same* solution. The Fourier series formula shows how every particle-like solution can be decomposed into its constituent vibrational modes, while the d'Alembert formula demonstrates how vibrating solutions combine into moving wave packets.

The coexistence of particle and wave features is reminiscent of the long-running historical debate over the nature of light. Newton and his disciples proposed a particle-based theory, anticipating the modern concept of photons. However, until the beginning of the twentieth century, most physicists advocated a wave-like or vibrational viewpoint. Einstein's explanation of the photoelectric effect served to resurrect the particle interpretation. Only with the establishment of quantum mechanics was the debate resolved — light, and, indeed, all subatomic particles manifest *both* particle and wave features, depending upon the experiment and the physical situation. But a theoretical basis for the perplexing wave-particle duality could have been found already in Fourier's and d'Alembert's competing solution formulae for the classical wave equation!

Exercises

♢ 4.2.13. (*a*) Solve the initial-boundary value problem from Example 4.3 using the d'Alembert method.
(*b*) Verify that your solution coincides with the Fourier series solution derived above.
(*c*) Justify our earlier observation that, at each time t, the solution $u(t, x)$ is a piecewise affine function of x.

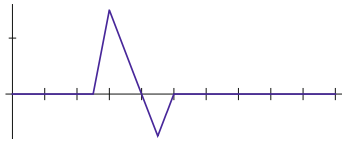
4.2.14. Sketch the solution of the wave equation $u_{tt} = u_{xx}$ and describe its behavior when the initial displacement is the box function $u(0, x) = \begin{cases} 1, & 1 < x < 2, \\ 0, & \text{otherwise}, \end{cases}$ while the initial velocity is 0 in each of the following scenarios: (*a*) on the entire line $-\infty < x < \infty$; (*b*) on the half-line $0 \le x < \infty$, with homogeneous Dirichlet boundary condition at the end; (*c*) on the half-line $0 \le x < \infty$, with homogeneous Neumann boundary condition at the end; (*d*) on the bounded interval $0 \le x \le 5$ with homogeneous Dirichlet boundary conditions; (*e*) on the bounded interval $0 \le x \le 5$ with homogeneous Neumann boundary conditions.

4.2.15. Answer Exercise 4.2.14 when the initial velocity is the box function, while the initial displacement is zero.

4.2.16. Consider the initial-boundary value problem

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}, \qquad \begin{array}{ll} u(t,0) = 0 = u(t,10), & t > 0, \\ u(0,x) = f(x), \quad u_t(0,x) = 0, & 0 < x < 10, \end{array}$$

for the wave equation, where the initial data has the following form:



$$f(x) = \begin{cases} 3x - 7.5, & 2.5 \le x \le 3, \\ 6 - 1.5x, & 3 \le x \le 4.5, \\ 1.5x - 7.5, & 4.5 \le x \le 5, \\ 0, & \text{otherwise.} \end{cases}$$

Discuss what happens to the solution. You do *not* need to write down an explicit formula for the solution, but for full credit you must explain (sketches can help) at least three or four interesting things that happen to the solution as time progresses.

4.2.17. Repeat Exercise 4.2.16 for the Neumann boundary conditions.

4.2.18. Suppose the initial displacement of a string of length ℓ looks like the graph to the right. Assuming that the ends of the string are held fixed, graph the string's profile at times $t = \ell/c$ and $2\ell/c$.

♣ 4.2.19. Consider the wave equation $u_{tt} = u_{xx}$ on the interval $0 \le x \le 1$, with homogeneous Dirichlet boundary conditions at both ends. (*a*) Use the d'Alembert formula to explicitly solve the initial value problem $u(0,x) = x - x^2$, $u_t(0,x) = 0$. (*b*) Graph the solution profile at some representative times, and discuss what you observe. (*c*) Find the Fourier series at each t of your solution and compare the two. (*d*) How many terms do you need to sum to obtain a reasonable approximation to the exact solution?

♣ 4.2.20. Solve Exercise 4.2.19 for the initial conditions $u(0,x) = 0$, $u_t(0,x) = x^2 - x$.

♣ 4.2.21. Solve (*i*) Exercise 4.2.19, (*ii*) Exercise 4.2.20, when the solution is subject to homogeneous Neumann boundary conditions.

♢ 4.2.22. Under what conditions is the solution to the Neumann boundary value problem for the wave equation on a bounded interval $[0, \ell]$ periodic in time? What is the period?

4.2.23. Discuss and sketch the behavior of the solution to the Neumann boundary value problem $u_{tt} = 4u_{xx}$, $0 < x < 1$, $u_x(t,0) = 0 = u_x(t,1)$, $u(0,x) = f(x)$, $u_t(0,x) = g(x)$, for

(*a*) a localized initial displacement: $f(x) = \begin{cases} 1, & .2 < x < .3, \\ 0, & \text{otherwise.} \end{cases}$ $g(x) = 0$;

(*b*) a localized initial velocity: $f(x) = 0$, $g(x) = \begin{cases} 1, & .2 < x < .3, \\ 0, & \text{otherwise.} \end{cases}$.

4.2.24. (*a*) Explain how to solve the Neumann initial-boundary value problem

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}, \qquad \frac{\partial u}{\partial x}(t,0) = 0 = \frac{\partial u}{\partial x}(t,1), \qquad u(0,x) = f(x), \qquad \frac{\partial u}{\partial t}(0,x) = g(x),$$

on the interval $0 \le x \le 1$.

(*b*) Let $f(x) = \begin{cases} x - \frac{1}{4}, & \frac{1}{4} \le x \le \frac{1}{2}, \\ \frac{3}{4} - x, & \frac{1}{2} \le x \le \frac{3}{4}, \\ 0, & \text{otherwise,} \end{cases}$ and $g(x) = 0$. Sketch the graph of the solution at a few representative times, and discuss what is happening. Is the solution periodic in time? If so, what is the period?

(*c*) Do the same when $f(x) = 0$ and $g(x) = x$.

4.2.25. (*a*) Write down a formula for the solution $u(t,x)$ to the initial-boundary value problem

$$\frac{\partial^2 u}{\partial t^2} - 4\,\frac{\partial^2 u}{\partial x^2} = 0, \quad u(0,x) = \sin x, \quad \frac{\partial u}{\partial t}\,(0,x) = \frac{\partial u}{\partial x}\,(t,0) = \frac{\partial u}{\partial x}\,(t,\pi) = 0, \quad 0 < x < \pi, \quad t > 0.$$

(*b*) Find $u\left(\frac{\pi}{2}, \frac{\pi}{2}\right)$. (*c*) Prove that $h(t) = u\left(t, \frac{\pi}{2}\right)$ is a periodic function of t and find its period. (*d*) Does $\dfrac{\partial u}{\partial x}$ have any discontinuities? If so, discuss their behavior.

4.2.26. Answer Exercise 4.2.25 for the mixed boundary conditions $u(t,0) = 0 = u_x(t,\pi)$.

♡ 4.2.27. (*a*) Explain how to use d'Alembert's formula (4.77) to solve the periodic initial-boundary value problem for the wave equation given in Exercise 4.2.6.
(*b*) Do the d'Alembert and Fourier series formulae represent the same solution? If so, can you justify it? If not, explain why they are different.

♢ 4.2.28. Show that the solution $u(t,x)$ to the wave equation on an interval $[0,\ell]$, subject to periodic boundary conditions $u(t,0) = u(t,\ell)$, $u_x(t,0) = u_x(t,\ell)$, is a periodic function of t if and only if there is no net initial velocity: $\displaystyle\int_0^\ell g(x)\,dx = 0$.

4.2.29. (*a*) Explain how to solve the wave equation on a half-line $x > 0$ when subject to Dirichlet boundary conditions $u(t,0) = 0$. (*b*) Assuming $c = 1$, find the solution satisfying $u(0,x) = (x-2)\,e^{-5\,(x-2.2)^2}$, $u_t(0,x) = 0$. (*c*) Sketch a picture of your solution at some representative times, and discuss what is happening.

4.2.30. Solve Exercise 4.2.29 for homogeneous Neumann boundary conditions at $x = 0$.

♢ 4.2.31. (*a*) Given that $f(x)$ is odd and 2ℓ–periodic, explain why $f(0) = 0 = f(\ell)$.
(*b*) Given that $f(x)$ is even and 2ℓ–periodic, explain why $f'(0) = 0 = f'(\ell)$.

♢ 4.2.32. (*a*) Prove that if $f(-x) = -f(x)$, $f(x + 2\ell) = f(x)$, for all x, then $u(t,x) = \frac{1}{2}\,[\,f(x - c\,t) + f(x + c\,t)\,]$ satisfies the Dirichlet boundary conditions (4.79).
(*b*) Prove that if $g(-x) = -g(x)$, $g(x + 2\ell) = g(x)$ for all x, then $u(t,x) = \dfrac{1}{2\,c}\displaystyle\int_{x-c\,t}^{x+c\,t} g(z)\,dz$ also satisfies the Dirichlet boundary conditions.

4.2.33. If both $u(0,x) = f(x)$ and $u_t(0,x) = g(x)$ are even functions, show that the solution $u(t,x)$ of the wave equation is even in x for all t.

4.2.34. (*a*) Prove that the solution $u(t,x)$ to the wave equation for $x \in \mathbb{R}$ is an even function of t if and only if its initial velocity, at $t = 0$, is zero.
(*b*) Under what conditions is $u(t,x)$ an odd function of t?

♢ 4.2.35. Let $u(t,x)$ be a classical solution to the wave equation $u_{tt} = c^2 u_{xx}$ on the interval $0 < x < \ell$, satisfying homogeneous Dirichlet boundary conditions. The *total energy* of u at time t is

$$E(t) = \int_0^\ell \frac{1}{2}\left[\left(\frac{\partial u}{\partial t}\right)^2 + c^2\left(\frac{\partial u}{\partial x}\right)^2\right] dx. \tag{4.81}$$

Establish the *Law of Conservation of Energy* by showing that $E(t) = E(0)$ is a constant function.

♢ 4.2.36. (*a*) Use Exercise 4.2.35 to prove that the only C^2 solution to the initial-boundary value problem $v_{tt} = c^2 v_{xx}$, $v(t,0) = v(t,\ell) = 0$, $v(0,x) = 0$, $v_t(0,x) = 0$, is the trivial solution $v(t,x) \equiv 0$. (*b*) Establish the following *Uniqueness Theorem* for the wave equation: given $f(x), g(x) \in \mathrm{C}^2$, there is at most one C^2 solution $u(t,x)$ to the initial-boundary value problem $u_{tt} = c^2 u_{xx}$, $u(t,0) = u(t,\ell) = 0$, $u(0,x) = f(x)$, $u_t(0,x) = g(x)$.

4.2.37. Referring back to Exercises 4.2.35 and 4.2.36: (*a*) Does conservation of energy hold for solutions to the homogeneous Neumann initial-boundary value problem?
(*b*) Can you establish a uniqueness theorem for the Neumann problem?

4.2.38. Explain how to solve the Dirichlet initial-boundary value problem

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

for the wave equation subject to an external forcing on the interval $[0, \ell]$.

4.3 The Planar Laplace and Poisson Equations

The two-dimensional *Laplace equation* is the second-order linear partial differential equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \tag{4.82}$$

named in honor of the influential eighteenth-century French mathematician Pierre–Simon Laplace. It, along with its higher-dimensional versions, is arguably the most important differential equation in all of mathematics. A real-valued solution $u(x, y)$ to the Laplace equation is known as a *harmonic function*. The space of harmonic functions can thus be identified as the kernel of the second-order linear partial differential operator

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\,, \tag{4.83}$$

known as the *Laplace operator*, or *Laplacian* for short. The inhomogeneous or forced version, namely

$$-\Delta[u] = -\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} = f(x,y), \tag{4.84}$$

is known as *Poisson's equation*, named after Siméon–Denis Poisson, who was taught by Laplace. The mathematical and physical reasons for including the minus sign will gradually become clear.

Besides their theoretical importance, the Laplace and Poisson equations arise as the basic equilibrium equations in a remarkable variety of physical systems. For example, we may interpret $u(x, y)$ as the displacement of a *membrane*, e.g., a drum skin; the inhomogeneity $f(x, y)$ in the Poisson equation represents an external forcing over the surface of the membrane. Another example is in the thermal equilibrium of flat plates; here $u(x, y)$ represents the temperature and $f(x, y)$ an external heat source. In fluid mechanics, $u(x, y)$ represents the potential function whose gradient $\mathbf{v} = \nabla u$ is the velocity vector field of a steady planar fluid flow. Similar considerations apply to two-dimensional electrostatic and gravitational potentials. The dynamical counterparts to the Laplace equation are the two-dimensional versions of the heat and wave equations, to be analyzed in Chapter 11.

Since both the Laplace and Poisson equations describe equilibrium configurations, they almost always appear the context of boundary value problems. We seek a solution $u(x, y)$ to the partial differential equation defined at points (x, y) belonging to a bounded, open domain $\Omega \subset \mathbb{R}^2$. The solution is required to satisfy suitable conditions on the boundary of the domain, denoted by $\partial\Omega$, which will consist of one or more simple closed curves, as illustrated in Figure 4.7. As in one-dimensional boundary value problems, there are several especially important types of boundary conditions.

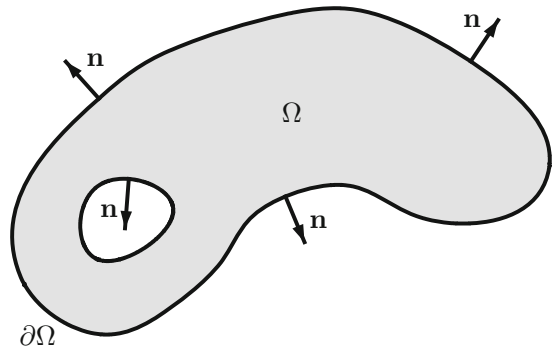


Figure 4.7. A planar domain with outward unit normals on its boundary.

The first are the *fixed* or *Dirichlet boundary conditions*, which specify the value of the function u on the boundary:

$$u(x, y) = h(x, y) \qquad \text{for} \qquad (x, y) \in \partial\Omega. \tag{4.85}$$

Under mild regularity conditions on the domain Ω, the boundary values h, and the forcing function f, the Dirichlet conditions (4.85) serve to uniquely specify the solution $u(x, y)$ to the Laplace or the Poisson equation. Physically, in the case of a free or forced membrane, the Dirichlet boundary conditions correspond to gluing the edge of the membrane to a wire at height $h(x, y)$ over each boundary point $(x, y) \in \partial\Omega$, as illustrated in Figure 4.8. A physical realization can be easily obtained by dipping the wire in a soap solution; the resulting soap film spanning the wire forms a *minimal surface*, which, if the wire is reasonably close to planar shape,† is the solution to the Dirichlet problem prescribed by the wire. Similarly, in the modeling of thermal equilibrium, a Dirichlet boundary condition represents the imposition of a prescribed temperature distribution, represented by the function h, along the boundary of the plate.

The second important class consists of the *Neumann boundary conditions*

$$\frac{\partial u}{\partial \mathbf{n}} = \nabla u \cdot \mathbf{n} = k(x, y) \qquad \text{on} \qquad \partial\Omega, \tag{4.86}$$

in which the *normal derivative* of the solution u on the boundary is prescribed. In general, $\mathbf{n}$ denotes the *unit outwards normal* to the boundary $\partial\Omega$, i.e., the vector of unit length, $\| \mathbf{n} \| = 1$, that is orthogonal to the tangent to the boundary and points *away* from the domain; see Figure 4.7. For example, in thermomechanics, a Neumann boundary condition specifies the heat flux out of a plate through its boundary. The "no-flux" or homogeneous Neumann boundary conditions, where $k(x, y) \equiv 0$, correspond to a fully insulated boundary. In the case of a membrane, homogeneous Neumann boundary conditions correspond to a free, unattached edge of a drum. In fluid mechanics, the Neumann conditions prescribe the fluid flux through the boundary; in particular, homogeneous Neumann boundary conditions

† More generally, the minimal surface formed by the soap film solves the vastly more complicated nonlinear *minimal surface equation* $(1 + u_x^2)u_{xx} - 2\,u_x u_y u_{xy} + (1 + u_y^2)u_{yy} = 0$, which, for surfaces with small variation, i.e., with $\| \nabla u \| \ll 1$, can be approximated by the Laplace equation.

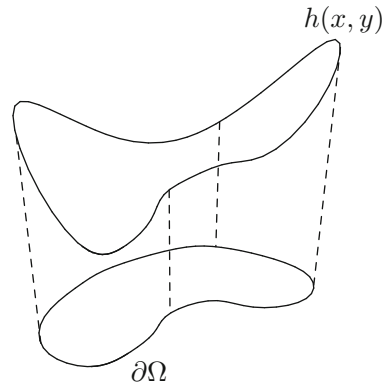


Figure 4.8. Dirichlet boundary conditions.

correspond to a solid boundary that the fluid cannot penetrate. More generally, the *Robin boundary conditions*

$$\frac{\partial u}{\partial \mathbf{n}} + \beta(x, y)\, u = k(x, y) \qquad \text{on} \qquad \partial\Omega,$$

also known as *impedance boundary conditions* due to their applications in electromagnetism, are used to model insulated plates in heat baths, or membranes attached to springs.

Finally, one can mix the previous kinds of boundary conditions, imposing, say, Dirichlet conditions on part of the boundary and Neumann conditions on the complementary part. A typical *mixed boundary value problem* has the form

$$-\Delta u = f \quad \text{in} \quad \Omega, \qquad u = h \quad \text{on} \quad D, \qquad \frac{\partial u}{\partial \mathbf{n}} = k \quad \text{on} \quad N, \tag{4.87}$$

with the boundary $\partial\Omega = D \cup N$ being the disjoint union of a "Dirichlet segment", denoted by D, and a "Neumann segment" N. For example, if u represents the equilibrium temperature in a plate, then the Dirichlet segment of the boundary is where the temperature is fixed, while the Neumann segment is insulated, or, more generally, has prescribed heat flux. Similarly, when modeling the displacement of a membrane, the Dirichlet segment is where the edge of the drum is attached to a support, while the homogeneous Neumann segment is left hanging free.

Exercises

4.3.1. (a) Solve the boundary value problem $\Delta u = 1$ for $x^2 + y^2 < 1$ and $u(x, y) = 0$ for $x^2 + y^2 = 1$ directly. *Hint*: The solution is a simple polynomial.
(b) Graph your solution, interpreting it as the equilibrium displacement of a circular drum under a constant gravitational force.

4.3.2. Set up the boundary value problem corresponding to the equilibrium of a circular membrane subject to a constant downwards gravitational force, half of whose boundary is glued to a flat semicircular wire, while the other half is unattached.

4.3.3. Set up the boundary value problem corresponding to the thermal equilibrium of a rectangular plate that is insulated on two of its sides, has 0° at its top edge and 100° at the

bottom edge. Where do you expect the maximum temperature to be located? What is its value? Can you find a formula for the temperature inside the plate? *Hint*: The solution is constant along horizontal lines.

4.3.4. Set up the boundary value problem corresponding to the thermal equilibrium of an insulated semi-circular plate with unit diameter, whose curved edge is kept at 0° and whose straight edge is at 50°.

4.3.5. Explain why the solution to the homogeneous Neumann boundary value problem for the Laplace equation is *not* unique.

4.3.6. Write down the Dirichlet boundary value problem for the Laplace equation on the unit square $0 \le x, y \le 1$ that is satisfied by $u(x, y) = 1 + x\,y$.

4.3.7. Write down the Neumann boundary value problem for the Poisson equation on the unit disk $x^2 + y^2 \le 1$ that is satisfied by $u(x, y) = x^3 + x\,y^2$.

♢ 4.3.8. Suppose $u(x, y)$ is a solution to the Laplace equation.
(*a*) Show that any translate $U(x, y) = u(x - a, y - b)$, where $a, b \in \mathbb{R}$, is also a solution.
(*b*) Show that the rotated function $U(x, y) = u(x\cos\theta + y\sin\theta, -x\sin\theta + y\cos\theta)$, where $-\pi < \theta \le \pi$, is also a solution.

♢ 4.3.9. (*a*) Show that if $u(x, y)$ solves the Laplace equation, then so does the rescaled function $U(x, y) = c\,u(\alpha\,x, \alpha\,y)$ for any constants c, α.
(*b*) Discuss the effect of scaling on the Dirichlet boundary value problem.
(*c*) What happens if we use different scaling factors in x and y?

Separation of Variables

Our first approach to solving the Laplace equation

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \tag{4.88}$$

will be based on the method of *separation of variables*. As in (4.64), we seek solutions that can be written as a product

$$u(x, y) = v(x)\,w(y) \tag{4.89}$$

of a function of x alone times a function of y alone. We compute

$$\frac{\partial^2 u}{\partial x^2} = v''(x)\,w(y), \qquad \frac{\partial^2 u}{\partial y^2} = v(x)\,w''(y),$$

and so

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = v''(x)\,w(y) + v(x)\,w''(y) = 0.$$

We then *separate the variables* by placing all the terms involving x on one side of the equation and all the terms involving y on the other; this is accomplished by dividing by $v(x)\,w(y)$ and then writing the resulting equation in the separated form

$$\frac{v''(x)}{v(x)} = -\frac{w''(y)}{w(y)} = \lambda. \tag{4.90}$$

As we argued in (4.65), the only way a function of x alone can be equal to a function of y alone is if both functions are equal to a common *separation constant* λ. Thus, the factors $v(x)$ and $w(y)$ must satisfy the elementary ordinary differential equations

$$v'' - \lambda\, v = 0, \qquad\qquad w'' + \lambda\, w = 0.$$

As before, the solution formulas depend on the sign of the separation constant λ. We list the resulting collection of separable harmonic functions in the following table:

Separable Solutions to Laplace's Equation

λ	$v(x)$	$w(y)$	$u(x,y) = v(x)\,w(y)$
$\lambda = -\omega^2 < 0$	$\cos \omega\, x,\ \sin \omega\, x$	$e^{-\omega\, y},\ e^{\omega\, y},$	$e^{\omega\, y}\cos \omega\, x,\quad e^{\omega\, y}\sin \omega\, x,$ $e^{-\omega\, y}\cos \omega\, x,\ \ e^{-\omega\, y}\sin \omega\, x$
$\lambda = 0$	$1,\ x$	$1,\ y$	$1,\ x,\ y,\ x\, y$
$\lambda = \omega^2 > 0$	$e^{-\omega\, x},\ e^{\omega\, x}$	$\cos \omega\, y,\ \sin \omega\, y$	$e^{\omega\, x}\cos \omega\, y,\quad e^{\omega\, x}\sin \omega\, y,$ $e^{-\omega\, x}\cos \omega\, y,\ \ e^{-\omega\, x}\sin \omega\, y$

Since Laplace's equation is a homogeneous linear system, any linear combination of solutions is also a solution. So, we can build more general solutions as finite linear combinations, or, provided we pay proper attention to convergence issues, infinite series in the separable solutions. Our goal is to solve boundary value problems, and so we must ensure that the resulting combination satisfies the boundary conditions. But this is not such an easy task, unless the underlying domain has a rather special geometry.

In fact, the only bounded domains on which we can explicitly solve boundary value problems using the preceding separable solutions are rectangles. So, we will concentrate on boundary value problems for Laplace's equation

$$\Delta u = 0 \qquad \text{on a rectangle} \qquad R = \{\, 0 < x < a,\ \ 0 < y < b \,\}. \tag{4.91}$$

To make progress, we will allow nonzero boundary values on only one of the four sides of the rectangle. To illustrate, we will focus on the following Dirichlet boundary conditions:

$$u(x,0) = f(x), \qquad u(x,b) = 0, \qquad u(0,y) = 0, \qquad u(a,y) = 0. \tag{4.92}$$

Once we know how to solve this type of problem, we can employ linear superposition to solve the general Dirichlet boundary value problem on a rectangle; see Exercise 4.3.12 for details. Other boundary conditions can be treated in a similar fashion — with the proviso that the condition on each side of the rectangle is either entirely Dirichlet or entirely Neumann or, more generally, entirely Robin with constant transfer coefficient.

To solve the boundary value problem (4.91–92), the first step is to narrow down the separable solutions to only those that respect the three homogeneous boundary conditions. The separable function $u(x,y) = v(x)\,w(y)$ will vanish on the top, right, and left sides of the rectangle, provided

$$v(0) = v(a) = 0 \qquad \text{and} \qquad w(b) = 0.$$

Referring to the preceding table, the first condition $v(0) = 0$ requires

$$v(x) = \begin{cases} \sin \omega x, & \lambda = -\omega^2 < 0, \\ x, & \lambda = 0, \\ \sinh \omega x, & \lambda = \omega^2 > 0, \end{cases}$$

where $\sinh z = \frac{1}{2}(e^z - e^{-z})$ is the usual hyperbolic sine function. However, the second and third cases cannot satisfy the second boundary condition $v(a) = 0$, and so we discard them. The first case leads to the condition

$$v(a) = \sin \omega a = 0, \qquad \text{and hence} \qquad \omega a = \pi,\ 2\pi,\ 3\pi,\ \dots .$$

The corresponding separation constants and solutions (up to constant multiple) are

$$\lambda_n = -\omega^2 = -\frac{n^2 \pi^2}{a^2}, \qquad v_n(x) = \sin \frac{n \pi x}{a}, \qquad n = 1, 2, 3, \dots . \tag{4.93}$$

Note: So far, we have merely recomputed the known eigenvalues and eigenfunctions of the familiar boundary value problem $v'' - \lambda v = 0,\ v(0) = v(a) = 0$.

Next, since $\lambda = -\omega^2 < 0$, we have $w(y) = c_1 e^{\omega y} + c_2 e^{-\omega y}$ for constants c_1, c_2. The third boundary condition $w(b) = 0$ then requires that, up to constant multiple,

$$w_n(y) = \sinh \omega (b - y) = \sinh \frac{n \pi (b - y)}{a}. \tag{4.94}$$

We conclude that the harmonic functions

$$u_n(x, y) = \sin \frac{n \pi x}{a} \sinh \frac{n \pi (b - y)}{a}, \qquad n = 1, 2, 3, \dots , \tag{4.95}$$

provide a complete list of separable solutions that satisfy the three homogeneous boundary conditions. It remains to analyze the inhomogeneous boundary condition along the bottom edge of the rectangle. To this end, let us try a linear superposition of the relevant separable solutions in the form of an infinite series

$$u(x, y) = \sum_{n=1}^{\infty} c_n u_n(x, y) = \sum_{n=1}^{\infty} c_n \sin \frac{n \pi x}{a} \sinh \frac{n \pi (b - y)}{a},$$

whose coefficients $c_1, c_2, \dots$ are to be prescribed by the remaining boundary condition. At the bottom edge, $y = 0$, we find

$$u(x, 0) = \sum_{n=1}^{\infty} c_n \sinh \frac{n \pi b}{a} \sin \frac{n \pi x}{a} = f(x), \qquad 0 \le x \le a, \tag{4.96}$$

which takes the form of a Fourier sine series for the function $f(x)$. Let

$$b_n = \frac{2}{a} \int_0^a f(x) \sin \frac{n \pi x}{a} \, dx \tag{4.97}$$

be its Fourier sine coefficients, whence $c_n = b_n / \sinh(n \pi b / a)$. We thus anticipate that the solution to the boundary value problem can be expressed as the infinite series

$$u(x, y) = \sum_{n=1}^{\infty} \frac{b_n \sin \frac{n \pi x}{a} \sinh \frac{n \pi (b - y)}{a}}{\sinh \frac{n \pi b}{a}}. \tag{4.98}$$

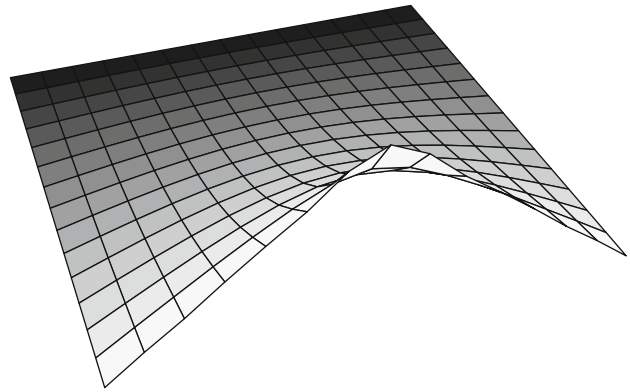


Figure 4.9. Square membrane on a wire.

Does this series actually converge to the solution to the boundary value problem? Fourier analysis says that, under very mild conditions on the boundary function $f(x)$, the answer is yes. Suppose that its Fourier coefficients are uniformly bounded,

$$| b_n | \leq M \qquad \text{for all} \qquad n \geq 1, \tag{4.99}$$

which, according to (4.27), is true whenever $f(x)$ is piecewise continuous or, more generally, integrable: $\displaystyle\int_0^a | f(x) | \, dx < \infty$. In this case, as you are asked to prove in Exercise 4.3.20, the coefficients of the Fourier sine series (4.98) go to zero exponentially fast:

$$\frac{b_n \sinh \dfrac{n\pi(b-y)}{a}}{\sinh \dfrac{n\pi b}{a}} \longrightarrow 0 \qquad \text{as} \quad n \longrightarrow \infty \quad \text{for all} \quad 0 < y \leq b, \tag{4.100}$$

and so, at each point inside the rectangle, the series can be well approximated by partial summation. Theorem 3.31 tells us that, for each $0 < y \leq b$, the solution $u(x, y)$ is an infinitely differentiable function of x. Moreover, by term-wise differentiation of the series with respect to y and use of Proposition 3.28, we also establish that the solution is infinitely differentiable with respect to y; see Exercise 4.3.21. (In fact, as we shall see, solutions to the Laplace equation are *always analytic* functions inside their domain of definition — even when their boundary values are rather rough.) Since the individual terms all satisfy the Laplace equation, we conclude that the series (4.98) is indeed a classical solution to the boundary value problem.

Example 4.4. A membrane is stretched over a wire in the shape of a unit square with one side bent in half, as graphed in Figure 4.9. The precise boundary conditions are

$$u(x,y) = \begin{cases} x, & 0 \leq x \leq \frac{1}{2}, & y = 0, \\ 1 - x, & \frac{1}{2} \leq x \leq 1, & y = 0, \\ 0, & 0 \leq x \leq 1, & y = 1, \\ 0, & x = 0, & 0 \leq y \leq 1, \\ 0, & x = 1, & 0 \leq y \leq 1. \end{cases}$$

The Fourier sine series of the inhomogeneous boundary function is readily computed:

$$f(x) = \begin{cases} x, & 0 \le x \le \frac{1}{2}, \\ 1 - x, & \frac{1}{2} \le x \le 1, \end{cases}$$
$$= \frac{4}{\pi^2}\left(\sin \pi x - \frac{\sin 3\pi x}{9} + \frac{\sin 5\pi x}{25} - \cdots\right) = \frac{4}{\pi^2}\sum_{j=0}^{\infty} (-1)^j \frac{\sin(2j+1)\pi x}{(2j+1)^2}.$$

Specializing (4.98) to $a = b = 1$, we conclude that the solution to the boundary value problem can be expressed as a Fourier series

$$u(x,y) = \frac{4}{\pi^2}\sum_{j=0}^{\infty} (-1)^j \frac{\sin(2j+1)\pi x \, \sinh(2j+1)\pi(1-y)}{(2j+1)^2 \sinh(2j+1)\pi}.$$

In Figure 4.9 we plot the sum of the first 10 terms in the series. This gives a reasonably good approximation to the actual solution, except when we are very close to the raised corner of the boundary wire — which is the point of maximal displacement of the membrane.

Exercises

4.3.10. Solve the following boundary value problems for Laplace's equation on the square
$$\Omega = \{\, 0 \le x \le \pi, \quad 0 \le y \le \pi \,\}.$$
(a) $u(x,0) = \sin^3 x$, $u(x,\pi) = 0$, $u(0,y) = 0$, $u(\pi,y) = 0$.
(b) $u(x,0) = 0$, $u(x,\pi) = 0$, $u(0,y) = \sin y$, $u(\pi,y) = 0$.
(c) $u(x,0) = 0$, $u(x,\pi) = 1$, $u(0,y) = 0$, $u(\pi,y) = 0$.
(d) $u(x,0) = 0$, $u(x,\pi) = 0$, $u(0,y) = 0$, $u(\pi,y) = y(\pi - y)$.

♢ 4.3.11. (a) Explain how to use linear superposition to solve the boundary value problem
$$\Delta u = 0, \qquad u(x,0) = f(x), \qquad u(x,b) = g(x), \qquad u(0,y) = h(y), \qquad u(a,y) = k(y),$$
on the rectangle $R = \{\, 0 < x < a, \; 0 < y < b \,\}$, by splitting it into four separate boundary value problems for which each of the solutions vanishes on three sides of the rectangle.
(b) Write down a series formula for the resulting solution.

4.3.12. Solve the following Dirichlet problems for Laplace's equation on the unit square $S = \{\, 0 < x, y < 1 \,\}$. *Hint*: Use superposition as in Exercise 4.3.11.
(a) $u(x,0) = \sin \pi x$, $u(x,1) = 0$, $u(0,y) = \sin \pi y$, $u(1,y) = 0$;
(b) $u(x,0) = 1$, $u(x,1) = 0$, $u(0,y) = 1$, $u(1,y) = 0$;
(c) $u(x,0) = 1$, $u(x,1) = 1$, $u(0,y) = 0$, $u(1,y) = 0$;
(d) $u(x,0) = x$, $u(x,1) = 1 - x$, $u(0,y) = y$, $u(1,y) = 1 - y$.

4.3.13. Solve the following mixed boundary value problems for Laplace's equation $\Delta u = 0$ on the square $S = \{\, 0 < x, y < \pi \,\}$.
(a) $u(x,0) = \sin \frac{1}{2} x$, $u_y(x,\pi) = 0$, $u(0,y) = 0$, $u_x(\pi,y) = 0$;
(b) $u(x,0) = \sin \frac{1}{2} x$, $u_y(x,\pi) = 0$, $u_x(0,y) = 0$, $u_x(\pi,y) = 0$;
(c) $u(x,0) = x$, $u(x,\pi) = 0$, $u_x(0,y) = 0$, $u_x(\pi,y) = 0$;
(d) $u(x,0) = x$, $u(x,\pi) = 0$, $u(0,y) = 0$, $u_x(\pi,y) = 0$.

4.3.14. Find the solution to the boundary value problem
$$\Delta u = 0, \qquad \begin{array}{c} u_y(x,0) = u_y(x,2) = 0, \\ u(0,y) = 2\cos \pi y - 1, \quad u(1,y) = 0, \end{array} \qquad \begin{array}{l} 0 < x < 1, \\ 0 < y < 2. \end{array}$$

4.3.15. Find the solution to the boundary value problem

$$\Delta u = 0, \qquad \begin{array}{c} u(x,0) = 2\cos 7\pi x - 4, \quad u(x,1) = 5\cos 3\pi x, \\ u_x(0,y) = u_x(1,y) = 0, \end{array} \qquad 0 < x, y < 1.$$

4.3.16. Let $u(x,y)$ be the solution to the boundary value problem

$$\Delta u = 0,\ u(x,-1) = f(x),\ u(x,1) = 0,\ u(-1,y) = 0,\ u(1,y) = 0,\ -1 < x < 1,\ -1 < y < 1.$$

(a) *True or false*: If $f(-x) = -f(x)$ is odd, then $u(0,y) = 0$ for all $-1 \le y \le 1$.
(b) *True or false*: If $f(0) = 0$, then $u(0,y) = 0$ for all $-1 \le y \le 1$.
(c) Under what conditions on $f(x)$ is $u(x,0) = 0$ for all $-1 \le x \le 1$?

4.3.17. Use separation of variables to solve the following boundary value problem:

$$u_{xx} + 2\,u_y + u_{yy} = 0, \quad u(x,0) = 0, \quad u(x,1) = f(x), \quad u(0,y) = 0, \quad u(1,y) = 0.$$

4.3.18. Use separation of variables to solve the Helmholtz boundary value problem $\Delta u = u$, $u(x,0) = 0$, $u(x,1) = f(x)$, $u(0,y) = 0$, $u(1,y) = 0$, on the unit square $0 < x, y < 1$.

♢ 4.3.19. Provide the details for the derivation of (4.94).

♢ 4.3.20. Justify the statement that if $|\, b_n \,| \le M$ are uniformly bounded, then the coefficients given in (4.100) go to zero exponentially fast as $n \to \infty$ for any $0 < y \le b$.

♢ 4.3.21. Let $u(x,y)$ denote the solution to the boundary value problem (4.91–92). (a) Write down the Fourier sine series for $\partial u/\partial y$. (b) Prove that $\partial u/\partial y$ is an infinitely differentiable function of x. (c) Justify the same result for the functions $\partial^k u/\partial y^k$ for each $k \ge 0$. *Hint*: Don't forget that $u(x,y)$ solves the Laplace equation.

Polar Coordinates

The method of separation of variables can be successfully exploited in certain other very special geometries. One particularly important case is a circular disk. To be specific, let us take the disk to have radius 1 and be centered at the origin. Consider the Dirichlet boundary value problem

$$\Delta u = 0, \qquad x^2 + y^2 < 1, \qquad \text{and} \qquad u = h, \qquad x^2 + y^2 = 1, \tag{4.101}$$

so that the function $u(x,y)$ satisfies the Laplace equation on the unit disk and the specified Dirichlet boundary conditions on the unit circle. For example, $u(x,y)$ might represent the displacement of a circular drum that is attached to a wire of height

$$h(x,y) = h(\cos\theta, \sin\theta) \equiv h(\theta), \qquad -\pi < \theta \le \pi, \tag{4.102}$$

at each point $(x,y) = (\cos\theta, \sin\theta)$ on its edge.

The rectangular separable solutions are not particularly helpful in this situation, and so we look for solutions that are better adapted to a circular geometry. This inspires us to adopt *polar coordinates*

$$x = r\cos\theta, \qquad y = r\sin\theta, \qquad \text{or} \qquad r = \sqrt{x^2 + y^2}\,, \qquad \theta = \tan^{-1}\frac{y}{x}\,, \tag{4.103}$$

and write the solution $u(r,\theta)$ as a function thereof.

Warning: We will often retain the same symbol, e.g., u, when rewriting a function in a different coordinate system. This is the convention of tensor analysis, physics, and

differential geometry, [**3**], that treats the function (scalar field) as an intrinsic object, which is concretely realized through its formula in any chosen coordinate system. For instance, if $u(x,y) = x^2 + 2y$ in rectangular coordinates, then its expression in polar coordinates is $u(r,\theta) = (r\cos\theta)^2 + 2r\sin\theta$, *not* $r^2 + 2\theta$. This convention avoids the inconvenience of having to devise new symbols when changing coordinates.

We need to relate derivatives with respect to x and y to those with respect to r and θ. Performing a standard multivariate chain rule computation based on (4.103), we obtain

$$\begin{aligned} \frac{\partial}{\partial r} &= \cos\theta\,\frac{\partial}{\partial x} + \sin\theta\,\frac{\partial}{\partial y}\,, \\ \frac{\partial}{\partial \theta} &= -\,r\sin\theta\,\frac{\partial}{\partial x} + r\cos\theta\,\frac{\partial}{\partial y}\,, \end{aligned} \qquad \text{so} \qquad \begin{aligned} \frac{\partial}{\partial x} &= \cos\theta\,\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\,\frac{\partial}{\partial \theta}\,, \\ \frac{\partial}{\partial y} &= \sin\theta\,\frac{\partial}{\partial r} + \frac{\cos\theta}{r}\,\frac{\partial}{\partial \theta}\,. \end{aligned} \tag{4.104}$$

Applying the squares of the latter differential operators to $u(r,\theta)$, we find, after a calculation in which many of the terms cancel, the *polar coordinate form of the Laplace equation*:

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\,\frac{\partial u}{\partial r} + \frac{1}{r^2}\,\frac{\partial^2 u}{\partial \theta^2} = 0. \tag{4.105}$$

The boundary conditions are imposed on the unit circle $r = 1$, and so, by (4.102), take the form

$$u(1,\theta) = h(\theta). \tag{4.106}$$

Keep in mind that, in order to be single-valued functions of x, y, the solution $u(r,\theta)$ and its boundary values $h(\theta)$ must both be 2π–periodic functions of the angular coordinate:

$$u(r,\theta + 2\pi) = u(r,\theta), \qquad\qquad h(\theta + 2\pi) = h(\theta). \tag{4.107}$$

Polar *separation of variables* is based on the ansatz

$$u(r,\theta) = v(r)\,w(\theta), \tag{4.108}$$

which assumes that the solution is a product of functions of the individual variables. Substituting (4.108) into the polar form (4.105) of Laplace's equation yields

$$v''(r)\,w(\theta) + \frac{1}{r}\,v'(r)\,w(\theta) + \frac{1}{r^2}\,v(r)\,w''(\theta) = 0.$$

We now separate variables by moving all the terms involving r onto one side of the equation and all the terms involving θ onto the other. This is accomplished by first multiplying the equation by $r^2/\big(v(r)\,w(\theta)\big)$ and then moving the final term to the right-hand side:

$$\frac{r^2\,v''(r) + r\,v'(r)}{v(r)} = -\,\frac{w''(\theta)}{w(\theta)} = \lambda.$$

As in the rectangular case, a function of r can equal a function of θ if and only if both are equal to a common separation constant, which we call λ. The partial differential equation thus splits into a pair of ordinary differential equations

$$r^2\,v'' + r\,v' - \lambda\,v = 0, \qquad\qquad w'' + \lambda\,w = 0, \tag{4.109}$$

that will prescribe the separable solution (4.108). Observe that both have the form of an eigenfunction equation in which the separation constant λ plays the role of the eigenvalue. We are, as always, interested only in nonzero solutions.

We have already solved the eigenvalue problem for $w(\theta)$. According to (4.107), $w(\theta + 2\pi) = w(\theta)$ must be a 2π–periodic function. Therefore, by our earlier discussion, this periodic boundary value problem has the nonzero eigenfunctions

$$1, \qquad \sin n\theta, \qquad \cos n\theta, \qquad n = 1, 2, \dots, \tag{4.110}$$

corresponding to the eigenvalues (separation constants)

$$\lambda = n^2, \qquad n = 0, 1, 2, \dots.$$

With the value of λ fixed, the linear ordinary differential equation for the radial component,

$$r^2 v'' + r\, v' - n^2 v = 0, \tag{4.111}$$

does not have constant coefficients. But, fortunately, it has the form of a second-order *Euler ordinary differential equation*, [**23**, **89**], and hence can be readily solved by substituting the power ansatz $v(r) = r^k$. (See also Exercise 4.3.23.) Note that

$$v'(r) = k r^{k-1}, \qquad v''(r) = k(k-1)\, r^{k-2},$$

and hence, by substituting into the differential equation,

$$r^2 v'' + r\, v' - n^2 v = \big[\, k(k-1) + k - n^2 \,\big] r^k = (k^2 - n^2) r^k.$$

Thus, r^k is a solution if and only if

$$k^2 - n^2 = 0, \qquad \text{and hence} \qquad k = \pm n.$$

For $n \neq 0$, we have found the two linearly independent solutions:

$$v_1(r) = r^n, \qquad v_2(r) = r^{-n}, \qquad n = 1, 2, \dots. \tag{4.112}$$

When $n = 0$, the power ansatz yields only the constant solution. But in this case, the equation $r^2 v'' + r\, v' = 0$ is effectively of first order and linear in v', and hence readily integrated. This provides the two independent solutions

$$v_1(r) = 1, \qquad v_2(r) = \log r, \qquad n = 0. \tag{4.113}$$

Combining (4.110) and (4.112–113), we produce the complete list of separable polar coordinate solutions to the Laplace equation:

$$\begin{array}{llll} 1, & r^n \cos n\theta, & r^n \sin n\theta, & \\ \log r, & r^{-n} \cos n\theta, & r^{-n} \sin n\theta, & n = 1, 2, 3, \dots. \end{array} \tag{4.114}$$

Now, the solutions in the top row of (4.114) are continuous (in fact analytic) at the origin, where $r = 0$, whereas the solutions in the bottom row have singularities as $r \to 0$. The latter are not of use in the present situation, since we require that the solution remain bounded and smooth — even at the center of the disk. Thus, we should use only the nonsingular solutions to concoct a candidate series solution

$$u(r, \theta) \;=\; \frac{a_0}{2} \;+\; \sum_{n=1}^{\infty} \big(a_n r^n \cos n\theta + b_n r^n \sin n\theta \big). \tag{4.115}$$

The coefficients a_n, b_n will be prescribed by the boundary conditions (4.106). Substituting $r = 1$, we obtain

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

We recognize this as a standard Fourier series (3.29) (with θ replacing x) for the 2π periodic function $h(\theta)$. Therefore,

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} h(\theta) \cos n\theta \, d\theta, \qquad\qquad b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} h(\theta) \sin n\theta \, d\theta, \tag{4.116}$$

are precisely its Fourier coefficients, cf. (3.35). In this manner, we have produced a series solution (4.115) to the boundary value problem (4.105–106).

Remark: Introducing the complex variable

$$z = x + \mathrm{i}\, y = r\, e^{\,\mathrm{i}\,\theta} = r\cos\theta + \mathrm{i}\, r \sin\theta \tag{4.117}$$

allows us to write

$$z^n = r^n\, e^{\,\mathrm{i}\, n\theta} = r^n \cos n\theta + \mathrm{i}\, r^n \sin n\theta. \tag{4.118}$$

Therefore, the nonsingular separable solutions are the *harmonic polynomials*

$$r^n \cos n\theta = \operatorname{Re} z^n, \qquad\qquad r^n \sin n\theta = \operatorname{Im} z^n. \tag{4.119}$$

The first few are listed in the following table:

n	$\operatorname{Re} z^n$	$\operatorname{Im} z^n$
0	1	0
1	x	y
2	$x^2 - y^2$	$2\,x\,y$
3	$x^3 - 3\,x\,y^2$	$3\,x^2\,y - y^3$
4	$x^4 - 4\,x^2y^2 + y^4$	$4\,x^3\,y - 4\,x\,y^3$

Their general expression is obtained using the Binomial Formula:

$$\begin{aligned} z^n &= (x + \mathrm{i}\, y)^n \\ &= x^n + n\, x^{n-1}(\mathrm{i}\, y) + \binom{n}{2} x^{n-2}(\mathrm{i}\, y)^2 + \binom{n}{3} x^{n-3}(\mathrm{i}\, y)^3 + \;\cdots\; + (\mathrm{i}\, y)^n \\ &= x^n + \mathrm{i}\, n\, x^{n-1} y - \binom{n}{2} x^{n-2} y^2 - \mathrm{i} \binom{n}{3} x^{n-3} y^3 + \;\cdots\;, \end{aligned}$$

where

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!} \tag{4.120}$$

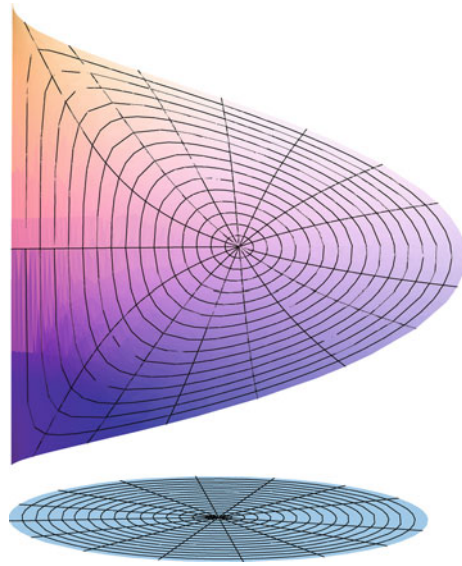


Figure 4.10. Membrane attached to a helical wire.

are the usual *binomial coefficients*. Separating the real and imaginary terms, we produce the explicit formulae

$$\begin{aligned} r^n \cos n\theta &= \operatorname{Re} z^n = x^n - \binom{n}{2} x^{n-2} y^2 + \binom{n}{4} x^{n-4} y^4 + \cdots, \\ r^n \sin n\theta &= \operatorname{Im} z^n = n x^{n-1} y - \binom{n}{3} x^{n-3} y^3 + \binom{n}{5} x^{n-5} y^5 + \cdots, \end{aligned} \tag{4.121}$$

for the two independent harmonic polynomials of degree n.

Example 4.5. Consider the Dirichlet boundary value problem on the unit disk with

$$u(1, \theta) = \theta \qquad \text{for} \qquad -\pi < \theta < \pi. \tag{4.122}$$

The boundary data can be interpreted as a wire in the shape of a single turn of a spiral helix sitting over the unit circle. The wire has a single jump discontinuity, of magnitude 2π, at the boundary point $(-1, 0)$. The required Fourier series

$$h(\theta) = \theta \sim 2\left(\sin\theta - \frac{\sin 2\theta}{2} + \frac{\sin 3\theta}{3} - \frac{\sin 4\theta}{4} + \cdots\right)$$

was already computed in Example 3.3. Therefore, invoking our solution formula (4.115–116), we have

$$u(r, \theta) = 2\left(r\sin\theta - \frac{r^2 \sin 2\theta}{2} + \frac{r^3 \sin 3\theta}{3} - \frac{r^4 \sin 4\theta}{4} + \cdots\right) \tag{4.123}$$

is the desired solution, which is plotted in Figure 4.10. In fact, this series can be explicitly summed. In view of (4.119) and the usual formula (A.13) for the complex logarithm, we have

$$u = 2 \operatorname{Im}\left(z - \frac{z^2}{2} + \frac{z^3}{3} - \frac{z^4}{4} + \cdots\right) = 2 \operatorname{Im} \log(1 + z) = 2\psi, \tag{4.124}$$

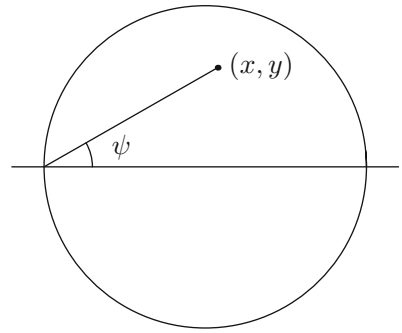


Figure 4.11. Geometric construction of the solution.

where

$$\psi = \tan^{-1} \frac{y}{1+x}$$

is the angle that the line passing through the two points (x, y) and $(-1, 0)$ makes with the x-axis, as sketched in Figure 4.11. You should try to convince yourself that, on the unit circle, $2\,\psi = \theta$ has the correct boundary values. Observe that, even though the boundary values are discontinuous, the solution is an analytic function inside the disk.

In fact, unlike the rectangular series (4.98), the general polar series solution formula (4.115) can, in fact, be summed in closed form! If we substitute the explicit Fourier formulae (4.116) into (4.115) — remembering to change the integration variable to, say, ϕ to avoid a notational conflict — we obtain

$$\begin{aligned} u(r,\theta) &= \frac{a_0}{2} + \sum_{n=1}^{\infty} \big(a_n r^n \cos n\theta + b_n r^n \sin n\theta \big) \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} h(\phi)\, d\phi + \sum_{n=1}^{\infty} \left[\frac{r^n \cos n\theta}{\pi} \int_{-\pi}^{\pi} h(\phi) \cos n\phi \, d\phi + \frac{r^n \sin n\theta}{\pi} \int_{-\pi}^{\pi} h(\phi) \sin n\phi \, d\phi \right] \\ &= \frac{1}{\pi} \int_{-\pi}^{\pi} h(\phi) \left[\frac{1}{2} + \sum_{n=1}^{\infty} r^n \big(\cos n\theta \cos n\phi + \sin n\theta \sin n\phi \big) \right] d\phi \\ &= \frac{1}{\pi} \int_{-\pi}^{\pi} h(\phi) \left[\frac{1}{2} + \sum_{n=1}^{\infty} r^n \cos n(\theta - \phi) \right] d\phi. \end{aligned} \tag{4.125}$$

We next show how to sum the final series. Using (4.118), we can write it as the real part of a geometric series:

$$\begin{aligned} \frac{1}{2} + \sum_{n=1}^{\infty} r^n \cos n\theta &= \operatorname{Re} \left(\frac{1}{2} + \sum_{n=1}^{\infty} z^n \right) = \operatorname{Re} \left(\frac{1}{2} + \frac{z}{1-z} \right) = \operatorname{Re} \left(\frac{1+z}{2\,(1-z)} \right) \\ &= \operatorname{Re} \left(\frac{(1+z)(1-\overline{z})}{2\,|\,1-z\,|^2} \right) = \frac{\operatorname{Re}\,(1+z-\overline{z}-|\,z\,|^2)}{2\,|\,1-z\,|^2} = \frac{1-|\,z\,|^2}{2\,|\,1-z\,|^2} = \frac{1-r^2}{2\,(1+r^2-2\,r\cos\theta)}\,, \end{aligned}$$

which is known as the *Poisson kernel*. Substituting back into (4.125) establishes the important *Poisson Integral Formula* for the solution to the boundary value problem.

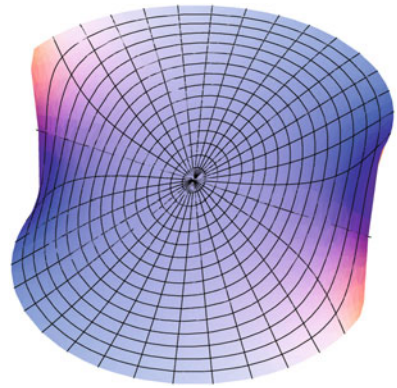


Figure 4.12. Equilibrium temperature of a disk.

Theorem 4.6. *The solution to the Laplace equation in the unit disk subject to Dirichlet boundary conditions* $u(1,\theta) = h(\theta)$ *is*

$$u(r,\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} h(\phi)\,\frac{1-r^2}{1+r^2-2r\cos(\theta-\phi)}\,d\phi. \tag{4.126}$$

Example 4.7. A uniform metal disk of unit radius has half of its circular boundary held at 1°, while the other half is held at 0°. Our task is to find the equilibrium temperature $u(x,y)$. In other words, we seek the solution to the Dirichlet boundary value problem

$$\Delta u = 0, \qquad x^2+y^2 < 1, \qquad u(x,y) = \begin{cases} 1, & x^2+y^2 = 1, \quad y > 0, \\ 0, & x^2+y^2 = 1, \quad y < 0. \end{cases} \tag{4.127}$$

In polar coordinates, the boundary data is a (periodic) step function

$$h(\theta) = \begin{cases} 1, & 0 < \theta < \pi, \\ 0, & -\pi < \theta < 0. \end{cases}$$

Therefore, according to the Poisson formula (4.126), the solution is given by†

$$u(r,\theta) = \frac{1}{2\pi}\int_{0}^{\pi} \frac{1-r^2}{1+r^2-2r\cos(\theta-\phi)}\,d\phi = \begin{cases} 1 - \dfrac{1}{\pi}\tan^{-1}\left(\dfrac{1-r^2}{2r\sin\theta}\right), & 0 < \theta < \pi, \\ \dfrac{1}{2}, & \theta = 0, \pm\pi, \\ -\dfrac{1}{\pi}\tan^{-1}\left(\dfrac{1-r^2}{2r\sin\theta}\right), & -\pi < \theta < 0, \end{cases} \tag{4.128}$$

† The detailed derivation of the final expressions is left to the reader as Exercise 4.3.40.

where we use the principal branch $-\frac{1}{2}\pi < \tan^{-1} t < \frac{1}{2}\pi$ of the inverse tangent. Reverting to rectangular coordinates, we find that the equilibrium temperature has the explicit formula

$$u(x,y) = \begin{cases} 1 - \dfrac{1}{\pi}\,\tan^{-1}\left(\dfrac{1-x^2-y^2}{2\,y}\right), & x^2+y^2<1, \quad y>0, \\ \dfrac{1}{2}\,, & x^2+y^2<1, \quad y=0, \\ -\dfrac{1}{\pi}\,\tan^{-1}\left(\dfrac{1-x^2-y^2}{2\,y}\right), & x^2+y^2<1, \quad y<0. \end{cases} \tag{4.129}$$

The result is depicted in Figure 4.12.

Averaging, the Maximum Principle, and Analyticity

Let us investigate some important consequences of the Poisson integral formula (4.126). First, setting $r = 0$ yields

$$u(0,\theta) = \frac{1}{2\,\pi}\int_{-\pi}^{\pi} h(\phi)\,d\phi. \tag{4.130}$$

The left-hand side is the value of u at the origin — the center of the disk — and so independent of θ; the right-hand side is the *average* of its boundary values around the unit circle. This formula is a particular instance of an important general fact.

Theorem 4.8. *Let $u(x,y)$ be harmonic inside a disk of radius a centered at a point (x_0,y_0) with piecewise continuous (or, more generally, integrable) boundary values on the circle $C = \{\,(x-x_0)^2 + (y-y_0)^2 = a^2\,\}$. Then its value at the center of the disk is equal to the average of its values on the boundary circle*:

$$u(x_0,y_0) = \frac{1}{2\,\pi\,a}\oint_C u\,ds = \frac{1}{2\,\pi}\int_{-\pi}^{\pi} u(x_0 + a\cos\theta, y_0 + a\sin\theta)\,d\theta. \tag{4.131}$$

Proof: We use the scaling and translation symmetries of the Laplace equation, cf. Exercises 4.3.8–9, to map the disk of radius a centered at (x_0,y_0) to the unit disk centered at the origin. Specifically, we set

$$U(x,y) = u(x_0 + a\,x, y_0 + a\,y). \tag{4.132}$$

An easy chain rule computation proves that $U(x,y)$ also satisfies the Laplace equation on the unit disk $x^2 + y^2 < 1$, with boundary values

$$h(\theta) = U(\cos\theta,\sin\theta) = u(x_0 + a\cos\theta, y_0 + a\sin\theta).$$

Therefore, by (4.130),

$$U(0,0) = \frac{1}{2\,\pi}\int_{-\pi}^{\pi} h(\theta)\,d\theta = \frac{1}{2\,\pi}\int_{-\pi}^{\pi} U(\cos\theta,\sin\theta)\,d\theta.$$

Replacing U by its formula (4.132) produces the desired result. *Q.E.D.*

An important consequence of the integral formula (4.131) is the *Strong Maximum Principle* for harmonic functions.

Theorem 4.9. *Let u be a* nonconstant *harmonic function defined on a bounded domain Ω and continuous on $\partial\Omega$. Then u achieves its maximum and minimum values only at boundary points of the domain. In other words, if*

$$m = \min\{\, u(x,y) \mid (x,y) \in \partial\Omega \,\}, \qquad M = \max\{\, u(x,y) \mid (x,y) \in \partial\Omega \,\},$$

are, respectively, its maximum and minimum values on the boundary, then

$$m < u(x,y) < M \qquad \text{at all interior points} \qquad (x,y) \in \Omega.$$

Proof: Let $M^\star \geq M$ be the maximum value of u on all of $\overline{\Omega} = \Omega \cup \partial\Omega$, and assume $u(x_0, y_0) = M^\star$ at some interior point $(x_0, y_0) \in \Omega$. Theorem 4.8 implies that $u(x_0, y_0)$ equals its average over any circle C centered at (x_0, y_0) that bounds a closed disk contained in Ω. Since u is continuous and $\leq M^\star$ on C, its average must be strictly less than $M^\star$ — except in the trivial case in which it is constant and equal to $M^\star$ on all of C. Thus, our assumption implies that $u(x,y) = M^\star = u(x_0, y_0)$ for all (x,y) belonging to any circle $C \subset \Omega$ centered at (x_0, y_0). Since Ω is connected, this allows us to conclude[†] that $u(x,y) = M^\star$ is constant throughout Ω, in contradiction to our original assumption.

A similar argument works for the minimum; alternatively, one can interchange maximum and minimum by replacing u by $-u$. *Q.E.D.*

Physically, if we interpret $u(x,y)$ as the vertical displacement of a membrane stretched over a wire, then Theorem 4.9 says that, in the absence of external forcing, the membrane cannot have any internal bumps — its highest and lowest points are necessarily on the boundary of the domain. This reconfirms our physical intuition: the restoring force exerted by the stretched membrane will serve to flatten any bump, and hence a membrane with a local maximum or minimum cannot be in equilibrium. A similar interpretation holds for heat conduction. A body in thermal equilibrium will achieve its maximum and minimum temperature only at boundary points. Indeed, thermal energy would flow away from any internal maximum, or towards any local minimum, and so if the body contained a local maximum or minimum in its interior, it could not remain in thermal equilibrium.

The Maximum Principle immediately implies the uniqueness of solutions to the Dirichlet boundary value problem for both the Laplace and Poisson equations:

Theorem 4.10. *If u and $\widetilde{u}$ both satisfy the same Poisson equation $-\Delta u = f = -\Delta \widetilde{u}$ within a bounded domain Ω, and $u = \widetilde{u}$ on $\partial\Omega$, then $u \equiv \widetilde{u}$ throughout Ω.*

Proof: By linearity, the difference $v = u - \widetilde{u}$ satisfies the homogeneous boundary value problem $\Delta v = 0$ in Ω and $v = 0$ on $\partial\Omega$. Our assumption implies that the maximum and minimum boundary values of v are both $0 = m = M$. Theorem 4.9 implies that $v(x,y) \equiv 0$ at all $(x,y) \in \Omega$, and hence $u \equiv \widetilde{u}$ everywhere in Ω. *Q.E.D.*

Finally, let us discuss the analyticity of harmonic functions. In view of (4.119), the n^{th} order term in the polar series solution (4.115), namely,

$$a_n r^n \cos n\theta + b_n r^n \sin n\theta = a_n \operatorname{Re} z^n + b_n \operatorname{Im} z^n = \operatorname{Re}\big[\,(a_n - \mathrm{i}\, b_n) z^n\,\big],$$

is, in fact, a homogeneous polynomial in (x,y) of degree n. This means that, when written in rectangular coordinates x and y, (4.115) is, in fact, a *power series* for the harmonic

[†] You are asked to supply the details in Exercise 4.3.42.

function $u(x,y)$. It is well known, [**8**, **23**, **97**], that any convergent power series converges to an analytic function — in this case $u(x,y)$. Moreover, the power series must, in fact, be the *Taylor series* for $u(x,y)$ based at the origin, and so its coefficients are multiples of the derivatives of u at $x = y = 0$. Details are worked out in Exercise 4.3.49.

We can adapt this argument to prove analyticity of *all* solutions to the Laplace equation. Note especially the contrast with the wave equation, which has many non-analytic solutions.

Theorem 4.11. *A harmonic function is analytic at every point in the interior of its domain of definition.*

Proof: Let $u(x,y)$ be a solution to the Laplace equation on the open domain $\Omega \subset \mathbb{R}^2$. Let $\mathbf{x}_0 = (x_0, y_0) \in \Omega$, and choose $a > 0$ such that the closed disk of radius a centered at $\mathbf{x}_0$ is entirely contained within Ω:

$$D_a(\mathbf{x}_0) = \{\, \| \mathbf{x} - \mathbf{x}_0 \| \leq a \,\} \subset \Omega,$$

where $\| \cdot \|$ is the usual Euclidean norm. Then the function $U(x,y)$ defined by (4.132) is harmonic on the unit disk, with well-defined boundary values. Thus, by the preceding remarks, $U(x,y)$ is analytic at every point inside the unit disk, and hence so is

$$u(x,y) = U\left(\frac{x - x_0}{a}\,, \frac{y - y_0}{a} \right)$$

at every point (x,y) in the interior of the disk $D_a(\mathbf{x}_0)$. Since $\mathbf{x}_0 \in \Omega$ was arbitrary, this establishes the analyticity of u throughout the domain. *Q.E.D.*

This concludes our discussion of the method of separation of variables for the planar Laplace equation and some of its important consequences. The method can be used in a few other special coordinate systems. See [**78**, **79**] for a complete account, including the fascinating connections with the underlying symmetry properties of the equation.

Exercises

4.3.22. Solve the following Euler differential equations by use of the power ansatz:

(*a*) $x^2 u'' + 5x u' - 5u = 0$, (*b*) $2x^2 u'' - x u' - 2u = 0$, (*c*) $x^2 u'' - u = 0$,

(*d*) $x^2 u'' + x u' - 3u = 0$, (*e*) $3x^2 u'' - 5x u' - 3u = 0$, (*f*) $\dfrac{d^2u}{dx^2} + \dfrac{2}{x}\dfrac{du}{dx} = 0$.

♢ 4.3.23. (*i*) Show that if $u(x)$ solves the *Euler differential equation*

$$a x^2 \frac{d^2u}{dx^2} + b x \frac{du}{dx} + c u = 0, \tag{4.133}$$

then $v(y) = u(e^y)$ solves a linear constant-coefficient differential equation.

(*ii*) Use this technique to solve the Euler differential equations in Exercise 4.3.22.

4.3.24. (*a*) Use the method in Exercise 4.3.23 to solve an Euler equation whose characteristic equation has a double root $r_1 = r_2 = r$. (*b*) Solve the specific equations

(*i*) $x^2 u'' - x u' + u = 0$, (*ii*) $\dfrac{d^2u}{dx^2} + \dfrac{1}{x}\dfrac{du}{dx} = 0$.

4.3.25. Solve the following boundary value problems:
(a) $\Delta u = 0,\ x^2 + y^2 < 1,\quad u = x^3,\ x^2 + y^2 = 1;$
(b) $\Delta u = 0,\ x^2 + y^2 < 2,\quad u = \log(x^2 + y^2),\ x^2 + y^2 = 1;$
(c) $\Delta u = 0,\ x^2 + y^2 < 4,\quad u = x^4,\ x^2 + y^2 = 4;$
(d) $\Delta u = 0,\ x^2 + y^2 < 1,\quad \dfrac{\partial u}{\partial \mathbf{n}} = x,\ x^2 + y^2 = 1.$

4.3.26. Let $u(x, y)$ be the solution to the boundary value problem $u_{xx} + u_{yy} = 0,\ x^2 + y^2 < 1,$ $u(x, y) = x^2,\ x^2 + y^2 = 1$. Find $u(0, 0)$.

♡ 4.3.27. (a) Find the equilibrium temperature on a disk of radius 1 when half the boundary is held at 1° and the other half is held at -1°. (b) Find the equilibrium temperature on a half-disk of radius 1 when the temperature is held to 1° on the curved edge and 0° on the straight edge. (c) Find the equilibrium temperature on a half disk of radius 1 when the temperature is held to 0° on the curved edge and 1° on the straight edge.

4.3.28. Find the solution to Laplace's equation $u_{xx} + u_{yy} = 0$ on the semi-disk $x^2 + y^2 < 1$, $y > 0$, that satisfies the boundary conditions $u(x, 0) = 0$ for $-1 < x < 1$ and $u(x, y) = y^3$ for $x^2 + y^2 = 1,\ y > 0$.

4.3.29. Find the equilibrium temperature on a half-disk of radius 1 when the temperature is held to 1° on the curved edge, while the straight edge is insulated.

4.3.30. Solve the Dirichlet boundary value problem for the Laplace equation on the pie wedge $W = \{\, 0 < \theta < \frac{1}{4}\pi,\ \ 0 < r < 1 \,\}$, when the nonzero boundary data $u(1, \theta) = h(\theta)$ appears only on the curved portion of its boundary.

4.3.31. Find a harmonic function $u(x, y)$ defined on the annulus $\frac{1}{2} < r < 1$ subject to the constant Dirichlet boundary conditions $u = a$ on $r = \frac{1}{2}$ and $u = b$ on $r = 1$.

4.3.32. Boiling water flows continually through a long circular metal pipe of inner radius 1 cm and outer radius 1.2 cm placed in an ice water bath. *True or false*: The temperature at the midpoint, at radius 1.1 cm, is 50°. If false, what is the temperature at this point?

4.3.33. Write out the series solution to the boundary value problem $u(1, \theta) = 0,\ u(2, \theta) = h(\theta)$, for the Laplace equation on an annulus $1 < r < 2$. *Hint*: Use all of the separable solutions listed in (4.114).

4.3.34. Solve the following boundary value problems for the Laplace equation on the annulus $1 < r < 2$: (a) $u(1, \theta) = 0,\ u(2, \theta) = 1,$ (b) $u(1, \theta) = 0,\ u(2, \theta) = \cos\theta,$
(c) $u(1, \theta) = \sin 2\theta,\ u(2, \theta) = \cos 2\theta,$ (d) $u_r(1, \theta) = 0,\ u(2, \theta) = 1,$
(e) $u_r(1, \theta) = 0,\ u(2, \theta) = \sin 2\theta,$ (f) $u_r(1, \theta) = 0,\ u_r(2, \theta) = 1,$
(g) $u_r(1, \theta) = 2,\ u_r(2, \theta) = 1.$

4.3.35. Solve the following boundary value problems for the Laplace equation on the semi-annular domain $D = \{\, 1 < x^2 + y^2 < 2,\ y > 0 \,\}$:
(a) $u(x, y) = 0,\quad x^2 + y^2 = 1,\quad u(x, y) = 1,\quad x^2 + y^2 = 2,\quad u(x, 0) = 0;$
(b) $u(x, y) = 0,\quad x^2 + y^2 = 1 \text{ or } 2,\quad u(x, 0) = 0,\quad x > 0,\quad u(x, 0) = 1,\quad x < 0.$

4.3.36. Solve the following boundary value problem:
$(x^2 + y^2)(u_{xx} + u_{yy}) + 2x\,u_x + 2y\,u_y = 0,\ x^2 + y^2 < 1,\ u(x, y) = 1 + 3x,\ x^2 + y^2 = 1.$

♢ 4.3.37. Justify the chain rule computation (4.104). Then justify formula (4.105) for the Laplacian in polar coordinates.

4.3.38. Suppose $\displaystyle\int_{-\pi}^{\pi} |\, h(\theta) \,|\, d\theta < \infty$. Prove that (4.115) converges uniformly to the solution to the boundary value problem (4.101) on any smaller disk $D_{r_\star} = \{\, r \le r_\star < 1 \,\} \subsetneq D_1$.

4.3.39. Prove directly that (4.124) satisfies the boundary conditions (4.122).

♢ 4.3.40. Justify the integration formula in (4.128).

4.3.41. Provide a complete proof that (4.129) is indeed the solution to the boundary value problem (4.127).

♢ 4.3.42. Complete the proof of Theorem 4.9 by showing that $u(x, y) = M^\star$ for all $(x, y) \in \Omega$. *Hint*: Join (x_0, y_0) to (x, y) by a curve $C \subset \Omega$ of finite length, and use the preceding part of the proof to inductively deduce the existence of a finite sequence of points $(x_i, y_i) \in C$, $i = 0, \dots, n$, with $(x_n, y_n) = (x, y)$, and such that $u(x_i, y_i) = M^\star$.

♢ 4.3.43. Derive the analogue of the Poisson integral formula for the solution to the Neumann boundary value problem $\Delta u = 0$, $x^2 + y^2 < 1$, $\partial u/\partial \mathbf{n} = h$, $x^2 + y^2 = 1$, on the unit disk. Pay careful attention to the existence and uniqueness of solutions in your formulation.

4.3.44. Give an example of a solution to Poisson's equation on the unit disk that achieves its maximum at an interior point. Interpret your construction physically.

4.3.45. Let $p(x, y)$ be a polynomial (not necessarily harmonic). Suppose $u(x, y)$ is harmonic and equals $p(x, y)$ on the unit circle $x^2 + y^2 = 1$. Prove that $u(x, y)$ is a harmonic polynomial.

4.3.46. Write down an integral formula for the solution to the Dirichlet boundary value problem on a disk of radius $R > 0$, namely, $\Delta u = 0$, $x^2 + y^2 < R^2$, $u = h$, $x^2 + y^2 = R^2$.

4.3.47. State and prove a one-dimensional version of Theorem 4.8. Does the analogue of Theorem 4.9 hold?

4.3.48. A unit area square plate has 100° temperature on its top edge and 0° on its three other edges. *True or false*: The temperature at the center equals the average edge temperature.

♢ 4.3.49. Let $u(x, y)$ be a harmonic function on the unit disk with boundary values $h(\theta)$ when $r = 1$. Using the fact that (4.115) is the Taylor series for $u(x, y)$ at the origin: (*a*) Find integral formulas for its partial derivatives $u_x(0, 0)$, $u_y(0, 0)$, involving the boundary values $h(\theta)$. (*b*) Generalize part (*a*) to the second-order derivatives $u_{xx}(0, 0)$, $u_{xy}(0, 0)$, $u_{yy}(0, 0)$.

4.3.50. Prove that if $u(x, y)$ is a bounded harmonic function defined on all of $\mathbb{R}^2$, then u is constant. *Hint*: First generalize Exercise 4.3.49(*a*) to find the value of its gradient, $\nabla u(x_0, y_0)$, in terms of the values of u on a circle of radius a centered at (x_0, y_0). Then see what happens when the radius of the circle goes to ∞.

4.4 Classification of Linear Partial Differential Equations

We have, at last, been introduced to the three paradigmatic linear second-order partial differential equations for functions of two variables. The homogeneous versions are

(*a*)	The wave equation:	$u_{tt} - c^2 u_{xx} = 0,$	*hyperbolic*,
(*b*)	The heat equation:	$u_t - \gamma u_{xx} = 0,$	*parabolic*,
(*c*)	Laplace's equation:	$u_{xx} + u_{yy} = 0,$	*elliptic.*

The last column indicates the equation's *type*, in accordance with the standard taxonomy of partial differential equations; an explanation will appear momentarily. The wave, heat, and Laplace equations are the prototypical representatives of these three fundamental genres. Each genre has its own distinctive analytic features, physical manifestations, and even numerical solution schemes. Equations governing vibrations, such as the wave equation,

are typically hyperbolic. Equations modeling diffusion, such as the heat equation, are parabolic. Hyperbolic and parabolic equations both typically represent dynamical processes, and so one of the independent variables is identified as time. On the other hand, equations modeling equilibrium phenomena, including the Laplace and Poisson equations, are usually elliptic, and involve only spatial variables. Elliptic partial differential equations are associated with boundary value problems, whereas parabolic and hyperbolic equations require initial and initial-boundary value problems.

The classification theory of real linear second-order partial differential equations for a scalar-valued function $u(t, x)$ depending on two variables† proceeds as follows. The most general such equation has the form

$$L[u] = A\,u_{tt} + B\,u_{tx} + C\,u_{xx} + D\,u_t + E\,u_x + F\,u = G, \tag{4.134}$$

where the coefficients A, B, C, D, E, F are all allowed to be functions of (t, x), as is the inhomogeneity or forcing function $G(t, x)$. The equation is *homogeneous* if and only if $G \equiv 0$. We assume that at least one of the leading coefficients A, B, C is not identically zero, since otherwise, the equation degenerates to a first-order equation.

The key quantity that determines the *type* of such a partial differential equation is its *discriminant*

$$\Delta = B^2 - 4\,AC. \tag{4.135}$$

This should (and for good reason) remind the reader of the discriminant of the quadratic equation

$$Q(x, y) = A\,x^2 + B\,x\,y + C\,y^2 + D\,x + E\,y + F = 0, \tag{4.136}$$

whose solutions trace out a plane curve — a conic section. In the nondegenerate cases, the discriminant (4.135) fixes its geometric type:

- a hyperbola when $\Delta > 0$,
- a parabola when $\Delta = 0$,
- an ellipse when $\Delta < 0$.

This motivates the choice of terminology used to classify second-order partial differential equations.

Definition 4.12. At a point (t, x), the linear second-order partial differential equation (4.134) is called

- *hyperbolic* if $\Delta(t, x) > 0$,
- *parabolic* if $\Delta(t, x) = 0$, but $A^2 + B^2 + C^2 \neq 0$,
- *elliptic* if $\Delta(t, x) < 0$,
- *singular* if $A = B = C = 0$.

In particular:

- The wave equation $u_{tt} - u_{xx} = 0$ has discriminant $\Delta = 4$, and is hyperbolic.
- The heat equation $u_{xx} - u_t = 0$ has discriminant $\Delta = 0$, and is parabolic.
- The Poisson equation $u_{tt} + u_{xx} = -f$ has discriminant $\Delta = -4$, and is elliptic.

† For equilibrium equations, we identify t with the space variable y.

Example 4.13. When the coefficients A, B, C vary, the type of the partial differential equation may not remain fixed over the entire domain. Equations that change type are less common, as well as being much harder to analyze and solve, both analytically and numerically. One example arising in the theory of supersonic aerodynamics, [**44**], is the *Tricomi equation*

$$x\,\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} = 0. \tag{4.137}$$

Comparing with (4.134), we find that

$$A = x, \quad B = 0, \quad C = -1, \qquad \text{while} \qquad D = E = F = G = 0.$$

The discriminant in this particular case is

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

and hence the equation is hyperbolic when $x > 0$, elliptic when $x < 0$, and parabolic on the transition line $x = 0$. In the physical model, the hyperbolic region corresponds to subsonic flow, while the supersonic regions are of elliptic type. The transitional parabolic boundary represents the shock line between the sub- and super-sonic regions — the familiar sonic boom as an airplane crosses the sound barrier.

While this tripartite classification into hyperbolic, parabolic, and elliptic equations initially appears in the bivariate context, the terminology, underlying properties, and associated physical models carry over to second-order partial differential equations in higher dimensions. Most of the partial differential equations arising in applications fall into one of these three categories, and it is fair to say that the field of partial differential equations splits into three distinct subfields. Or rather four subfields, the last containing all the equations, including higher-order equations, that do not fit into the preceding categorization. (One important example appears in Section 8.5.)

Remark: The classification into hyperbolic, parabolic, elliptic, and singular types carries over as stated to *quasilinear* second-order equations, whose coefficients $A, \dots, G$ are allowed to depend on u and its first-order derivatives, u_t, u_x. Here the type of the equation can vary with both the point in the domain and the particular solution being considered. Even more generally, for a *fully nonlinear* second-order partial differential equation

$$H(t, x, u, u_t, u_x, u_{tt}, u_{tx}, u_{xx}) = 0, \tag{4.138}$$

one defines its *discriminant* to be

$$\Delta = \left(\frac{\partial H}{\partial u_{tx}}\right)^2 - 4\,\frac{\partial H}{\partial u_{tt}}\,\frac{\partial H}{\partial u_{xx}}\,. \tag{4.139}$$

Its sign determines the type of the equation as above — again depending on the point in the domain and the solution under consideration.

Exercises

4.4.1. Plot the following conic sections and classify their type:
(*a*) $x^2 + 3y^2 = 1$, (*b*) $x\,y + x + y = 4$, (*c*) $x^2 - x\,y + y^2 = x - 2\,y$,
(*d*) $x^2 + 2\,x\,y + y^2 + y = 1$, (*e*) $x^2 - 2\,y^2 = 6\,x + 8\,y + 1$.

4.4.2. Determine the type of the following partial differential equations:
(a) $u_{tt} + 3\,u_{xx} = 0$, (b) $u_{tx} + u_t + u_x = u$, (c) $u_{tt} + u_t + u_x = 0$,
(d) $u_{tt} - u_{tx} + u_{xx} = u$, (e) $u_{tt} + 4\,u_{tx} + 4\,u_{xx} = u_t$, (f) $u_{tx} + u_{xx} = 0$.

4.4.3. Consider the partial differential equation $x\,u_{tt} + (t + x)\,u_{xx} = 0$. At what points of the plane is the equation elliptic? hyperbolic? parabolic? degenerate?

4.4.4. Answer Exercise 4.4.3 for the equations
(a) $x^2\,u_{xx} + x\,u_x + u_{yy} = 0$, (b) $\partial_x(x\,u_x) = \partial_y(y\,u_y)$, (c) $u_t = \partial_x[\,(x + t)u_x\,]$,
(d) $\nabla \cdot (c(x, y)\nabla u) = u$, where $c(x, y)$ is a given function.

4.4.5. Steady flow of air past an airplane is modeled by the partial differential equation $(m^2 - 1)u_{xx} + u_{yy} = 0$, in which x is the flight direction, y the transverse direction, and $m \geq 0$ is the *Mach number* — the ratio of the airplane's speed to the speed of sound. Show that the equation is hyperbolic for subsonic flight, but elliptic for supersonic flight.

4.4.6. Show that the second-order partial differential equation

$$-\frac{\partial}{\partial x}\left(p(x,y)\,\frac{\partial u}{\partial x}\right) - \frac{\partial}{\partial y}\left(q(x,y)\,\frac{\partial u}{\partial y}\right) + r(x,y)\,u = f(x,y)$$

is elliptic if and only if $p(x, y)$ and $q(x, y)$ are nonzero and have the same sign.

♢ 4.4.7. *True or false*: The type of a linear second-order partial differential equation is not affected by a change of independent variables: $\tau = \varphi(t, x)$, $\xi = \psi(t, x)$.

4.4.8. Let $v(t, x) = a(t, x)\,u(t, x) + b(t, x)$, where a, b are fixed functions with $a \neq 0$. Suppose u is a solution to a second-order linear partial differential equation. Prove that v also solves a linear partial differential equation of the same type.

♢ 4.4.9. *True or false*: The polar coordinate form (4.105) of the Laplace equation is elliptic.

4.4.10. Rewrite the Laplace equation $u_{xx} + u_{yy} = 0$ in terms of *parabolic coordinates* ξ, η, as defined by the equations $x = \xi^2 - \eta^2$, $y = 2\,\xi\,\eta$. Is the resulting equation elliptic?

♢ 4.4.11. Prove that the complex change of variables $x = x$, $t = \mathrm{i}\,y$, maps the Laplace equation $u_{xx} + u_{yy} = 0$ to the wave equation $u_{tt} = u_{xx}$. Explain why the type of a partial differential equation is *not* necessarily preserved under a complex change of variables.

♡ 4.4.12. Suppose, against all advice, we pose the elliptic Laplace equation as an initial value problem, namely

$$u_{tt} = -\,u_{xx} \qquad \text{for} \qquad 0 < x < 1, \qquad t > 0,$$
$$u(0, x) = f(x), \qquad u_t(0, x) = 0, \qquad 0 \leq x \leq 1, \qquad u(t, 0) = 0 = u(t, 1), \qquad t \geq 0.$$

(a) Prove that for any positive integer $n > 0$, the function $u_n(t, x) = \dfrac{\sin n\,\pi\,t\,\cosh n\,\pi\,x}{n}$ satisfies the initial value problem. Determine the initial condition $u_n(0, x) = f_n(x)$.
(b) Prove that, as $n \to \infty$, the initial condition $f_n(x) \to 0$ becomes vanishingly small, whereas, at any $t > 0$, the solution value $u_n\left(t, \frac{1}{2}\right) \to \infty$.
(c) Explain why this represents an ill-posed problem.

4.4.13. The *minimal surface equation* $(1+u_x^2)u_{xx} - 2\,u_x u_y u_{xy} + (1+u_y^2)u_{yy} = 0$ is (a) hyperbolic, (b) parabolic, (c) elliptic, (d) singular, (e) of variable type depending on the point in the domain, or (f) of variable type depending on the solution and the point in the domain.

Characteristics and the Cauchy Problem

In Chapter 2, we discovered that the characteristic curves guide the behavior of solutions to first-order partial differential equations. Characteristics play a similarly fundamental role in the analysis of more general hyperbolic partial differential equations and systems.

In particular, they provide a mechanism for distinguishing among the various classes of second-order partial differential equations.

As above, we will focus our attention on partial differential equations involving two independent variables. The starting point is the general initial value problem, also known as the Cauchy problem, in honor of the prolific nineteenth-century French mathematician Augustin–Louis Cauchy, justly famous for his wide-ranging contributions throughout mathematics and its applications, including the Cauchy–Schwarz inequality, many of the fundamental concepts in complex analysis, as well as the foundations of elasticity and materials science. The general *Cauchy problem* specifies appropriate initial data along a smooth curve[†] $\Gamma \subset \mathbb{R}^2$ and seeks a solution to the partial differential equation that assumes the given initial data on Γ. In all our examples, the curve in question has been a straight line, e.g., the x–axis, but one could easily envisage more general situations. If the partial differential equation has order n, then the *Cauchy data* consists of the values of the dependent variable u along with all its partial differential equations up to order $n-1$ on the curve Γ. For most curves, there is a unique solution $u(t,x)$ to the partial differential equation that achieves the specified values along Γ. More rigorously, if we are in the analytic category, meaning that the partial differential equation, the curve, and the Cauchy data are all specified by analytic functions, then the fundamental *Cauchy–Kovalevskaya Theorem* guarantees the existence of an analytic solution $u(t,x)$ to the Cauchy problem near any point on the initial curve. The statement of proof of this important theorem, due to Cauchy and, in general form, the influential nineteenth-century Russian mathematician Sofia Kovalevskaya, relies on the construction of convergent power series for the desired solution and would take us too far afield. We refer the interested reader to [**35**, **44**]. The exceptional curves, for which the Cauchy–Kovalevskaya Existence Theorem does not apply, are called the characteristics of the underlying partial differential equations.

More prosaically, a curve Γ will be called *non-characteristic* for the given partial differential equation if one can determine the values of *all* the derivatives of u along Γ from the specified Cauchy data. Indeed, the determination of the values of the higher-order derivatives along the curve is a necessary preliminary step towards establishing the Cauchy–Kovalevskaya existence result. As we will now show, this requirement serves to distinguish the characteristic and non-characteristic curves for the examples we have already encountered, and hence to lead to their characterization in much more general contexts.

To illustrate the preceding requirement, let us begin with a first-order linear partial differential equation of the form

$$\frac{\partial u}{\partial t} + c(t,x)\,\frac{\partial u}{\partial x} = f(t,x). \tag{4.140}$$

Let $\Gamma \subset \mathbb{R}^2$ be a smooth curve parametrized[§] by $\mathbf{x}(s) = \big(\,t(s), x(s)\,\big)^T$, where smoothness necessitates that its tangent vector not vanish: $\mathbf{x}'(s) = (\,dt/ds, dx/ds)^T \neq \mathbf{0}$. Since the equation is of order $n = 1$, the Cauchy data requires specifying the values of the dependent variable u only along Γ — in other words, the function

$$h(s) = u\big(t(s), x(s)\big). \tag{4.141}$$

[†] More generally, for partial differential equations in $m > 2$ independent variables, the curve is replaced by a hypersurface $S \subset \mathbb{R}^m$ of dimension $m-1$.

[§] The parameter s could be the arc length, but this is not required. See also Exercise 4.4.20.

The curve will be non-characteristic if we can then determine the values of the derivatives of u along Γ, starting with

$$\frac{\partial u}{\partial t}\big(t(s),x(s)\big), \qquad \frac{\partial u}{\partial x}\big(t(s),x(s)\big). \tag{4.142}$$

To this end, let us differentiate the Cauchy data (4.141): applying the chain rule, we obtain

$$h'(s) = \frac{d}{ds}\,u\big(t(s),x(s)\big) = \frac{\partial u}{\partial t}\big(t(s),x(s)\big)\,\frac{dt}{ds} + \frac{\partial u}{\partial x}\big(t(s),x(s)\big)\,\frac{dx}{ds}\,. \tag{4.143}$$

On the other hand, we are assuming that $u(t,x)$ solves the partial differential equation (4.140) at all points in its domain of definition. In particular, at points on the curve Γ, the partial differential equation requires

$$\frac{\partial u}{\partial t}\big(t(s),x(s)\big) + c\big(t(s),x(s)\big)\,\frac{\partial u}{\partial x}\big(t(s),x(s)\big) = f\big(t(s),x(s)\big). \tag{4.144}$$

We can regard (4.143–144) as a pair of inhomogeneous linear algebraic equations, which can be uniquely solved for the as yet unknown quantities (4.142), *unless* the determinant of their coefficient matrix vanishes:

$$\det\begin{pmatrix} 1 & c\big(t(s),x(s)\big) \\ dt/ds & dx/ds \end{pmatrix} = \frac{dx}{ds} - c\big(t(s),x(s)\big)\,\frac{dt}{ds} = 0. \tag{4.145}$$

This condition serves to define a *characteristic curve* for the first-order partial differential equation (4.140). In particular, if the curve is parametrized by $s = t$, i.e., can be identified with the graph of a function $x = g(t)$, then the characteristic condition (4.145) reduces to

$$\frac{dx}{dt} = c(t,x), \tag{4.146}$$

thus reproducing our original definition of characteristic curve, as in (2.18) and, more generally, Exercise 2.2.26. On the other hand, if the determinant (4.145) is nonzero, then one can solve (4.143–144) for the values of the first-order derivatives (4.142) along Γ. Further differentiation of these conditions proves that one can, in fact, determine the values of all the higher-order derivatives of the solution u along the curve, which is hence non-characteristic.

Next, consider a nonsingular linear second-order partial differential equation of the form (4.134). Since the equation has order $n = 2$, the Cauchy data along a curve Γ parametrized as above consists of the values of the function and its first derivatives:

$$u\big(t(s),x(s)\big), \qquad \frac{\partial u}{\partial t}\big(t(s),x(s)\big), \qquad \frac{\partial u}{\partial x}\big(t(s),x(s)\big). \tag{4.147}$$

However, the latter cannot be specified independently. Indeed, given the value of the dependent variable, $h(s) = u\big(t(s),x(s)\big)$, along Γ, its derivative

$$h'(s) = \frac{d}{ds}\,u\big(t(s),x(s)\big) = \frac{\partial u}{\partial t}\big(t(s),x(s)\big)\,\frac{dt}{ds} + \frac{\partial u}{\partial x}\big(t(s),x(s)\big)\,\frac{dx}{ds} \tag{4.148}$$

prescribes a particular combination of the two first-order derivatives. Thus, once the value of one derivative of u on Γ is known, the other is automatically fixed by the relation (4.148). For example, if $dx/ds \neq 0$, we can use (4.148) to determine $u_x\big(t(s),x(s)\big)$, knowing $u\big(t(s),x(s)\big)$ and $u_t\big(t(s),x(s)\big)$. Similarly, if we differentiate the values of the first-order

derivatives with respect to the curve parameter, we can determine two combinations of second-order derivatives along the curve Γ:

$$\begin{aligned} \frac{d}{ds}\,\frac{\partial u}{\partial t}\big(t(s),x(s)\big) &= \frac{\partial^2 u}{\partial t^2}\big(t(s),x(s)\big)\,\frac{dt}{ds} + \frac{\partial^2 u}{\partial t\,\partial x}\big(t(s),x(s)\big)\,\frac{dx}{ds}\,, \\ \frac{d}{ds}\,\frac{\partial u}{\partial x}\big(t(s),x(s)\big) &= \frac{\partial^2 u}{\partial t\,\partial x}\big(t(s),x(s)\big)\,\frac{dt}{ds} + \frac{\partial^2 u}{\partial x^2}\big(t(s),x(s)\big)\,\frac{dx}{ds}\,. \end{aligned} \tag{4.149}$$

On the other hand, the partial differential equation (4.134) induces yet a third relation among the second-order partial derivatives u_{tt}, u_{tx}, u_{xx}. These three linear equations can be uniquely solved for values of these derivatives on Γ if and only if the determinant of their coefficient matrix is nonzero:

$$\det\begin{pmatrix} A(t,x) & B(t,x) & C(t,x) \\ dt/ds & dx/ds & 0 \\ 0 & dt/ds & dx/ds \end{pmatrix} = A(t,x)\left(\frac{dx}{ds}\right)^2 - B(t,x)\,\frac{dt}{ds}\,\frac{dx}{ds} + C(t,x)\left(\frac{dt}{ds}\right)^2 = 0. \tag{4.150}$$

We conclude that a smooth curve $\mathbf{x}(s) = \big(t(s),x(s)\big)^T \subset \mathbb{R}^2$ is a *characteristic curve* for the nonsingular linear second-order partial differential equation (4.134) whenever its tangent vector $\mathbf{x}'(s) = (dt/ds, dx/ds)^T \neq \mathbf{0}$ satisfies the quadratic *characteristic equation* (4.150). Conversely, if the curve is non-characteristic, meaning that its tangent does not satisfy (4.150) anywhere, then one can, with some further work, determine all the higher-order derivatives of the solution $u(t,x)$ along Γ, and then, at least in the analytic category, prove existence of a solution to the Cauchy problem, [**35**].

According to Exercise 4.4.20, the status of a curve as characteristic or not does not depend on the choice of parametrization. In particular, if the curve is given by the graph of the function $x = x(t)$, which we parametrize by $s = t$, then the characteristic equation (4.150) takes the form of a quadratically nonlinear first-order ordinary differential equation

$$A(t,x)\left(\frac{dx}{dt}\right)^2 - B(t,x)\,\frac{dx}{dt} + C(t,x) = 0, \tag{4.151}$$

whose solutions are characteristic curves of the second-order partial differential equation.

Warning: If $A(t,x) = 0$, then the partial differential equation admits characteristic curves with vertical tangents that cannot be parametrized by $s = t$. For example, if $A(t,x) \equiv 0$, then the vertical lines e.g., $t =$ constant, $x = s$, are characteristic, satisfying (4.150), but do not appear as solutions to (4.151).

For example, consider the hyperbolic wave equation

$$u_{tt} - c^2\,u_{xx} = 0.$$

According to (4.151), any characteristic curve that is given by the graph of $x(t)$ must solve

$$\left(\frac{dx}{dt}\right)^2 - c^2 = 0, \qquad \text{which implies that} \qquad \frac{dx}{dt} = \pm\,c.$$

Thus, in accordance with our previous analysis, the characteristic curves are the straight lines of slope $\pm\,c$, and there are two characteristic curves passing through each point of the (t,x)–plane. On the other hand, the elliptic Laplace equation

$$u_{tt} + u_{xx} = 0$$

has no (real) characteristic curves, since the characteristic equation (4.150) reduces to

$$\left(\frac{dx}{ds}\right)^2 + \left(\frac{dt}{ds}\right)^2 = 0,$$

and t_s and x_s are not allowed to vanish simultaneously. Finally, for the parabolic heat equation

$$u_{xx} - u_t = 0,$$

the characteristic curve equation (4.150) is simply

$$\left(\frac{dt}{ds}\right)^2 = 0$$

(since the first-derivative term plays no role), and so there is only one characteristic curve passing through each point, namely the vertical line $t =$ constant. Observe that the standard initial value problem $u(0, x) = f(x)$ for the heat equation takes place on a characteristic curve — the x–axis — but does not take the form of a Cauchy problem, which would also require specifying the first-order derivatives $u_t(0, x), u_x(0, x)$ there. And indeed, the standard initial value problem is not well-posed near the characteristic x–axis for negative $t < 0$.

In general, the number of real solutions to the nondegenerate quadratic characteristic curve equation (4.150) depends on its discriminant $\Delta = B^2 - 4\,AC$: In the hyperbolic case, $\Delta > 0$, and there are two real characteristic curves passing through each point; in the parabolic case, $\Delta = 0$, and there is just one real characteristic curve passing through each point; in the elliptic case, $\Delta < 0$, and there are no real characteristic curves. In this manner, elliptic, parabolic, and hyperbolic partial differential equations are distinguished by the number of (real) characteristic curves passing through a point — namely, zero, one, and two, respectively. First-order partial differential equations are also viewed as *hyperbolic*, since they always admit real characteristic curves.

With further analysis, [**35**, **70**, **122**], it can be shown that, as with the wave equation, signals and disturbances propagate along characteristic curves. Thus, hyperbolic equations share many qualitative properties with the wave equation, with signals moving in two different directions. For example, light rays move along characteristic curves, and are thereby subject to the optical phenomena of refraction and focusing. Similarly, since the characteristic curves for the parabolic heat equation are the vertical lines, this indicates that the effect of a disturbance at a point $(t, x) = (t_0, x_0)$ is simultaneously felt along the entire contemporaneous vertical line $t = t_0$. This has the implication that disturbances in the heat equation propagate at infinite speed — a counterintuitive fact that will be further expounded on in Section 8.1. Elliptic equations have no characteristics, and as a consequence, do not support propagating signals; indeed, the effect of a localized disturbance is immediately felt throughout the domain. For example, even when an external force is concentrated near a single point, it displaces the entire membrane.

Exercises

4.4.14. Find and graph the real characteristic curves for each of the partial differential equations in Exercise 4.4.2.

4.4.15. Graph the characteristic curves for the Tricomi equation (4.137) in its hyperbolic region. What happens to the characteristics as one approaches the parabolic transition boundary?

4.4.16. *True or false*: The characteristic curves of the *Helmholtz equation* $u_{xx} + u_{yy} - u = 0$ are circles.

4.4.17. (*a*) At what points of the plane is the partial differential equation $x\,u_{xx} + y\,u_{yy} = 0$ elliptic? parabolic? hyperbolic? (*b*) How many characteristics are there through the point $(1, -1)$? (*c*) Find them explicitly.

4.4.18. Consider the partial differential equation $u_{xx} + y\,u_{xy} = y^2$.
(*a*) On which regions of the (x, y)–plane is the equation elliptic? parabolic? hyperbolic?
(*b*) Find the characteristics in the hyperbolic region.
(*c*) Find the general solution in the hyperbolic region. *Hint*: Use characteristic coordinates.

4.4.19. Find a partial differential equation whose characteristic curves are:
(*a*) the lines $x - y = a$, $x + 2\,y = b$, where $a, b \in \mathbb{R}$ are arbitrary constants;
(*b*) the exponential curves $y = c\,e^x$ for $c \in \mathbb{R}$;
(*c*) the concentric circles $x^2 + y^2 = a$ for $a \geq 0$, and the rays $y = b\,x$.

♢ 4.4.20. Prove that any reparametrization of a characteristic curve for a given second-order linear partial differential equation is also a characteristic curve.

4.4.21. *True or false*: You can uniquely recover a second-order partial differential equation by knowing all its characteristic curves.

♢ 4.4.22. Prove that any invertible change of variables, as in Exercise 4.4.7, maps the characteristic curves of the original linear partial differential equation to the characteristic curves of the transformed equation. Thus, characteristic curves are intrinsic: they do not depend on the parametrization, nor on the coordinates used to represent the partial differential equation.

6.3 Green's Functions for the Planar Poisson Equation

Now we develop the Green's function approach to solving boundary value problems involving the two-dimensional Poisson equation (4.84). As before, the Green's function is characterized as the solution to the homogeneous boundary value problem in which the inhomogeneity is a concentrated unit impulse — a delta function. The solution to the general forced boundary value problem is then obtained via linear superposition, that is, as a convolution integral with the Green's function.

However, before proceeding, we need to quickly review some basic facts concerning vector calculus in the plane. The student may wish to consult a standard multivariable calculus text, e.g., [**8**, **108**], for additional details.

Calculus in the Plane

Let $\mathbf{x} = (x, y)$ denote the usual Cartesian coordinates on $\mathbb{R}^2$. The term *scalar field* is synonymous with a real-valued function $u(x, y)$, defined on a domain $\Omega \subset \mathbb{R}^2$. A vector-valued function

$$\mathbf{v}(\mathbf{x}) = \mathbf{v}(x, y) = \begin{pmatrix} v_1(x, y) \\ v_2(x, y) \end{pmatrix} \tag{6.70}$$

is known as a (planar) *vector field*. A vector field assigns a vector $\mathbf{v}(x, y) \in \mathbb{R}^2$ to each point $(x, y) \in \Omega$ in its domain of definition, and hence defines a function $\mathbf{v}\colon \Omega \to \mathbb{R}^2$. Physical examples include velocity vector fields of fluid flows, heat flux fields in thermodynamics, and gravitational and electrostatic force fields.

The *gradient* operator ∇ maps a scalar field $u(x, y)$ to the vector field

$$\nabla u = \begin{pmatrix} \partial u/\partial x \\ \partial u/\partial y \end{pmatrix}. \tag{6.71}$$

The scalar field u is often referred to as a *potential function* for its gradient vector field $\mathbf{v} = \nabla u$. On a connected domain Ω, the potential, when it exists, is uniquely determined up to addition of a constant.

The *divergence* of the planar vector field $\mathbf{v} = (\, v_1, v_2 \,)^T$ is the scalar field

$$\nabla \cdot \mathbf{v} = \operatorname{div} \mathbf{v} = \frac{\partial v_1}{\partial x} + \frac{\partial v_2}{\partial y}\,. \tag{6.72}$$

Its *curl* is defined as

$$\nabla \times \mathbf{v} = \operatorname{curl} \mathbf{v} = \frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y}\,. \tag{6.73}$$

Notice that the curl of a planar vector field is a scalar field. (In contrast, in three dimensions, the curl of a vector field is another vector field.) Given a smooth potential $u \in \mathrm{C}^2$, the curl of its gradient vector field automatically vanishes:

$$\nabla \times \nabla u = \frac{\partial}{\partial x}\frac{\partial u}{\partial y} - \frac{\partial}{\partial y}\frac{\partial u}{\partial x} \equiv 0,$$

by the equality of mixed partials. Thus, a necessary condition for a vector field $\mathbf{v}$ to admit a potential is that it be *irrotational*, meaning $\nabla \times \mathbf{v} = 0$; this condition is sufficient if

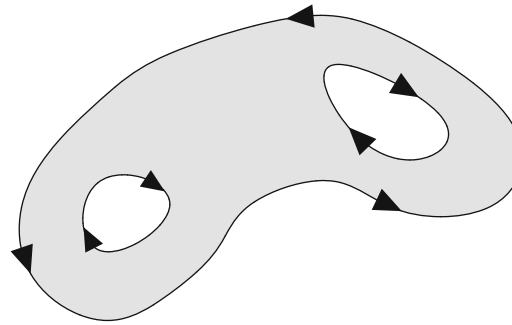


Figure 6.10. Orientation of the boundary of a planar domain.

the underlying domain Ω is *simply connected*, i.e., has no holes. On the other hand, the divergence of a gradient vector field coincides with the Laplacian of the potential function:

$$\nabla \cdot \nabla u = \Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\,. \tag{6.74}$$

A vector field is *incompressible* if it has zero divergence: $\nabla \cdot \mathbf{v} = 0$; for the velocity vector field of a steady-state fluid flow, incompressibility means that the fluid does not change volume. (Water is, for all practical purposes, an incompressible fluid.) Therefore, an irrotational vector field with potential u is also incompressible if and only if the potential solves the Laplace equation $\Delta u = 0$.

Remark: Because of formula (6.74), the Laplacian operator is also sometimes written as $\Delta = \nabla^2$. The factorization of the Laplacian into the product of the divergence and the gradient operators is, in fact, of great importance, and underlies its "self-adjointness", a fundamental property whose ramifications will be explored in depth in Chapter 9.

Let $\Omega \subset \mathbb{R}^2$ be a bounded domain whose boundary $\partial\Omega$ consists of one or more piecewise smooth closed curves. We orient the boundary so that the domain is always on one's left as one goes around the boundary curve(s). Figure 6.10 sketches a domain with two holes; its three boundary curves are oriented according to the directions of the arrows. Note that the outer boundary curve is traversed in a counterclockwise direction, while the two inner boundary curves are oriented clockwise.

Green's Theorem, first formulated by George Green to use in his seminal study of partial differential equations and potential theory, relates certain double integrals over a domain to line integrals around its boundary. It should be viewed as the extension of the Fundamental Theorem of Calculus to double integrals.

Theorem 6.13. *Let* $\mathbf{v}(\mathbf{x})$ *be a smooth*[†] *vector field defined on a bounded domain* $\Omega \subset \mathbb{R}^2$. *Then the line integral of* $\mathbf{v}$ *around the boundary* $\partial\Omega$ *equals the double integral of its curl over the domain*:

$$\iint_\Omega \nabla \times \mathbf{v}\; dx\, dy = \oint_{\partial\Omega} \mathbf{v} \cdot d\mathbf{x}, \tag{6.75}$$

[†] To be precise, we require $\mathbf{v}$ to be continuously differentiable within the domain, and continuous up to the boundary, so $\mathbf{v} \in \mathrm{C}^0(\overline{\Omega}) \cap \mathrm{C}^1(\Omega)$, where $\overline{\Omega} = \Omega \cup \partial\Omega$ denotes the closure of the domain Ω.

or, in full detail,

$$\iint_\Omega \left(\frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y} \right) dx\, dy = \oint_{\partial\Omega} v_1\, dx + v_2\, dy\,. \tag{6.76}$$

Example 6.14. Let us apply Green's Theorem 6.13 to the particular vector field $\mathbf{v} = (\, y, 0\,)^T$. Since $\nabla \times \mathbf{v} \equiv -1$, we obtain

$$\oint_{\partial\Omega} y\, dx = \iint_\Omega (-1)\, dx\, dy = -\operatorname{area} \Omega. \tag{6.77}$$

This means that we can determine the area of a planar domain by computing the negative of the indicated line integral around its boundary.

For later purposes, we rewrite the basic Green identity (6.75) in an equivalent "divergence form". Given a planar vector field $\mathbf{v} = (\, v_1, v_2\,)^T$, let

$$\mathbf{v}^\perp = \begin{pmatrix} -\,v_2 \\ v_1 \end{pmatrix} \tag{6.78}$$

denote the "perpendicular" vector field. We note that its curl

$$\nabla \times \mathbf{v}^\perp = \frac{\partial v_1}{\partial x} + \frac{\partial v_2}{\partial y} = \nabla \cdot \mathbf{v} \tag{6.79}$$

coincides with the divergence of the original vector field.

When we replace $\mathbf{v}$ in Green's identity (6.75) by $\mathbf{v}^\perp$, the result is

$$\iint_\Omega \nabla \cdot \mathbf{v}\; dx\, dy = \iint_\Omega \nabla \times \mathbf{v}^\perp\, dx\, dy = \oint_{\partial\Omega} \mathbf{v}^\perp \cdot\, d\mathbf{x} = \oint_{\partial\Omega} \mathbf{v} \cdot \mathbf{n}\; ds,$$

where $\mathbf{n}$ denotes the *unit outwards normal* to the boundary of our domain, while ds denotes the arc-length element along the boundary curve. This yields the *divergence form* of Green's Theorem:

$$\iint_\Omega \nabla \cdot \mathbf{v}\; dx\, dy = \oint_{\partial\Omega} \mathbf{v} \cdot \mathbf{n}\; ds. \tag{6.80}$$

Physically, if $\mathbf{v}$ represents the velocity vector field of a steady-state fluid flow, then the line integral in (6.80) represents the net fluid flux out of the region Ω. As a result, the divergence $\nabla \cdot \mathbf{v}$ represents the local change in area of the fluid at each point, which serves to justify our earlier statement on incompressibility.

Consider next the product vector field $u\,\mathbf{v}$ obtained by multiplying a vector field $\mathbf{v}$ by a scalar field u. An elementary computation proves that its divergence is

$$\nabla \cdot (u\,\mathbf{v}) = u\,\nabla \cdot \mathbf{v} + \nabla u \cdot \mathbf{v}. \tag{6.81}$$

Replacing $\mathbf{v}$ by $u\,\mathbf{v}$ in the divergence formula (6.80), we deduce what is usually referred to as *Green's formula*

$$\iint_\Omega \left(u\,\nabla \cdot \mathbf{v} + \nabla u \cdot \mathbf{v} \right) dx\, dy = \oint_{\partial\Omega} u\,(\mathbf{v} \cdot \mathbf{n})\, ds, \tag{6.82}$$

which is valid for arbitrary bounded domains Ω, and arbitrary C^1 scalar and vector fields defined thereon. Rearranging the terms produces

$$\iint_\Omega \nabla u \cdot \mathbf{v}\; dx\, dy = \oint_{\partial\Omega} u\,(\mathbf{v} \cdot \mathbf{n})\, ds - \iint_\Omega u\,\nabla \cdot \mathbf{v}\, dx\, dy. \tag{6.83}$$

We will view this identity as an *integration by parts* formula for double integrals. Indeed, comparing with the one-dimensional integration by parts formula

$$\int_a^b u'(x)\, v(x)\, dx = u(x)\, v(x) \Big|_{x=a}^{b} - \int_a^b u(x)\, v'(x)\, dx, \tag{6.84}$$

we observe that the single integrals have become double integrals; the derivatives are vector derivatives (gradient and divergence), while the boundary contributions at the endpoints of the interval are replaced by a line integral around the entire boundary of the two-dimensional domain.

A useful special case of (6.82) is that in which $\mathbf{v} = \nabla v$ is the gradient of a scalar field v. Then, in view of (6.74), Green's formula (6.82) becomes

$$\iint_\Omega \big(u\, \Delta v + \nabla u \cdot \nabla v \big)\, dx\, dy = \oint_{\partial\Omega} u\, \frac{\partial v}{\partial \mathbf{n}}\, ds, \tag{6.85}$$

where $\partial v/\partial \mathbf{n} = \nabla v \cdot \mathbf{n}$ is the *normal derivative* of the scalar field v on the boundary of the domain. In particular, setting $v = u$, we deduce

$$\iint_\Omega \big(u\, \Delta u + \| \nabla u \|^2 \big)\, dx\, dy = \oint_{\partial\Omega} u\, \frac{\partial u}{\partial \mathbf{n}}\, ds. \tag{6.86}$$

As an application, we establish a basic uniqueness theorem for solutions to the boundary value problems for the Poisson equation:

Theorem 6.15. *Suppose $\widetilde{u}$ and u both satisfy the same inhomogeneous Dirichlet or mixed boundary value problem for the Poisson equation on a connected, bounded domain Ω. Then $\widetilde{u} = u$. On the other hand, if $\widetilde{u}$ and u satisfy the same Neumann boundary value problem, then $\widetilde{u} = u + c$ for some constant c.*

Proof: Since, by assumption, $-\Delta\widetilde{u} = f = -\Delta u$, the difference $v = \widetilde{u} - u$ satisfies the Laplace equation $\Delta v = 0$ in Ω, and satisfies the homogeneous boundary conditions. Therefore, applying (6.86) to v, we find

$$\iint_\Omega \| \nabla v \|^2\, dx\, dy = \oint_{\partial\Omega} v\, \frac{\partial v}{\partial \mathbf{n}}\, ds = 0,$$

since, at every point on the boundary, either $v = 0$ or $\partial v/\partial \mathbf{n} = 0$. Since the integrand is continuous and everywhere nonnegative, we immediately conclude that $\| \nabla v \|^2 = 0$, and hence $\nabla v = \mathbf{0}$ throughout Ω. On a connected domain, the only functions annihilated by the gradient operator are the constants:

Lemma 6.16. *If $v(x, y)$ is a C^1 function defined on a connected domain $\Omega \subset \mathbb{R}^2$, then $\nabla v \equiv 0$ if and only if $v(x, y) \equiv c$ is a constant.*

Proof: Let $\mathbf{a}, \mathbf{b}$ be any two points in Ω. Then, by connectivity, we can find a curve C connecting them. The Fundamental Theorem for line integrals, [**8**, **108**], states that

$$\int_C \nabla v \cdot d\mathbf{x} = v(\mathbf{b}) - v(\mathbf{a}).$$

Thus, if $\nabla v \equiv 0$, then $v(\mathbf{b}) = v(\mathbf{a})$ for all $\mathbf{a}, \mathbf{b} \in \Omega$, which implies that v must be constant. *Q.E.D.*

Returning to our proof, we conclude that $\widetilde{u} = u + v = u + c$, which proves the result in the Neumann case. In the Dirichlet or mixed problems, there is at least one point on the boundary where $v = 0$, and hence the only possible constant is $v = c = 0$, proving that $\widetilde{u} = u$. *Q.E.D.*

Thus, the Dirichlet and mixed boundary value problems admit at most one solution, while the Neumann boundary value problem has either no solutions or infinitely many solutions. Proof of existence of solutions is more challenging, and will be left to a more advanced text, e.g., [**35**, **44**, **61**, **70**].

If we subtract from formula (6.85) the formula

$$\iint_\Omega \big(v\,\Delta u + \nabla u \cdot \nabla v \big)\, dx\, dy = \oint_{\partial\Omega} v\, \frac{\partial u}{\partial \mathbf{n}}\, ds, \tag{6.87}$$

obtained by interchanging u and v, we obtain the identity

$$\iint_\Omega \big(u\,\Delta v - v\,\Delta u \big)\, dx\, dy = \oint_{\partial\Omega} \left(u\, \frac{\partial v}{\partial \mathbf{n}} - v\, \frac{\partial u}{\partial \mathbf{n}} \right) ds, \tag{6.88}$$

which will play a major role in our analysis of the Poisson equation. Setting $v = 1$ in (6.87) yields

$$\iint_\Omega \Delta u\, dx\, dy = \oint_{\partial\Omega} \frac{\partial u}{\partial \mathbf{n}}\, ds. \tag{6.89}$$

Suppose u solves the Neumann boundary value problem

$$-\Delta u = f, \quad \text{in} \quad \Omega \qquad\qquad \frac{\partial u}{\partial \mathbf{n}} = h \quad \text{on} \quad \partial\Omega.$$

Then (6.89) requires that

$$\iint_\Omega f\, dx\, dy \;+\; \oint_{\partial\Omega} h\, ds = 0, \tag{6.90}$$

which thus forms a necessary condition for the existence of a solution u to the inhomogeneous Neumann boundary value problem. Physically, if u represents the equilibrium temperature of a plate, then the integrals in (6.89) measure the net gain or loss in heat energy due to, respectively, the external heat source and the heat flux through the boundary. Equation (6.90) is telling us that, for the plate to remain in thermal equilibrium, there can be no net change in its total heat energy.

The Two–Dimensional Delta Function

Now let us return to the business at hand — solving the Poisson equation on a bounded domain $\Omega \subset \mathbb{R}^2$. We will subject the solution to either homogeneous Dirichlet boundary conditions or homogeneous mixed boundary conditions. (As we just noted, the Neumann boundary value problem does not admit a unique solution, and hence does not possess a Green's function.) The Green's function for the boundary value problem arises when the forcing function is a unit impulse concentrated at a single point in the domain.

Thus, our first task is to establish the proper form for a unit impulse in our two-dimensional context. The *delta function* concentrated at a point $\boldsymbol{\xi} = (\xi, \eta) \in \mathbb{R}^2$ is denoted by

$$\delta_{(\xi,\eta)}(x, y) = \delta_{\boldsymbol{\xi}}(\mathbf{x}) = \delta(\mathbf{x} - \boldsymbol{\xi}) = \delta(x - \xi, y - \eta), \tag{6.91}$$

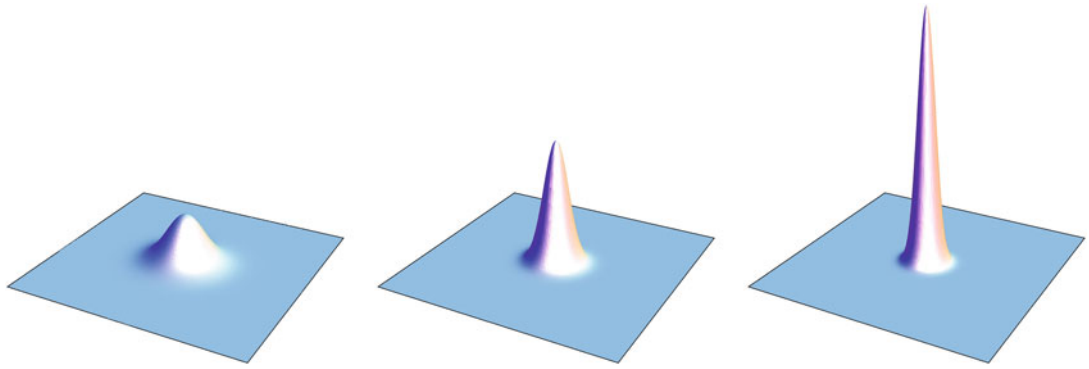


Figure 6.11. Gaussian functions converging to the delta function.

and is designed so that

$$\delta_{\boldsymbol{\xi}}(\mathbf{x}) = 0, \qquad \mathbf{x} \neq \boldsymbol{\xi}, \qquad \iint_{\Omega} \delta_{(\xi,\eta)}(x,y)\, dx\, dy = 1, \qquad \boldsymbol{\xi} \in \Omega. \tag{6.92}$$

In particular, $\delta(x,y) = \delta_{\mathbf{0}}(x,y)$ represents the delta function at the origin. As in the one-dimensional version, there is no ordinary function that satisfies both criteria; rather, $\delta(x,y)$ is to be viewed as the limit of a sequence of more and more highly concentrated functions $g_n(x,y)$, with

$$\lim_{n\to\infty} g_n(x,y) = 0, \quad \text{for} \quad (x,y) \neq (0,0), \qquad \text{while} \qquad \iint_{\mathbb{R}^2} g_n(x,y)\, dx\, dy = 1.$$

A good example of a suitable sequence is provided by the *radial Gaussian functions*

$$g_n(x,y) = \frac{n}{\pi}\, e^{-n\,(x^2+y^2)}\,. \tag{6.93}$$

As plotted in Figure 6.11, as $n \to \infty$, the Gaussian profiles become more and more concentrated near the origin, while maintaining a unit volume underneath their graphs. The fact that their integral over $\mathbb{R}^2$ equals 1 is a consequence of (2.99).

Alternatively, one can assign the delta function a dual interpretation as the linear functional

$$L_{(\xi,\eta)}[u] = L_{\boldsymbol{\xi}}[u] = u(\boldsymbol{\xi}) = u(\xi,\eta), \tag{6.94}$$

which assigns to each continuous function $u \in \mathrm{C}^0(\overline{\Omega})$ its value at the point $\boldsymbol{\xi} = (\xi,\eta) \in \Omega$. Then, using the L^2 inner product

$$\langle\, u\,,\, v\,\rangle = \iint_{\Omega} u(x,y)\, v(x,y)\, dx\, dy \tag{6.95}$$

between scalar fields $u, v \in \mathrm{C}^0(\overline{\Omega})$, we formally identify the linear functional $L_{(\xi,\eta)}$ with the delta "function" by the integral formula

$$\langle\, \delta_{(\xi,\eta)}\,,\, u\,\rangle = \iint_{\Omega} \delta_{(\xi,\eta)}(x,y)\, u(x,y)\, dx\, dy = \begin{cases} u(\xi,\eta), & (\xi,\eta) \in \Omega, \\ 0, & (\xi,\eta) \in \mathbb{R}^2 \setminus \overline{\Omega}, \end{cases} \tag{6.96}$$

for any $u \in \mathrm{C}^0(\overline{\Omega})$. As in the one-dimensional version, we will avoid defining the integral when the delta function is concentrated at a boundary point of the domain.

Since double integrals can be evaluated as repeated one-dimensional integrals, we can conveniently view

$$\delta_{(\xi,\eta)}(x,y) = \delta_\xi(x)\,\delta_\eta(y) = \delta(x-\xi)\,\delta(y-\eta) \tag{6.97}$$

as the product† of a pair of one-dimensional delta functions. Indeed, if the impulse point

$$(\xi,\eta) \in R = \big\{\, a < x < b,\ c < y < d \,\big\} \subset \Omega$$

is contained in a rectangle that lies within the domain, then

$$\begin{aligned}\iint_\Omega \delta_{(\xi,\eta)}(x,y)\,u(x,y)\,dx\,dy &= \iint_R \delta_{(\xi,\eta)}(x,y)\,u(x,y)\,dx\,dy\\ &= \int_a^b \left(\int_c^d \delta(x-\xi)\,\delta(y-\eta)\,u(x,y)\,dy \right) dx = \int_a^b \delta(x-\xi)\,u(x,\eta)\,dx = u(\xi,\eta).\end{aligned}$$

The Green's Function

As in the one-dimensional context, the Green's function is defined as the solution to the inhomogeneous differential equation when subject to a concentrated unit delta impulse at a prescribed point $\boldsymbol{\xi} = (\xi,\eta) \in \Omega$ inside the domain. In the current situation, the Poisson equation takes the form

$$-\Delta u = \delta_{\boldsymbol{\xi}}, \qquad \text{or, explicitly,} \qquad -\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} = \delta(x-\xi)\,\delta(y-\eta). \tag{6.98}$$

The function $u(x,y)$ is also subject to some homogeneous boundary conditions, e.g., the Dirichlet conditions $u = 0$ on $\partial\Omega$. The resulting solution is called the *Green's function* for the boundary value problem, and written

$$G_{\boldsymbol{\xi}}(\mathbf{x}) = G(\mathbf{x};\boldsymbol{\xi}) = G(x,y;\xi,\eta). \tag{6.99}$$

Once we know the Green's function, the solution to the general Poisson boundary value problem

$$-\Delta u = f \qquad \text{in} \qquad \Omega, \qquad\qquad u = 0 \qquad \text{on} \qquad \partial\Omega \tag{6.100}$$

is reconstructed as follows. We regard the forcing function

$$f(x,y) = \iint_\Omega \delta(x-\xi)\,\delta(y-\eta) f(\xi,\eta)\,d\xi\,d\eta$$

as a superposition of delta impulses, whose strength equals the value of f at the impulse point. Linearity implies that the solution to the boundary value problem is the corresponding superposition of Green's function responses to each of the constituent impulses. The net result is the fundamental *superposition formula*

$$u(x,y) = \iint_\Omega G(x,y;\xi,\eta)\,f(\xi,\eta)\,d\xi\,d\eta \tag{6.101}$$

† This is an exception to our earlier injunction not to multiply delta functions. Multiplication is allowed when they depend on *different* variables.

for the solution to the boundary value problem. Indeed,

$$\begin{aligned} -\Delta u(x,y) &= \iint_\Omega -\Delta G(x,y;\xi,\eta)\, f(\xi,\eta)\, d\xi\, d\eta \\ &= \iint_\Omega \delta(x-\xi, y-\eta)\, f(\xi,\eta)\, d\xi\, d\eta = f(x,y), \end{aligned}$$

while the fact that $G(x,y;\xi,\eta) = 0$ for all $(x,y) \in \partial\Omega$ implies that $u(x,y) = 0$ on the boundary.

The Green's function inevitably turns out to be symmetric under interchange of its arguments:

$$G(\xi,\eta;x,y) = G(x,y;\xi,\eta). \tag{6.102}$$

As in the one-dimensional case, symmetry is a consequence of the self-adjointness of the boundary value problem, and will be explained in full in Chapter 9. Symmetry has the following intriguing physical interpretation: Let $\mathbf{x}, \boldsymbol{\xi} \in \Omega$ be any two points in the domain. We apply a concentrated unit force to the membrane at the first point and measure its deflection at the second; the result is exactly the same as if we applied the impulse at the second point and measured the deflection at the first. (Deflections at other points in the domain will typically have no obvious relation with one another.) Similarly, in electrostatics, the solution $u(x,y)$ is interpreted as the electrostatic potential for a system of charges in equilibrium. A delta function corresponds to a point charge, e.g., an electron. The symmetry property says that the electrostatic potential at $\mathbf{x}$ due to a point charge placed at position $\boldsymbol{\xi}$ is exactly the same as the potential at $\boldsymbol{\xi}$ due to a point charge at $\mathbf{x}$. The reader may wish to meditate on the physical plausibility of these striking facts.

Unfortunately, most Green's functions cannot be written down in closed form. One important exception occurs when the domain is the entire plane: $\Omega = \mathbb{R}^2$. The solution to the Poisson equation (6.98) is the *free-space Green's function* $G_0(x,y;\xi,\eta) = G_0(\mathbf{x};\boldsymbol{\xi})$, which measures the effect of a unit impulse, concentrated at $\boldsymbol{\xi}$, throughout two-dimensional space, e.g., the gravitational potential due to a point mass or the electrostatic potential due to a point charge. To motivate the construction, let us appeal to physical intuition. First, since the concentrated impulse is zero when $\mathbf{x} \neq \boldsymbol{\xi}$, the function must solve the homogeneous Laplace equation

$$-\Delta G_0 = 0 \qquad \text{for all} \qquad \mathbf{x} \neq \boldsymbol{\xi}. \tag{6.103}$$

Second, since the Poisson equation is modeling a homogeneous, uniform medium, in the absence of boundary conditions the effect of a unit impulse should depend only on the distance from its source. Therefore, we expect G_0 to be a function of the radial variable alone:

$$G_0(x,y;\xi,\eta) = v(r), \qquad \text{where} \qquad r = \| \mathbf{x} - \boldsymbol{\xi} \| = \sqrt{(x-\xi)^2 + (y-\eta)^2}\,.$$

According to (4.113), the only radially symmetric solutions to the Laplace equation are

$$v(r) = a + b \log r, \tag{6.104}$$

where a and b are constants. The constant term a has zero derivative, and so cannot contribute to the delta function singularity. Therefore, we expect the required solution to be a multiple of the logarithmic term. To determine the multiple, consider a closed disk of radius $\varepsilon > 0$ centered at $\boldsymbol{\xi}$,

$$D_\varepsilon = \big\{\, 0 \le r \le \varepsilon \,\big\} = \big\{\, \| \mathbf{x} - \boldsymbol{\xi} \| \le \varepsilon \,\big\},$$

with circular boundary

$$C_\varepsilon = \partial D_\varepsilon = \{\, r = \| \mathbf{x} - \boldsymbol{\xi} \| = \varepsilon \,\} = \{\, (\, \xi + \varepsilon \cos\theta, \eta + \varepsilon \sin\theta \,) \;\mid\; -\pi \le \theta \le \pi \,\}.$$

Then, by (6.74) and the divergence form (6.80) of Green's Theorem,

$$\begin{aligned} 1 &= \iint_{D_\varepsilon} \delta(x,y)\, dx\, dy = -b \iint_{D_\varepsilon} \Delta(\log r)\, dx\, dy = -b \iint_{D_\varepsilon} \nabla \cdot \nabla(\log r)\, dx\, dy \\ &= -b \oint_{C_\varepsilon} \frac{\partial(\log r)}{\partial \mathbf{n}}\, ds = -b \oint_{C_\varepsilon} \frac{\partial(\log r)}{\partial r}\, ds = -b \oint_{C_\varepsilon} \frac{1}{r}\, ds = -b \int_{-\pi}^{\pi} d\theta = -2\pi b, \end{aligned} \tag{6.105}$$

and hence $b = -1/(2\pi)$. We conclude that the free-space Green's function should have the logarithmic form

$$G_0(x,y;\xi,\eta) = -\frac{1}{2\pi} \log r = -\frac{1}{2\pi} \log \| \mathbf{x} - \boldsymbol{\xi} \| = -\frac{1}{4\pi} \log\big[\, (x-\xi)^2 + (y-\eta)^2 \,\big]. \tag{6.106}$$

A fully rigorous, albeit more difficult, justification of (6.106) comes from the following important result, known as *Green's representation formula*.

Theorem 6.17. *Let $\Omega \subset \mathbb{R}^2$ be a bounded domain, with piecewise C^1 boundary $\partial\Omega$. Suppose $u \in \mathrm{C}^2(\Omega) \cap \mathrm{C}^1(\overline{\Omega})$. Then, for any $(x,y) \in \Omega$,*

$$\begin{aligned} u(x,y) = &- \iint_\Omega G_0(x,y;\xi,\eta)\, \Delta u(\xi,\eta)\, d\xi\, d\eta \\ &+ \oint_{\partial\Omega} \left(G_0(x,y;\xi,\eta)\, \frac{\partial u}{\partial \mathbf{n}}\,(\xi,\eta) - \frac{\partial G_0}{\partial \mathbf{n}}\,(x,y;\xi,\eta)\, u(\xi,\eta) \right) ds, \end{aligned} \tag{6.107}$$

where the Laplacian and the normal derivatives on the boundary are all taken with respect to the integration variables $\boldsymbol{\xi} = (\xi,\eta)$.

In particular, if both u and $\partial u/\partial \mathbf{n}$ vanish on $\partial\Omega$, then (6.107) reduces to

$$u(x,y) = -\iint_{\mathbb{R}^2} G_0(x,y;\xi,\eta)\, \Delta u(\xi,\eta)\, d\xi\, d\eta.$$

Invoking the definition of the delta function on the left-hand side and formally applying the Green identity (6.88) to the right-hand side produces

$$\iint_{\mathbb{R}^2} \delta(x-\xi)\, \delta(y-\eta)\, u(\xi,\eta)\, d\xi\, d\eta = \iint_{\mathbb{R}^2} -\Delta G_0(x,y;\xi,\eta)\, u(\xi,\eta)\, d\xi\, d\eta. \tag{6.108}$$

It is in this dual sense that we justify the desired formula

$$-\Delta G_0(\mathbf{x};\boldsymbol{\xi}) = \frac{1}{2\pi}\, \Delta\big(\log \| \mathbf{x} - \boldsymbol{\xi} \|\big) = \delta(\mathbf{x} - \boldsymbol{\xi}). \tag{6.109}$$

Proof of Theorem 6.17: We first note that, even though $G_0(\mathbf{x},\boldsymbol{\xi})$ has a logarithmic singularity at $\mathbf{x} = \boldsymbol{\xi}$, the double integral in (6.107) is finite. Indeed, after introducing polar coordinates $\xi = x + r\cos\theta$, $\eta = y + r\sin\theta$, and recalling $d\xi\, d\eta = r\, dr\, d\theta$, we see that it equals

$$\frac{1}{2\pi} \iint (r \log r)\, \Delta u\, dr\, d\theta.$$

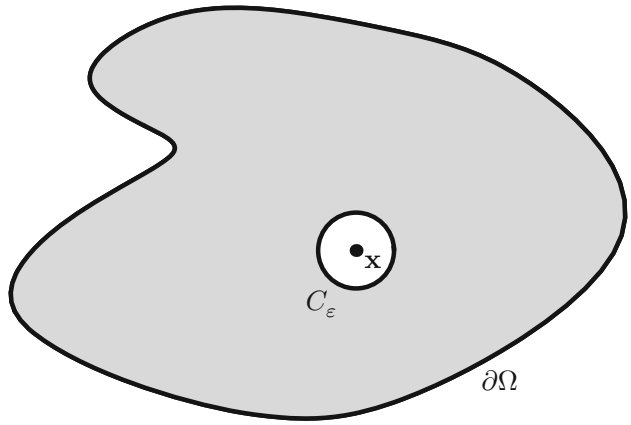


Figure 6.12. Domain $\Omega_\varepsilon = \Omega \setminus D_\varepsilon(\mathbf{x})$.

The product $r \log r$ is everywhere continuous — even at $r = 0$ — and so, provided Δu is well behaved, e.g., continuous, the integral is finite. There is, of course, no problem with the line integral in (6.107), since the contour does not go through the singularity.

Let us now avoid dealing directly with the singularity by working on a subdomain

$$\Omega_\varepsilon = \Omega \setminus D_\varepsilon(\mathbf{x}) = \{\, \boldsymbol{\xi} \in \Omega \mid \ \| \mathbf{x} - \boldsymbol{\xi} \| > \varepsilon \,\}$$

obtained by cutting out a small disk

$$D_\varepsilon(\mathbf{x}) = \{\, \boldsymbol{\xi} \mid \ \| \mathbf{x} - \boldsymbol{\xi} \| \leq \varepsilon \,\}$$

of radius $\varepsilon > 0$ centered at $\mathbf{x}$. We choose ε sufficiently small in order that $D_\varepsilon(\mathbf{x}) \subset \Omega$, and hence

$$\partial\Omega_\varepsilon = \partial\Omega \cup C_\varepsilon, \qquad \text{where} \qquad C_\varepsilon = \big\{\, \| \mathbf{x} - \boldsymbol{\xi} \| = \varepsilon \,\big\}$$

is the circular boundary of the disk. The subdomain Ω_ε is represented by the shaded region in Figure 6.12. Since the double integral is well defined, we can approximate it by integrating over Ω_ε:

$$\iint_\Omega G_0(x, y; \xi, \eta)\, \Delta u(\xi, \eta)\, d\xi\, d\eta = \lim_{\varepsilon \to 0} \iint_{\Omega_\varepsilon} G_0(x, y; \xi, \eta)\, \Delta u(\xi, \eta)\, d\xi\, d\eta. \tag{6.110}$$

Since G_0 has no singularities in Ω_ε, we are able to apply the Green formula (6.85) and then (6.103) to evaluate

$$\begin{aligned} &\iint_{\Omega_\varepsilon} G_0(x, y; \xi, \eta)\, \Delta u(\xi, \eta)\, d\xi\, d\eta \\ &\qquad = \oint_{\partial\Omega} \left(G_0(x, y; \xi, \eta)\, \frac{\partial u}{\partial \mathbf{n}}\,(\xi, \eta) - \frac{\partial G_0}{\partial \mathbf{n}}\,(x, y; \xi, \eta)\, u(\xi, \eta) \right) ds \\ &\qquad\qquad - \oint_{C_\varepsilon} \left(G_0(x, y; \xi, \eta)\, \frac{\partial u}{\partial \mathbf{n}}\,(\xi, \eta) - \frac{\partial G_0}{\partial \mathbf{n}}\,(x, y; \xi, \eta)\, u(\xi, \eta) \right) ds, \end{aligned} \tag{6.111}$$

where the line integral around C_ε is taken in the usual counterclockwise direction — the opposite orientation to that induced by its status as part of the boundary of Ω_ε. Now, on

the circle C_ε,

$$G_0(x,y;\xi,\eta) = -\frac{\log r}{2\pi}\,\Big|_{r=\varepsilon} = -\frac{\log \varepsilon}{2\pi}\,, \tag{6.112}$$

while, in view of Exercise 6.3.1,

$$\frac{\partial G_0}{\partial \mathbf{n}}\,(x,y;\xi,\eta) = -\frac{1}{2\pi}\,\frac{\partial(\log r)}{\partial r}\,\Big|_{r=\varepsilon} = -\frac{1}{2\pi\varepsilon}\,. \tag{6.113}$$

Therefore,

$$\oint_{C_\varepsilon} \frac{\partial G_0}{\partial \mathbf{n}}\,(x,y;\xi,\eta)\,u(\xi,\eta)\,ds = -\frac{1}{2\pi\varepsilon}\,\oint_{C_\varepsilon} u(\xi,\eta)\,ds,$$

which we recognize as minus the *average* of u on the circle of radius ε. As $\varepsilon \to 0$, the circles shrink down to their common center, and so, by continuity, the averages tend to the value $u(x,y)$ at the center; thus,

$$\lim_{\varepsilon\to 0} \oint_{C_\varepsilon} \frac{\partial G_0}{\partial \mathbf{n}}\,(x,y;\xi,\eta)\,u(\xi,\eta)\,ds = -u(x,y). \tag{6.114}$$

On the other hand, using (6.112), and then (6.89) on the disk D_ε, we have

$$\begin{aligned}\oint_{C_\varepsilon} G_0(x,y;\xi,\eta)\,\frac{\partial u}{\partial \mathbf{n}}\,(\xi,\eta)\,ds &= -\frac{\log\varepsilon}{2\pi}\,\oint_{C_\varepsilon}\frac{\partial u}{\partial \mathbf{n}}\,(\xi,\eta)\,ds\\ &= -\frac{\log\varepsilon}{2\pi}\,\iint_{D_\varepsilon}\Delta u(\xi,\eta)\,d\xi\,d\eta = -\,(\varepsilon^2\log\varepsilon)\,\overline{\Delta u}_\varepsilon,\end{aligned}$$

where

$$\overline{\Delta u}_\varepsilon = \frac{1}{2\pi\varepsilon^2}\,\iint_{D_\varepsilon}\Delta u(\xi,\eta)\,d\xi\,d\eta$$

is the *average* of Δu over the disk D_ε. As above, as $\varepsilon \to 0$, the averages over the disks converge to the value at their common center, $\overline{\Delta u}_\varepsilon \to \Delta u(x,y)$, and hence

$$\lim_{\varepsilon\to 0} \oint_{C_\varepsilon} G_0(x,y;\xi,\eta)\,\frac{\partial u}{\partial \mathbf{n}}\,(\xi,\eta)\,ds = \lim_{\varepsilon\to 0}\,(-\varepsilon^2\log\varepsilon)\,\overline{\Delta u}_\varepsilon = 0. \tag{6.115}$$

In view of (6.110, 114, 115), the $\varepsilon \to 0$ limit of (6.111) is exactly the Green representation formula (6.107). *Q.E.D.*

As noted above, the free space Green's function (6.106) represents the gravitational potential in empty two-dimensional space due to a unit point mass, or, equivalently, the two-dimensional electrostatic potential due to a unit point charge sitting at position $\boldsymbol{\xi}$. The corresponding gravitational or electrostatic force field is obtained by taking its gradient:

$$\mathbf{F} = \nabla G_0 = -\frac{\mathbf{x}-\boldsymbol{\xi}}{2\pi\,\|\,\mathbf{x}-\boldsymbol{\xi}\,\|^2}\,.$$

Its magnitude

$$\|\,\mathbf{F}\,\| = \frac{1}{2\pi\,\|\,\mathbf{x}-\boldsymbol{\xi}\,\|}$$

is inversely proportional to the distance from the mass or charge, which is the two-dimensional form of Newton's and Coulomb's three-dimensional inverse square laws.

The gravitational potential due to a two-dimensional mass, e.g., a flat plate, in the shape of a domain $\Omega \subset \mathbb{R}^2$ is obtained by superimposing delta function sources with strengths equal to the density of the material at each point. The result is the potential function

$$u(x,y) = -\frac{1}{4\,\pi} \iint_{\Omega} \rho(\xi,\eta)\, \log\big[\,(x-\xi)^2 + (y-\eta)^2\,\big]\, d\xi\, d\eta, \tag{6.116}$$

in which $\rho(\xi,\eta)$ denotes the density at position $(\xi,\eta) \in \Omega$.

Example 6.18. The gravitational potential due to a circular disk $D = \{\, x^2 + y^2 \le 1 \,\}$ of unit radius and unit density $\rho \equiv 1$ is

$$u(x,y) = -\frac{1}{4\,\pi} \iint_{D} \log\big[\,(x-\xi)^2 + (y-\eta)^2\,\big]\, d\xi\, d\eta. \tag{6.117}$$

A direct evaluation of this double integral is not so easy. However, we can write down the potential in closed form by recalling that it solves the Poisson equation

$$-\,\Delta u = \begin{cases} 1, & \|\,\mathbf{x}\,\| < 1, \\ 0, & \|\,\mathbf{x}\,\| > 1. \end{cases} \tag{6.118}$$

Moreover, u is clearly radially symmetric, and hence a function of r alone. Thus, in the polar coordinate expression (4.105) for the Laplacian, the θ derivative terms vanish, and so (6.118) reduces to

$$\frac{d^2u}{dr^2} + \frac{1}{r}\,\frac{du}{dr} = \begin{cases} -1, & r < 1, \\ 0, & r > 1, \end{cases}$$

which is effectively a first-order linear ordinary differential equation for du/dr. Solving separately on the two subintervals produces

$$u(r) = \begin{cases} a + b\log r - \frac{1}{4}\,r^2, & r < 1, \\ c + d\log r, & r > 1, \end{cases}$$

where a, b, c, d are constants. Continuity of $u(r)$ and $u'(r)$ at $r = 1$ implies $c = a - \frac{1}{4}$, $d = b - \frac{1}{2}$. Moreover, the potential for a non-concentrated mass cannot have a singularity at the origin, and so $b = 0$. Direct evaluation of (6.117) at $x = y = 0$, using polar coordinates, proves that $a = \frac{1}{4}$. We conclude that the gravitational potential (6.117) due to a uniform disk of unit radius, and hence total mass (area) π, is, explicitly,

$$u(x,y) = \begin{cases} \frac{1}{4}\,(1 - r^2) = \frac{1}{4}\,(1 - x^2 - y^2), & x^2 + y^2 \le 1, \\ -\frac{1}{2}\log r = -\frac{1}{4}\log(x^2 + y^2), & x^2 + y^2 \ge 1. \end{cases} \tag{6.119}$$

Observe that, outside the disk, the potential is exactly the same as the logarithmic potential due to a point mass of magnitude π located at the origin. Consequently, the gravitational force field outside a uniform disk is the same as if all its mass were concentrated at the origin.

With the free-space logarithmic potential in hand, let us return to the question of finding the Green's function for a boundary value problem on a bounded domain $\Omega \subset \mathbb{R}^2$. Since the logarithmic potential (6.106) is a particular solution to the Poisson equation (6.98), the general solution, according to Theorem 1.6, is given by $u = G_0 + z$, where z is an arbitrary solution to the homogeneous equation $\Delta z = 0$, i.e., an arbitrary harmonic function. Thus, constructing the Green's function has been reduced to the problem of finding the harmonic function z such that $G = G_0 + z$ satisfies the desired homogeneous boundary conditions. Let us explicitly formulate this result for the (inhomogeneous) Dirichlet problem.

Theorem 6.19. *The Green's function for the Dirichlet boundary value problem for the Poisson equation on a bounded domain $\Omega \subset \mathbb{R}^2$ has the form*

$$G(x,y;\xi,\eta) = G_0(x,y;\xi,\eta) + z(x,y;\xi,\eta), \tag{6.120}$$

where the first term is the logarithmic potential (6.106)*, while, for each $(\xi,\eta) \in \Omega$, the second term is the harmonic function that solves the boundary value problem*

$$\begin{aligned} \Delta z &= 0 \qquad \text{on} \qquad \Omega, \\ z(x,y;\xi,\eta) &= \frac{1}{4\,\pi} \log\big[\,(x-\xi)^2 + (y-\eta)^2\,\big] \qquad \text{for} \qquad (x,y) \in \partial\Omega. \end{aligned} \tag{6.121}$$

If $u(x,y)$ is a solution to the inhomogeneous Dirichlet problem

$$-\Delta u = f, \qquad \mathbf{x} \in \Omega, \qquad\qquad u = h, \qquad \mathbf{x} \in \partial\Omega, \tag{6.122}$$

then

$$u(x,y) = \iint_\Omega G(x,y;\xi,\eta)\, f(\xi,\eta)\, d\xi\, d\eta - \oint_{\partial\Omega} \frac{\partial G}{\partial \mathbf{n}}\,(x,y;\xi,\eta)\, h(\xi,\eta)\, ds, \tag{6.123}$$

where the normal derivative of G is taken with respect to $(\xi,\eta) \in \partial\Omega$.

Proof: To show that (6.120) is the Green's function, we note that

$$-\Delta G = -\Delta G_0 - \Delta z = \delta_{(\xi,\eta)} \qquad \text{in} \qquad \Omega, \tag{6.124}$$

while

$$G(x,y;\xi,\eta) = G_0(x,y;\xi,\eta) + z(x,y;\xi,\eta) = 0 \qquad \text{on} \qquad \partial\Omega. \tag{6.125}$$

Next, to establish the solution formula (6.123), since both z and u are C^2, we can use (6.88) (with $v = z$, keeping in mind that $\Delta z = 0$) to establish

$$\begin{aligned} 0 = -\iint_\Omega z(x,y;\xi,\eta)\, \Delta u(\xi,\eta)\, d\xi\, d\eta \\ + \oint_{\partial\Omega} \left(z(x,y;\xi,\eta)\, \frac{\partial u}{\partial \mathbf{n}}\,(\xi,\eta) - \frac{\partial z}{\partial \mathbf{n}}\,(x,y;\xi,\eta)\, u(\xi,\eta) \right) ds. \end{aligned}$$

Adding this to Green's representation formula (6.107), and using (6.125), we deduce that

$$u(x,y) = -\iint_\Omega G(x,y;\xi,\eta)\, \Delta u(\xi,\eta)\, d\xi\, d\eta - \oint_{\partial\Omega} \frac{\partial G(x,y;\xi,\eta)}{\partial \mathbf{n}}\, u(\xi,\eta)\, ds,$$

which, given (6.122), produces (6.123). *Q.E.D.*

The one subtle issue left unresolved is the existence of the solution. Read properly, Theorem 6.19 states that *if* a classical solution exists, *then* it is necessarily given by the Green's function formula (6.123). Proving existence of the solution — and also the existence of the Green's function, or equivalently, the solution z to (6.121) — requires further in-depth analysis, lying beyond the scope of this text. In particular, to guarantee existence, the underlying domain must have a reasonably nice boundary, e.g., a piecewise smooth curve without sharp cusps. Interestingly, lack of regularity at sharp cusps in the boundary underlies the electromagnetic phenomenon known as St. Elmo's fire, cf. [**121**]. Extensions to irregular domains, e.g., those with fractal boundaries, is an active area of contemporary research. Moreover, unlike one-dimensional boundary value problems, mere continuity of

the forcing function f is not quite sufficient to ensure the existence of a classical solution to the Poisson boundary value problem; differentiability does suffice, although this assumption can be weakened. We refer to [**61**, **70**], for a development of the Perron method based on approximating the solution by a sequence of *subsolutions*, which, by definition, solve the differential inequality $-\Delta u \le f$. An alternative proof, using the direct method of the calculus of variations, can be found in [**35**]. The latter proof relies on the characterization of the solution by a minimization principle, which we discuss in some detail in Chapter 9.

Exercises

♢ 6.3.1. Let C_R be a circle of radius R centered at the origin and $\mathbf{n}$ its unit outward normal. Let $f(r,\theta)$ be a function expressed in polar coordinates. Prove that $\partial f/\partial \mathbf{n} = \partial f/\partial r$ on C_R.

6.3.2. Let $f(x) > 0$ be a continuous, positive function on the interval $a \le x \le b$. Let Ω be the domain lying between the graph of $f(x)$ on the interval $[a,b]$ and the x–axis. Explain why (6.77) reduces to the usual calculus formula for the area under the graph of f.

6.3.3. Explain what happens to the conclusion of Lemma 6.16 if Ω is not a connected domain.

6.3.4. Can you find constants c_n such that the functions $g_n(x,y) = c_n[1 + n^2(x^2+y^2)]^{-1}$ converge to the two-dimensional delta function: $g_n(x,y) \to \delta(x,y)$ as $n \to \infty$?

6.3.5. Explain why the two-dimensional delta function satisfies the scaling law

$$\delta(\beta\, x, \beta\, y) = \frac{1}{\beta^2}\,\delta(x,y), \qquad \text{for} \qquad \beta > 0.$$

♢ 6.3.6. Write out a polar coordinate formula, in terms of $\delta(r - r_0)$ and $\delta(\theta - \theta_0)$, for the two-dimensional delta function $\delta(x - x_0, y - y_0) = \delta(x - x_0)\,\delta(y - y_0)$.

6.3.7. *True or false*: $\delta(\mathbf{x}) = \delta(\| \mathbf{x} \|)$.

♢ 6.3.8. Suppose that $\xi = f(x,y)$, $\eta = g(x,y)$ defines a one-to-one C^1 map from a domain $D \subset \mathbb{R}^2$ to the domain $\Omega = \{\, (\xi,\eta) = (f(x,y), g(x,y)) \,|\, (x,y) \in D \,\} \subset \mathbb{R}^2$, and has nonzero Jacobian determinant: $J(x,y) = f_x g_y - f_y g_x \neq 0$ for all $(x,y) \in D$. Suppose further that $(0,0) = (f(x_0,y_0), g(x_0,y_0)) \in \Omega$ for $(x_0,y_0) \in D$. Prove the following formula governing the effect of the map on the two-dimensional delta function:

$$\delta(f(x,y), g(x,y)) = \frac{\delta(x - x_0, y - y_0)}{|\, J(x_0,y_0)\,|}\,. \tag{6.126}$$

6.3.9. Suppose $f(x,y) = \begin{cases} 1, & 3x - 2y > 1, \\ 0, & 3x - 2y < 1. \end{cases}$ Compute its partial derivatives $\dfrac{\partial f}{\partial x}$ and $\dfrac{\partial f}{\partial y}$ in the sense of generalized functions.

6.3.10. Find a series solution to the rectangular boundary value problem (4.91–92) when the boundary data $f(x) = \delta(x - \xi)$ is a delta function at a point $0 < \xi < a$. Is your solution infinitely differentiable inside the rectangle?

6.3.11. Answer Exercise 6.3.10 when $f(x) = \delta'(x - \xi)$ is the derivative of the delta function.

6.3.12. A 1 meter square plate is subject to the Neumann boundary conditions $\partial u/\partial \mathbf{n} = 1$ on its entire boundary. What is the equilibrium temperature? Explain.

♢ 6.3.13. A *conservation law* for an equilibrium system in two dimensions is, by definition, a divergence expression

$$\frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} = 0 \tag{6.127}$$

that vanishes for all solutions.

(*a*) Given a conservation law prescribed by $\mathbf{v} = (X, Y)$ defined on a simply connected domain D, show that the line integral $\int_C \mathbf{v} \cdot \mathbf{n}\, ds = \int_C X\, dy - Y\, dx$ is path-independent, meaning that its value depends only on the endpoints of the curve C.

(*b*) Show that the Laplace equation can be written as a conservation law, and write down the corresponding path-independent line integral.

Note: Path-independent integrals are of importance in the study of cracks, dislocations, and other material singularities, [**49**].

♢ 6.3.14. In two-dimensional dynamics, a *conservation law* is an equation of the form

$$\frac{\partial T}{\partial t} + \frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} = 0, \tag{6.128}$$

in which T is the *conserved density*, while $\mathbf{v} = (X, Y)$ represents the associated *flux*.

(*a*) Prove that, on a bounded domain $\Omega \subset \mathbb{R}^2$, the rate of change of the integral $\iint_\Omega T\, dx\, dy$ of the conserved density depends only on the flux through the boundary $\partial\Omega$.

(*b*) Write the partial differential equation $u_t + u\, u_x + u\, u_y = 0$ as a conservation law. What is the integrated version?

The Method of Images

The preceding analysis exposes the underlying form of the Green's function, but we are still left with the determination of the harmonic component $z(x, y)$ required to match the logarithmic potential boundary values, cf. (6.121). We will discuss two principal analytic techniques employed to produce explicit formulas. The first is an adaptation of the method of separation of variables, which leads to infinite series expressions. We will not dwell on this approach here, although a couple of the exercises ask the reader to work through some of the details; see also the discussion leading up to (9.110). The second is the *Method of Images*, which will be developed in this section. Another approach is based on the theory of *conformal mapping*; it can be found in books on complex analysis, including [**53**, **98**]. While the first two methods are limited to a fairly small class of domains, they extend to higher-dimensional problems, as well as to certain other types of elliptic boundary value problems, whereas conformal mapping is, unfortunately, restricted to two-dimensional problems involving the Laplace and Poisson equations.

We already know that the singular part of the Green's function for the two-dimensional Poisson equation is provided by a logarithmic potential. The problem, then, is to construct the harmonic part, called $z(x, y)$ in (6.120), so that the sum has the correct homogeneous boundary values, or, equivalently, so that $z(x, y)$ has the same boundary values as the logarithmic potential. In certain cases, $z(x, y)$ can be thought of as the potential induced by one or more hypothetical electric charges (or, equivalently, gravitational point masses) that are located *outside* the domain Ω, arranged in such a manner that their combined electrostatic potential happens to coincide with the logarithmic potential on the boundary of the domain. The goal, then, is to place image charges of suitable strengths in the appropriate positions.

Here, we will only consider the case of a single image charge, located at a position $\boldsymbol{\eta} \notin \Omega$. We scale the logarithmic potential (6.106) by the charge strength, and, for added

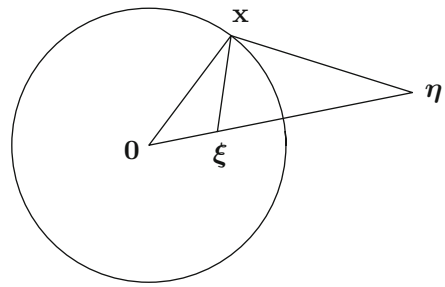


Figure 6.13. Method of Images for the unit disk.

flexibility, include an additional constant — the charge's potential baseline:

$$z(x,y) = a \log \| \mathbf{x} - \boldsymbol{\eta} \| + b, \qquad \boldsymbol{\eta} \in \mathbb{R}^2 \setminus \overline{\Omega}.$$

The function $z(x,y)$ is harmonic inside Ω, since the logarithmic potential is harmonic everywhere except at the external singularity $\boldsymbol{\eta}$. For the Dirichlet boundary value problem, then, for each point $\boldsymbol{\xi} \in \Omega$, we must find a corresponding image point $\boldsymbol{\eta} \in \mathbb{R}^2 \setminus \overline{\Omega}$ and constants $a, b \in \mathbb{R}$ such that[†]

$$\log \| \mathbf{x} - \boldsymbol{\xi} \| = a \log \| \mathbf{x} - \boldsymbol{\eta} \| + b \qquad \text{for all} \qquad \mathbf{x} \in \partial\Omega,$$

or, equivalently,

$$\| \mathbf{x} - \boldsymbol{\xi} \| = \lambda \| \mathbf{x} - \boldsymbol{\eta} \|^a \qquad \text{for all} \qquad \mathbf{x} \in \partial\Omega, \tag{6.129}$$

where $\lambda = e^b$. For each fixed $\boldsymbol{\xi}, \boldsymbol{\eta}, \lambda, a$, the equation in (6.129) will, typically, implicitly prescribe a plane curve, but it is not clear that one can always arrange that these curves all coincide with the boundary of our domain.

To make further progress, we appeal to a geometric construction based on similar triangles. Let us select $\boldsymbol{\eta} = c\,\boldsymbol{\xi}$ to be a point lying on the ray through $\boldsymbol{\xi}$. Its location is chosen so that the triangle with vertices $\mathbf{0}, \mathbf{x}, \boldsymbol{\eta}$ is similar to the triangle with vertices $\mathbf{0}, \boldsymbol{\xi}, \mathbf{x}$, noting that they have the same angle at the common vertex $\mathbf{0}$ — see Figure 6.13. Similarity requires that the triangles' corresponding sides have a common ratio, and so

$$\frac{\| \boldsymbol{\xi} \|}{\| \mathbf{x} \|} = \frac{\| \mathbf{x} \|}{\| \boldsymbol{\eta} \|} = \frac{\| \mathbf{x} - \boldsymbol{\xi} \|}{\| \mathbf{x} - \boldsymbol{\eta} \|} = \lambda. \tag{6.130}$$

The last equality implies that (6.129) holds with $a = 1$. Consequently, if we choose

$$\| \boldsymbol{\eta} \| = \frac{1}{\| \boldsymbol{\xi} \|}, \qquad \text{so that} \qquad \boldsymbol{\eta} = \frac{\boldsymbol{\xi}}{\| \boldsymbol{\xi} \|^2}, \tag{6.131}$$

then

$$\| \mathbf{x} \|^2 = \| \boldsymbol{\xi} \| \, \| \boldsymbol{\eta} \| = 1.$$

[†] To simplify the formulas, we have omitted the $1/(2\pi)$ factor, which can easily be reinstated at the end of the analysis.

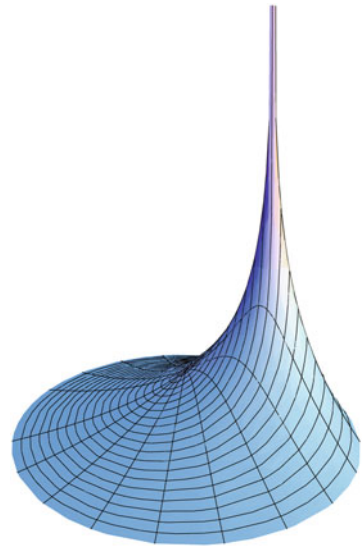


Figure 6.14. Green's function for the unit disk.

Thus $\mathbf{x}$ lies on the unit circle, and, as a result, $\lambda = \| \boldsymbol{\xi} \| = 1/\| \boldsymbol{\eta} \|$. The map taking a point $\boldsymbol{\xi}$ inside the disk to its image point $\boldsymbol{\eta}$ defined by (6.131) is known as *inversion* with respect to the unit circle.

We have now demonstrated that the potentials

$$\frac{1}{2\,\pi}\,\log \| \mathbf{x} - \boldsymbol{\xi} \| = \frac{1}{2\,\pi}\,\log\big(\| \boldsymbol{\xi} \|\, \| \mathbf{x} - \boldsymbol{\eta} \| \big) = \frac{1}{2\,\pi}\,\log \frac{\| \, \| \boldsymbol{\xi} \|^2 \, \mathbf{x} - \boldsymbol{\xi} \, \|}{\| \boldsymbol{\xi} \|}\,, \qquad \| \mathbf{x} \| = 1, \tag{6.132}$$

have the same boundary values on the unit circle. Consequently, their difference

$$G(\mathbf{x}; \boldsymbol{\xi}) = -\,\frac{1}{2\,\pi}\,\log \| \mathbf{x} - \boldsymbol{\xi} \| + \frac{1}{2\,\pi}\,\log \frac{\| \, \| \boldsymbol{\xi} \|^2 \, \mathbf{x} - \boldsymbol{\xi} \, \|}{\| \boldsymbol{\xi} \|} = \frac{1}{2\,\pi}\,\log \frac{\| \, \| \boldsymbol{\xi} \|^2 \, \mathbf{x} - \boldsymbol{\xi} \, \|}{\| \boldsymbol{\xi} \|\, \| \mathbf{x} - \boldsymbol{\xi} \|} \tag{6.133}$$

has the required properties for the Green's function for the Dirichlet problem on the unit disk. Writing this in terms of polar coordinates

$$\mathbf{x} = (r \cos\theta, r \sin\theta), \qquad \boldsymbol{\xi} = (\rho \cos\phi, \rho \sin\phi),$$

and applying the Law of Cosines to the triangles in Figure 6.13 produces the explicit formula

$$G(r, \theta; \rho, \phi) = \frac{1}{4\,\pi}\,\log\left(\frac{1 + r^2\rho^2 - 2\,r\rho\cos(\theta - \phi)}{r^2 + \rho^2 - 2\,r\rho\cos(\theta - \phi)} \right). \tag{6.134}$$

In Figure 6.14 we sketch the Green's function for the Dirichlet boundary value problem corresponding to a unit impulse being applied at a point halfway between the center and the edge of the disk. We also require its radial derivative

$$\frac{\partial G}{\partial r}\,(r, \theta; \rho, \phi) = -\,\frac{1}{2\,\pi}\,\frac{1 - r^2}{1 + r^2 - 2\,r\cos(\theta - \phi)}\,, \tag{6.135}$$

which coincides with its normal derivative on the unit circle. Thus, specializing (6.123), we arrive at a solution to the general Dirichlet boundary value problem for the Poisson equation in the unit disk.

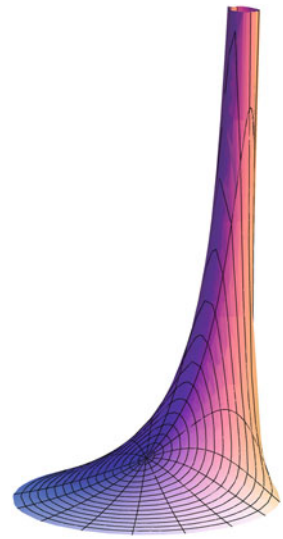


Figure 6.15. The Poisson kernel.

Theorem 6.20. *The solution to the inhomogeneous Dirichlet boundary value problem*

$$-\Delta u = f, \quad \text{for} \quad r = \| \mathbf{x} \| < 1, \qquad u = h, \quad \text{for} \quad r = 1,$$

is, when expressed in polar coordinates,

$$u(r,\theta) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \int_{0}^{1} f(\rho,\phi)\, \log\left(\frac{1 + r^2\rho^2 - 2r\rho\cos(\theta-\phi)}{r^2 + \rho^2 - 2r\rho\cos(\theta-\phi)} \right) \rho\, d\rho\, d\phi + \frac{1}{2\pi} \int_{-\pi}^{\pi} h(\phi)\, \frac{1 - r^2}{1 + r^2 - 2r\cos(\theta-\phi)}\, d\phi. \tag{6.136}$$

When $f \equiv 0$, formula (6.136) recovers the Poisson integral formula (4.126) for the solution to the Dirichlet boundary value problem for the Laplace equation. In particular, the boundary data $h(\theta) = \delta(\theta - \phi)$, corresponding to a concentrated unit heat source applied to a single point on the boundary, produces the *Poisson kernel*

$$u(r,\theta) = \frac{1 - r^2}{2\pi\big(1 + r^2 - 2r\cos(\theta-\phi)\big)}\,. \tag{6.137}$$

The reader may enjoy verifying that this function indeed solves the Laplace equation and has the correct boundary values in the limit as $r \to 1$.

Exercises

6.3.15. A circular disk of radius 1 is subject to a heat source of unit magnitude on the subdisk $r \le \frac{1}{2}$. Its boundary is kept at 0°.

(*a*) Write down an integral formula for the equilibrium temperature.

(*b*) Use radial symmetry to find an explicit formula for the equilibrium temperature.

6.3.16. A circular disk of radius 1 meter is subject to a unit concentrated heat source at its center and has completely insulated boundary. What is the equilibrium temperature?

♡ 6.3.17. (*a*) For $n > 0$, find the solution to the boundary value problem

$$-\Delta u = \frac{n}{\pi}\, e^{-n\,(x^2+y^2)}, \qquad x^2 + y^2 < 1, \qquad u(x,y) = 0, \qquad x^2 + y^2 = 1.$$

(*b*) Discuss what happens in the limit as $n \to \infty$.

♡ 6.3.18. (*a*) Use the Method of Images to construct the Green's function for a half-plane $\{\, y > 0 \,\}$ that is subject to homogeneous Dirichlet boundary conditions. *Hint*: The image point is obtained by reflection. (*b*) Use your Green's function to solve the boundary value problem

$$-\Delta u = \frac{1}{1+y}\,, \qquad y > 0, \qquad u(x,0) = 0.$$

6.3.19. Construct the Green's function for the half-disk $\Omega = \{\, x^2 + y^2 < 1,\ y > 0 \,\}$ when subject to homogeneous Dirichlet boundary conditions. *Hint*: Use three image points.

6.3.20. Prove directly that the Poisson kernel (6.137) solves the Laplace equation for all $r < 1$.

♡ 6.3.21. Provide the details for the following alternative method for solving the homogeneous Dirichlet boundary value problem for the Poisson equation on the unit square:

$$u_{xx} - u_{yy} = f(x,y), \quad u(x,0) = 0, \quad u(x,1) = 0, \quad u(0,y) = 0, \quad u(1,y) = 0, \quad 0 < x,\ y < 1.$$

(*a*) Write both $u(x,y)$ and $f(x,y)$ as Fourier sine series in y whose coefficients depend on x. (*b*) Substitute these series into the differential equation, and equate Fourier coefficients to obtain an infinite system of ordinary boundary value problems for the x-dependent Fourier coefficients of u. (*c*) Use the Green's functions for each boundary value problem to write out the solution and hence a series for the solution to the original boundary value problem. (*d*) Implement this method for the following forcing functions:
(*i*) $f(x,y) = \sin \pi y$, (*ii*) $f(x,y) = \sin \pi x\, \sin 2\pi y$, (*iii*) $f(x,y) = 1$.

♢ 6.3.22. Use the method of Exercise 6.3.21 to find a series representation for the Green's function of a unit square subject to Dirichlet boundary conditions.

6.3.23. Write out the details of how to derive (6.134) from (6.133).

6.3.24. *True or false*: If the gravitational potential at a point $\mathbf{a}$ is greater than its value at the point $\mathbf{b}$, then the magnitude of the gravitational force at $\mathbf{a}$ is greater than its value at $\mathbf{b}$.

♠ 6.3.25. (*a*) Write down integral formulas for the gravitational potential and force due to a square plate $S = \{\, -1 \le x, y \le 1 \,\}$ of unit density $\rho = 1$. (*b*) Use numerical integration to calculate the gravitational force at the points $(2,0)$ and $\left(\sqrt{2}, \sqrt{2}\right)$. Before starting, try to predict which point experiences the stronger force, and then check your prediction.

♠ 6.3.26. An equilateral triangular plate with unit area exerts a gravitational force on an observer sitting a unit distance away from its center. Is the force greater if the observer is located opposite a vertex of the triangle or opposite a side? Is the force greater than or less than that exerted by a circular plate of the same area? Use numerical integration to evaluate the double integrals.

6.3.27. Consider the wave equation $u_{tt} = c^2 u_{xx}$ on the line $-\infty < x < \infty$. Use the d'Alembert formula (2.82) to solve the initial value problem $u(0,x) = \delta(x-a)$, $u_t(0,x) = 0$. Can you realize your solution as the limit of classical solutions?

♢ 6.3.28. Consider the wave equation $u_{tt} = c^2 u_{xx}$ on the line $-\infty < x < \infty$. Use the d'Alembert formula (2.82) to solve the initial value problem $u(0,x) = 0$, $u_t(0,x) = \delta(x-a)$, modeling the effect of striking the string with a highly concentrated blow at the point $x = a$. Graph the solution at several times. Discuss the behavior of any discontinuities in the solution. In particular, show that $u(t,x) \neq 0$ on the domain of influence of the point $(a,0)$.

6.3.29. (*a*) Write down the solution $u(t,x)$ to the wave equation $u_{tt} = 4u_{xx}$ on the real line with initial data $u(0,x) = \begin{cases} 1 - |\,x\,|, & |\,x\,| \le 1, \\ 0, & \text{otherwise}, \end{cases}$ $\dfrac{\partial u}{\partial t}\,(0,x) = 0$. (*b*) Explain why $u(t,x)$ is not a classical solution to the wave equation. (*c*) Determine the derivatives $\partial^2 u/\partial t^2$ and $\partial^2 u/\partial x^2$ in the sense of distributions (generalized functions) and use this to justify the fact that $u(t,x)$ solves the wave equation in a distributional sense.

♡ 6.3.30. A piano string of length $\ell = 3$ and wave speed $c = 2$ with both ends fixed is hit by a hammer $\frac{1}{3}$ of the way along. The initial-boundary value problem that governs the resulting vibrations of the string is

$$\frac{\partial^2 u}{\partial t^2} = 4\,\frac{\partial^2 u}{\partial x^2}\,, \qquad u(t,0) = 0 = u(t,3), \qquad u(0,x) = 0, \qquad \frac{\partial u}{\partial t}\,(0,x) = \delta(x-1).$$

(*a*) What are the fundamental frequencies of vibration?
(*b*) Write down the solution to the initial-boundary value problem in Fourier series form.
(*c*) Write down the Fourier series for the velocity $\partial u/\partial t$ of your solution.
(*d*) Write down the d'Alembert formula for the solution, and sketch a picture of the string at four or five representative times.
(*e*) *True or false*: The solution is periodic in time. If true, what is the period? If false, explain what happens as t increases.

6.3.31. (*a*) Write down a Fourier series for the solution to the initial-boundary value problem

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

(*b*) Write down an analytic formula for the solution, i.e., sum your series. (*c*) In what sense does the series solution in part (*a*) converge to the true solution? Do the partial sums provide a good approximation to the actual solution?

6.3.32. Answer Exercise 6.3.31 for

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

Chapter 8
Linear and Nonlinear Evolution Equations

The term *evolution equation* refers to a dynamical partial differential equation that involves both time t and space $\mathbf{x} = (x_1, \dots, x_n)$ as independent variables and takes the form

$$\frac{\partial u}{\partial t} = K[u], \tag{8.1}$$

whose left-hand side is just the first-order time derivative of the dependent variable u, while the right-hand side, which can be linear or nonlinear, involves only u and its space derivatives and, possibly, t and $\mathbf{x}$. Examples already encountered include the linear and nonlinear transport equations in Chapter 2 and the heat equation. (But not the wave equation or Laplace equation.) In this chapter, we will analyze several important evolution equations, both linear and nonlinear, involving a single spatial variable.

Our first stop is to revisit the heat equation. We introduce the fundamental solution, which, for dynamical partial differential equations, assumes the role of the Green's function, in that its initial condition is a concentrated delta impulse. The fundamental solution leads to an integral superposition formula for the solutions produced by more general initial conditions or by external forcing. For the heat equation on the entire real line, the Fourier transform enables us to construct an explicit formula that identifies its fundamental solution as a Gaussian filter. We next present the Maximum Principle that rigorously justifies the entropic decay of temperature in a heated body and underlies much of the advanced mathematical analysis of parabolic partial differential equations. Finally, we discuss the Black–Scholes equation, the paradigmatic model for investment portfolios, first proposed in the early 1970s and now lying at the heart of the modern financial industry. We will find that the Black–Scholes equation can be transformed into the linear heat equation, whose fundamental solution is applied to establish the celebrated Black–Scholes formula for option pricing.

The following section provides a brief introduction to symmetry-based solution techniques for linear and nonlinear partial differential equations. Knowing a symmetry of a partial differential equation allows one to readily construct additional solutions from any known solution. Solutions that remain invariant under a one-parameter family of symmetries can be found by solving a reduced ordinary differential equation. The most important are the traveling wave solutions, which are invariant under translation symmetries, and similarity solutions, which are invariant under scaling symmetries.

The next evolution equation to appear is a paradigmatic model of nonlinear diffusion known as Burgers' equation. It can be regarded as a very simplified model of fluid dynamics, combining both nonlinear and viscous effects. We discover a remarkable nonlinear change

P.J. Olver, *Introduction to Partial Differential Equations*, Undergraduate Texts in Mathematics,
DOI 10.1007/978-3-319-02099-0_8,

of variables that maps Burgers' equation to the linear heat equation, and thereby facilitates its analysis, allowing us to construct explicit solutions, and investigate how they converge to shock wave solutions of the nonlinear transport equation in the inviscid limit.

Next, we turn our attention to the simplest third-order linear evolution equation, which arises as a model for wave mechanics. Unlike first- and second-order wave equations, its solutions are not simple traveling waves, but instead exhibit dispersion, in which oscillatory waves of different frequencies move at different speeds. As a result, initially localized disturbances will spread out or disperse, even while they conserve the underlying energy. Dispersion implies that the individual wave velocities differ from the group velocity, which measures the speed of propagation of energy in the system. An everyday manifestation of this phenomenon can be observed in the ripples caused by throwing a rock into a pond: the individual waves move faster than the overall disturbance. Finally, we present the remarkable Talbot effect, only recently discovered, in which solutions having discontinuous initial data and subject to periodic boundary conditions exhibit radically different profiles at rational and irrational times.

Our final example is the celebrated Korteweg–de Vries equation, which originally arose in the work of the nineteenth-century French applied mathematician Joseph Boussinesq as a model for surface waves on shallow water. It combines the effects of linear dispersion and nonlinear transport. Unlike the linearly dispersive model, the Korteweg–de Vries equation admits explicit, localized traveling wave solutions, now known as "solitons". Remarkably, despite the potentially complicated nonlinear nature of their interaction, two solitons emerge from a collision with their individual profiles preserved, the only residual effect being a relative phase shift. The Korteweg–de Vries equation is the prototype of a completely integrable partial differential equation, whose many remarkable properties were first discovered in the mid 1960s. A surprising number of such completely integrable nonlinear systems appear in a variety of applications, including dynamical models in fluids, plasmas, optics, and solid mechanics. Their analysis remains an extremely active area of contemporary research, [**2**, **36**].

8.1 The Fundamental Solution to the Heat Equation

One disadvantage of the Fourier series solution to the heat equation is that it is not nearly as explicit as one might desire for practical applications, numerical computations, or even further theoretical investigations and developments. An alternative approach is based on the idea of the *fundamental solution*, which plays the role of the Green's function in solving initial value problems. The fundamental solution measures the effect of a concentrated, instantaneous impulse, either in the initial conditions or as an external force on the system.

We restrict our attention to homogeneous boundary conditions — keeping in mind that these can always be included by use of linear superposition. The basic idea is to analyze the case in which the initial data $u(0, x) = \delta_\xi(x) = \delta(x - \xi)$ is a delta function, which we can interpret as a highly concentrated unit heat source, e.g., a soldering iron or laser beam, that is instantaneously applied at a position ξ along a metal bar. The heat will diffuse away from its initial concentration, and the resulting *fundamental solution* is denoted by

$$u(t, x) = F(t, x; \xi), \qquad \text{with} \qquad F(0, x; \xi) = \delta(x - \xi). \tag{8.2}$$

For each fixed ξ, the fundamental solution, considered as a function of $t > 0$ and x, must

satisfy the underlying partial differential equation, and so, for the heat equation,

$$\frac{\partial F}{\partial t} = \gamma \, \frac{\partial^2 F}{\partial x^2}\,, \tag{8.3}$$

along with the specified homogeneous boundary conditions.

As with the Green's function, once we have determined the fundamental solution, we can then use linear superposition to reconstruct the general solution to the initial-boundary value problem. Namely, we first write the initial data

$$u(0,x) = f(x) = \int_a^b \delta(x-\xi)\, f(\xi)\, d\xi \tag{8.4}$$

as a superposition of delta functions, as in (6.16). Linearity implies that the solution can be expressed as the corresponding superposition of the responses to those individual concentrated delta profiles:

$$u(t,x) = \int_a^b F(t,x;\xi)\, f(\xi)\, d\xi. \tag{8.5}$$

Assuming that we can differentiate under the integral sign, the fact that $F(t,x;\xi)$ satisfies the differential equation and the homogeneous boundary conditions for each fixed ξ immediately implies that the integral (8.5) is also a solution with the correct initial and (homogeneous) boundary conditions.

Unfortunately, most boundary value problems do not have fundamental solutions that can be written down in closed form. An important exception is the case of an infinitely long homogeneous bar, which requires solving the heat equation on the entire real line:

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}\,, \qquad \text{for} \qquad -\infty < x < \infty, \qquad t > 0. \tag{8.6}$$

For simplicity, we have chosen units in which the thermal diffusivity is $\gamma = 1$. The solution $u(t,x)$ is defined for all $x \in \mathbb{R}$, and has initial conditions

$$u(0,x) = f(x) \qquad \text{for} \qquad -\infty < x < \infty. \tag{8.7}$$

In order to specify the solution uniquely, we shall require that the temperature be square-integrable, i.e., in L^2, at all times, so that

$$\int_{-\infty}^{\infty} |\, u(t,x)\,|^2\, dx < \infty \qquad \text{for all} \qquad t \geq 0. \tag{8.8}$$

Roughly speaking, square-integrability requires that the temperature be vanishingly small at large distances, and hence plays the role of boundary conditions in this context.

To solve the initial value problem (8.6–7), we apply the Fourier transform, in the x variable, to both sides of the differential equation. In view of the effect of the Fourier transform on derivatives, cf. (7.43), the result is

$$\frac{\partial \widehat{u}}{\partial t} = -\,k^2\, \widehat{u}, \tag{8.9}$$

where

$$\widehat{u}(t,k) = \frac{1}{\sqrt{2\,\pi}} \int_{-\infty}^{\infty} u(t,x)\, e^{-\,\mathrm{i}\,k\,x}\, dx \tag{8.10}$$

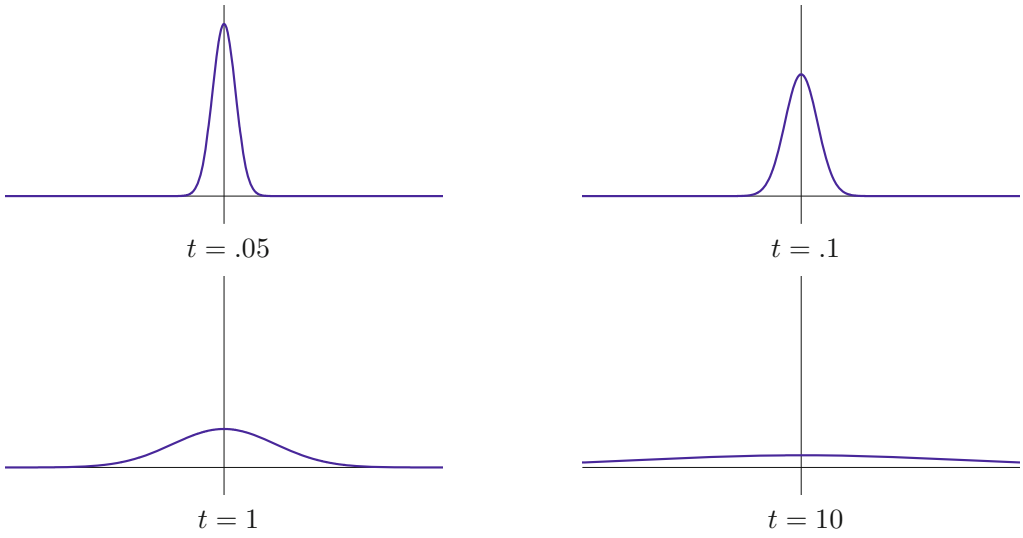


Figure 8.1. The fundamental solution to the one-dimensional heat equation.

is the Fourier transformed solution. For each fixed k, (8.9) can be viewed as a first-order linear ordinary differential equation for $\widehat{u}(t,k)$, with initial conditions

$$\widehat{u}(0,k) = \widehat{f}(k) = \frac{1}{\sqrt{2\,\pi}} \int_{-\infty}^{\infty} f(x)\, e^{-\mathrm{i}\,k\,x}\, dx \tag{8.11}$$

given by Fourier transforming the initial data (8.7). The solution to the initial value problem (8.9, 11) is immediate:

$$\widehat{u}(t,k) = e^{-k^2\,t}\, \widehat{f}(k). \tag{8.12}$$

We can thus recover the solution to the initial value problem (8.6–7) by applying the inverse Fourier transform to (8.12), leading to the explicit integral formula

$$u(t,x) = \frac{1}{\sqrt{2\,\pi}} \int_{-\infty}^{\infty} e^{\mathrm{i}\,k\,x}\, \widehat{u}(t,k)\, dk = \frac{1}{\sqrt{2\,\pi}} \int_{-\infty}^{\infty} e^{\mathrm{i}\,k\,x - k^2\,t}\, \widehat{f}(k)\, dk. \tag{8.13}$$

In particular, to construct the fundamental solution, we take the initial temperature profile to be a delta function $\delta_\xi(x) = \delta(x-\xi)$ concentrated at $x = \xi$. According to (7.37), its Fourier transform is

$$\widehat{\delta}_\xi(k) = \frac{e^{-\mathrm{i}\,k\,\xi}}{\sqrt{2\,\pi}}\,.$$

Plugging this into (8.13), and then referring to our table of Fourier transforms, we are led to the following explicit formula for the fundamental solution:

$$F(t,x;\xi) = \frac{1}{2\,\pi} \int_{-\infty}^{\infty} e^{\mathrm{i}\,k(x-\xi) - k^2\,t}\, dk = \frac{1}{2\sqrt{\pi\,t}}\, e^{-(x-\xi)^2/(4\,t)} \qquad \text{for} \qquad t > 0. \tag{8.14}$$

As you can verify, for each fixed ξ, the function $F(t,x;\xi)$ is indeed a solution to the heat equation for all $t > 0$. In addition,

$$\lim_{t\,\to\,0^+} F(t,x;\xi) = \begin{cases} 0, & x \neq \xi, \\ \infty, & x = \xi. \end{cases}$$

Furthermore, its integral

$$\int_{-\infty}^{\infty} F(t,x;\xi)\,dx = 1 \tag{8.15}$$

is constant — in accordance with the law of conservation of thermal energy; see Exercise 8.1.20. Therefore, as $t \to 0^+$, the fundamental solution satisfies the original limiting definition (6.8–9) of the delta function, and so $F(0,x;\xi) = \delta_\xi(x)$ has the desired initial temperature profile.

In Figure 8.1, we graph $F(t,x;0)$ at the indicated times. It starts life as a delta spike concentrated at the origin, and then immediately smooths out into a tall and narrow bell-shaped curve, centered at $x = 0$. As time increases, the solution shrinks and widens, eventually decaying everywhere to zero. Its amplitude is proportional to $t^{-1/2}$, while its overall width is proportional to $t^{1/2}$. The thermal energy (8.15), which is the area under the graph, remains fixed while gradually spreading out over the entire real line.

Remark: In probability, these exponentially bell-shaped curves are known as *normal* or *Gaussian distributions*, [**39**]. The width of the bell curve measures its *standard deviation*. For this reason, the fundamental solution to the heat equation is sometimes referred to as a *Gaussian filter*.

Remark: The fact that the fundamental solution depends only on the difference $x - \xi$, and hence has the same profile at all $\xi \in \mathbb{R}$, is a consequence of the translation invariance of the heat equation, reflecting the fact that it models the thermodynamics of a uniform medium. See Section 8.2 for additional symmetry properties of the heat equation and its solutions.

Remark: One of the striking properties of the heat equation is that thermal energy propagates with *infinite* speed. Indeed, because, at any $t > 0$, the fundamental solution is nonzero for all x, the effect of an initial concentration of heat will immediately be felt along the entire length of an infinite bar. (The graphs in Figure 8.1 are a little misleading because they fail to show the extremely small, but still positive, exponentially decreasing tails.) This effect, while more or less negligible at large distances, is nevertheless in clear violation of physical intuition — not to mention relativity, which postulates that signals cannot propagate faster than the speed of light. Despite this non-physical artifact, the heat equation remains an accurate model for heat propagation and similar diffusive phenomena, and so continues to be successfully used in applications.

With the fundamental solution in hand, we can adapt the linear superposition formula (8.5) to reconstruct the general solution

$$u(t,x) = \frac{1}{2\sqrt{\pi\, t}} \int_{-\infty}^{\infty} e^{-(x-\xi)^2/(4t)}\, f(\xi)\, d\xi \tag{8.16}$$

to our initial value problem (8.6). This solution formula is merely a restatement of (8.13) combined with the Fourier transform formula (8.11). Comparing with (7.54), we see that the solutions are obtained by convolution of the initial data with a one-parameter family of progressively wider and shorter Gaussian filters:

$$u(t,x) = F_0(t,x) * f(x), \qquad \text{where} \qquad F_0(t,x) = F(t,x;0) = \frac{e^{-x^2/(4t)}}{2\sqrt{\pi\, t}}\,.$$

Since $u(t,x)$ solves the heat equation, we conclude that Gaussian filter convolution has the same smoothing effect on the initial signal $f(x)$. Indeed, the convolution integral (8.16)

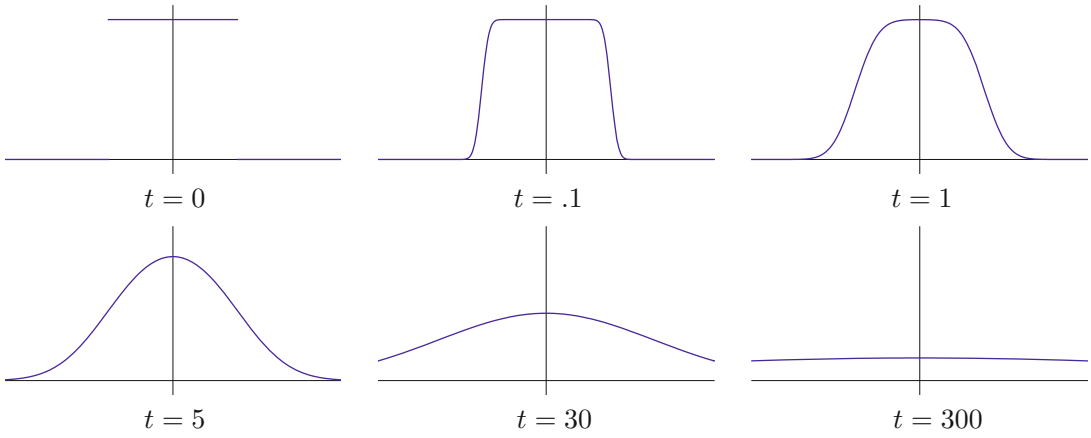


Figure 8.2. Error function solution to the heat equation.

serves to replace each initial value $f(x)$ by a weighted average of nearby values, the weight being determined by the Gaussian distribution. This has the effect of smoothing out high-frequency variations in the signal, and, consequently, the Gaussian convolution formula (8.16) provides an effective method for denoising rough signals and data.

Example 8.1. An infinite bar is initially heated to unit temperature along a finite interval. The initial temperature profile is thus a box function

$$u(0,x) = f(x) = \sigma(x-a) - \sigma(x-b) = \begin{cases} 1, & a < x < b, \\ 0, & \text{otherwise.} \end{cases}$$

The ensuing temperature is provided by the solution to the heat equation obtained by the integral formula (8.16):

$$u(t,x) = \frac{1}{2\sqrt{\pi\, t}} \int_a^b e^{-(x-\xi)^2/(4\,t)}\, d\xi = \frac{1}{2}\left[\operatorname{erf}\left(\frac{x-a}{2\sqrt{t}}\right) - \operatorname{erf}\left(\frac{x-b}{2\sqrt{t}}\right)\right], \tag{8.17}$$

where erf denotes the error function, as defined in (2.87). Graphs of the solution (8.17) for $a = -5$, $b = 5$, at the indicated times, are displayed in Figure 8.2. Observe the instantaneous smoothing of the sharp interface and instantaneous propagation of the disturbance, followed by a gradual decay to thermal equilibrium, with $u(t,x) \to 0$ as $t \to \infty$.

The Forced Heat Equation and Duhamel's Principle

The fundamental solution approach can be also applied to solve the inhomogeneous heat equation

$$u_t = u_{xx} + h(t,x), \tag{8.18}$$

modeling a bar subject to an external heat source $h(t,x)$, which might depend on both position and time. We begin by solving the particular case

$$u_t = u_{xx} + \delta(t-\tau)\,\delta(x-\xi), \tag{8.19}$$

whose inhomogeneity represents a heat source of unit magnitude that is concentrated at a position $x = \xi$ and applied at a single time $t = \tau > 0$. Physically, this models the effect of instantaneously applying a soldering iron to a single spot on the bar. Let us also impose homogeneous initial conditions

$$u(0, x) = 0 \tag{8.20}$$

as well as homogeneous boundary conditions of one of our standard types. The resulting solution

$$u(t, x) = G(t, x; \tau, \xi) \tag{8.21}$$

will be referred to as the *general fundamental solution* to the heat equation. Since a heat source that is applied at time τ will affect the solution only at later times $t \geq \tau$, we expect that

$$G(t, x; \tau, \xi) = 0 \qquad \text{for all} \qquad t < \tau. \tag{8.22}$$

Indeed, since $u(t, x)$ solves the unforced heat equation at all times $t < \tau$ subject to homogeneous boundary conditions and has zero initial temperature, this follows immediately from the uniqueness of the solution to the initial-boundary value problem.

Once we know the general fundamental solution (8.21), we are able to solve the problem for a general external heat source (8.18). We first write the forcing as a superposition

$$h(t, x) = \int_0^\infty \int_a^b \delta(t - \tau)\, \delta(x - \xi)\, h(\tau, \xi)\, d\xi\, d\tau \tag{8.23}$$

of concentrated instantaneous heat sources. Linearity allows us to conclude that the solution is given by the self-same superposition formula

$$u(t, x) = \int_0^t \int_a^b G(t, x; \tau, \xi)\, h(\tau, \xi)\, d\xi\, d\tau. \tag{8.24}$$

The fact that we only need to integrate over times $0 \leq \tau \leq t$ is a consequence of (8.22).

Remark: If we have a nonzero initial condition, $u(0, x) = f(x)$, then, by linear superposition, the solution

$$u(t, x) = \int_a^b F(t, x; \xi)\, f(\xi)\, d\xi + \int_0^t \int_a^b G(t, x; \tau, \xi)\, h(\tau, \xi)\, d\xi\, d\tau \tag{8.25}$$

is a combination of (a) the solution with no external heat source, but nonzero initial conditions, plus (b) the solution with homogeneous initial conditions but nonzero heat source.

Let us explicitly solve the forced heat equation on an infinite interval $-\infty < x < \infty$. We begin by computing the general fundamental solution. As before, we take the Fourier transform of both sides of the partial differential equation (8.18) with respect to x. In view of (7.37, 43), we find

$$\frac{\partial \widehat{u}}{\partial t} + k^2\, \widehat{u} = \frac{1}{\sqrt{2\,\pi}}\, e^{-\mathrm{i}\, k\, \xi}\, \delta(t - \tau), \tag{8.26}$$

which is an inhomogeneous first-order ordinary differential equation for the Fourier transform $\widehat{u}(t, k)$ of $u(t, x)$, while (8.20) implies the initial condition

$$\widehat{u}(0, k) = 0. \tag{8.27}$$

We solve the initial value problem (8.26–27) by the usual method, [**18**, **23**]. Multiplying the differential equation by the integrating factor $e^{k^2 t}$ yields

$$\frac{\partial}{\partial t}\left(e^{k^2 t}\, \widehat{u}\, \right) = \frac{1}{\sqrt{2\,\pi}}\, e^{k^2 t - \mathrm{i}\, k\, \xi}\, \delta(t-\tau).$$

Integrating both sides from 0 to t and using the initial condition, we obtain

$$\widehat{u}(t,k) = \frac{1}{\sqrt{2\,\pi}}\, e^{-k^2 (t-\tau) - \mathrm{i}\, k\, \xi}\, \sigma(t-\tau),$$

where $\sigma(s)$ is the usual step function (6.23). Finally, we apply the inverse Fourier transform formula (7.9), and then (8.14), to deduce that

$$\begin{aligned} u(t,x) = G(t,x;\tau,\xi) &= \frac{\sigma(t-\tau)}{2\,\pi} \int_{-\infty}^{\infty} e^{-k^2 (t-\tau) + \mathrm{i}\, k\, (x-\xi)}\, dk \\ &= \frac{\sigma(t-\tau)}{2\sqrt{\pi(t-\tau)}}\, \exp\left[-\frac{(x-\xi)^2}{4(t-\tau)} \right] = \sigma(t-\tau) F(t-\tau, x; \xi)\,. \end{aligned} \tag{8.28}$$

Thus, the general fundamental solution is obtained by translating the fundamental solution $F(t,x;\xi)$ for the initial value problem to a starting time of $t = \tau$ instead of $t = 0$. Finally, the superposition principle (8.24) produces the solution,

$$u(t,x) = \int_0^t \int_{-\infty}^{\infty} \frac{h(\tau,\xi)}{2\,\sqrt{\pi(t-\tau)}}\, \exp\left[-\frac{(x-\xi)^2}{4\,(t-\tau)} \right] d\xi\, d\tau, \tag{8.29}$$

to the heat equation with source term and zero initial condition on an infinite bar. A nonzero initial condition $u(0,x) = f(x)$ leads, as in the superposition formula (8.25), to an additional term of the form (8.16) in the solution formula.

Remark: The fact that an initial condition has the same aftereffect on the temperature as an instantaneous applied heat source of the same magnitude, thus implying the identification (8.28) of the two types of fundamental solution, is known as *Duhamel's Principle*, named after the nineteenth-century French mathematician Jean–Marie Duhamel. Duhamel's Principle remains valid over a broad range of linear evolution equations.

Example 8.2. An infinitely long bar with unit thermal diffusivity starts out uniformly at zero degrees. Beginning at time $t = 0$, a concentrated heat source of unit magnitude is continually applied at the origin. The resulting temperature is the solution $u(t,x)$ to the initial value problem

$$u_t = u_{xx} + \delta(x), \qquad u(0,x) = 0, \qquad t > 0, \qquad -\infty < x < \infty.$$

According to (8.29), the solution is given by

$$\begin{aligned} u(t,x) &= \int_0^t \int_{-\infty}^{\infty} \frac{\delta(\xi)}{2\,\sqrt{\pi(t-\tau)}}\, \exp\left[-\frac{(x-\xi)^2}{4\,(t-\tau)} \right] d\xi\, d\tau \\ &= \int_0^t \frac{1}{2\,\sqrt{\pi(t-\tau)}}\, \exp\left[-\frac{x^2}{4\,(t-\tau)} \right] d\tau = \sqrt{\frac{t}{\pi}}\, \exp\left[-\frac{x^2}{4\,t} \right] + \frac{x\, \mathrm{erf}\left(\dfrac{x}{2\,\sqrt{t}} \right) - |\,x\,|}{2}\,. \end{aligned}$$

Three snapshots can be seen in Figure 8.3. Observe that the solution is even in x and monotonically decreasing as $|\,x\,| \to \infty$. Moreover, it has a corner at the origin with limiting

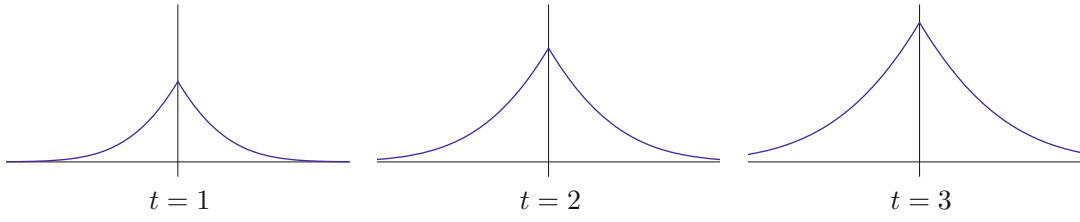


Figure 8.3. Effect of a concentrated heat source.

tangent lines of slopes $\pm \frac{1}{2}$, which implies that its second x derivative produces the delta-function forcing term. At each time t, the solution can be viewed as the linear superposition of a continuous family of fundamental solutions, corresponding to the cumulative effect of individual heat sources applied at each previous time $0 \le \tau \le t$. Moreover, it is not difficult to see that, at each fixed x, the temperature is monotonically increasing in t, with $u(t, x) \to \infty$ as $t \to \infty$, and hence the continuous heat source eventually produces an unbounded temperature in the entire infinite bar.

The Black–Scholes Equation and Mathematical Finance

The most important and influential partial differential equation in financial modeling and investment is the celebrated *Black–Scholes equation*

$$\frac{\partial u}{\partial t} + \frac{\sigma^2}{2}\, x^2\, \frac{\partial^2 u}{\partial x^2} + r\, x\, \frac{\partial u}{\partial x} - r\, u = 0, \tag{8.30}$$

first proposed in 1973 by the American economists Fischer Black and Myron Scholes, [**19**], and Robert Merton, [**71**]. The dependent variable $u(t, x)$ represents the monetary value of a single financial *option*, meaning a contract to either buy or sell an asset at a specified *exercise price* p at a certain future time $t_\star$. The value $u(t, x)$ of the option will depend on the current time $t \le t_\star$ and the current price $x \ge 0$ of the underlying asset. As with many financial models, one assumes the absence of arbitrage, meaning that there is no way to make a riskless profit. The constant $\sigma > 0$ represents the asset's *volatility*, while r denotes the (assumed fixed) *interest rate* for bank deposits, where investors could place their money with a guaranteed rate of return instead of buying the option. (Investors borrowing money to buy the asset would use a negative value of r.) The derivation of the Black–Scholes equation from basic financial modeling relies on the theory of stochastic differential equations, [**83**], which would take us too far afield to explain here; instead, we refer the interested reader to [**123**]. The Black–Scholes equation and its generalizations form the basis of much of the modern financial world, and, increasingly, the insurance industry.

Observe first that the Black–Scholes equation is a *backwards* diffusion process, since, upon solving for

$$\frac{\partial u}{\partial t} = -\frac{\sigma^2}{2}\, x^2\, \frac{\partial^2 u}{\partial x^2} - r\, x\, \frac{\partial u}{\partial x} + r\, u, \tag{8.31}$$

the coefficient of the diffusion term u_{xx} is *negative*. This implies that the initial value problem is well-posed only when time runs *backwards*. In other words, given a prescribed

value of the option at some specified time in the future, we can use the Black–Scholes equation to determine its current value. However, ill-posedness implies that we cannot predict future values from the current worth of the portfolio.

The "final value problem" for the Black–Scholes equation is to determine the option's value $u(t,x)$ at the current time t and asset value $x \geq 0$, given the *final condition*

$$u(t_\star, x) = f(x) \tag{8.32}$$

at the exercise time $t_\star > t$. For a so-called *European call option*, whereby the asset is to be bought at the exercise price $p > 0$ at the specified time, the final condition is

$$u(t_\star, x) = \max\{\, x - p,\ 0\,\}, \tag{8.33}$$

representing the investor's profit when $x > p$, or, when $x \leq p$, the option not being exercised so as to avoid a loss. Analogously, for a *put option*, where the asset is to be sold, the final condition is

$$u(t_\star, x) = \max\{\, p - x,\ 0\,\}. \tag{8.34}$$

The solution $u(t,x)$ will be defined for all $t < t_\star$ and all $x > 0$, subject to the boundary conditions

$$u(t,0) = 0, \qquad u(t,x) \ \sim\ x \qquad \text{as} \qquad x \to \infty,$$

where the asymptotic boundary condition means that the ratio $u(t,x)/x$ tends to a constant as $x \to \infty$.

Fortunately, the Black–Scholes equation can be solved explicitly by transforming it into the heat equation. The first step is to convert it to a forward diffusion process, by setting

$$\tau = \tfrac{1}{2}\sigma^2\,(t_\star - t), \qquad v(\tau, x) = u(t_\star - 2\,\tau/\sigma^2, x),$$

so that τ effectively runs forward from 0 as the actual time t runs backwards from $t_\star$. This substitution has the effect of converting the final condition (8.32) into an initial condition $v(0,x) = f(x)$. Moreover, a straightforward chain rule computation shows that v satisfies

$$\frac{\partial v}{\partial \tau} = x^2\,\frac{\partial^2 v}{\partial x^2} + \kappa\, x\,\frac{\partial v}{\partial x} - \kappa\, v, \qquad \text{where} \qquad \kappa = \frac{2\,r}{\sigma^2}\,.$$

The next step is to remove the explicit dependence on the independent variable x. The hint is that the right-hand side has the form of an Euler ordinary differential equation, [**23**, **89**]. According to Exercise 4.3.23, these terms can be placed into constant-coefficient form by the change of independent variables $x = e^y$. Indeed, writing

$$w(\tau, y) = v(\tau, e^y) = v(\tau, x) \qquad \text{when} \qquad x = e^y,$$

we apply the chain rule to compute the derivatives

$$\frac{\partial w}{\partial \tau} = \frac{\partial v}{\partial \tau}\,, \qquad \frac{\partial w}{\partial y} = e^y\,\frac{\partial v}{\partial x} = x\,\frac{\partial v}{\partial x}\,, \qquad \frac{\partial^2 w}{\partial y^2} = e^{2\,y}\,\frac{\partial^2 v}{\partial x^2} + e^y\,\frac{\partial v}{\partial x} = x^2\,\frac{\partial^2 v}{\partial x^2} + x\,\frac{\partial v}{\partial x}\,.$$

As a result, we find that w solves the partial differential equation

$$\frac{\partial w}{\partial \tau} = \frac{\partial^2 w}{\partial y^2} + (\kappa - 1)\,\frac{\partial w}{\partial y} - \kappa\, w. \tag{8.35}$$

This is getting closer to the heat equation, and, in fact, can be changed into it by setting

$$w(\tau, y) = e^{\alpha\,\tau + \beta\, y}\, z(\tau, y)$$

for suitable constants α, β. Indeed, differentiating and substituting into (8.35) yields

$$\frac{\partial z}{\partial \tau} + \alpha\, z = \frac{\partial^2 z}{\partial y^2} + 2\,\beta\,\frac{\partial z}{\partial y} + \beta^2 z + (\kappa - 1)\left(\frac{\partial z}{\partial y} + \beta\, z\right) - \kappa\, z.$$

The terms involving $\partial z/\partial y$ and z are eliminated by setting

$$\alpha = -\tfrac{1}{4}\,(\kappa + 1)^2, \qquad \beta = -\tfrac{1}{2}\,(\kappa - 1). \tag{8.36}$$

We conclude that the function

$$z(\tau, y) = e^{(\kappa+1)^2\tau/4 + (\kappa-1)y/2}\, w(\tau, y) \tag{8.37}$$

satisfies the heat equation

$$\frac{\partial z}{\partial \tau} = \frac{\partial^2 z}{\partial y^2}\,. \tag{8.38}$$

Unwinding the preceding argument, we have managed to prove the following:

Proposition 8.3. *If $z(\tau, y)$ is the solution to the initial value problem*

$$\frac{\partial z}{\partial \tau} = \frac{\partial^2 z}{\partial y^2}\,, \qquad z(0, y) = h(y) = e^{(\kappa-1)y/2} f(e^y), \tag{8.39}$$

for $\tau > 0$, $-\infty < y < \infty$, then

$$u(t, x) = x^{-(\kappa-1)/2} e^{-(\kappa+1)^2 \sigma^2 (t_\star - t)/8}\, z\big(\tfrac{1}{2}\,\sigma^2 (t_\star - t), \log x\big) \tag{8.40}$$

solves the final value problem (8.30, 32) *for the Black–Scholes equation for $t < t_\star$ and $0 < x < \infty$.*

Now, according to (8.16), the solution to the initial value problem (8.39) can be written as a convolution integral of the initial data with the heat equation's fundamental solution:

$$z(\tau, y) = \frac{1}{2\sqrt{\pi\,\tau}} \int_{-\infty}^{\infty} e^{-(y-\eta)^2/(4\tau)}\, h(\eta)\, d\eta = \frac{1}{2\sqrt{\pi\,\tau}} \int_{-\infty}^{\infty} e^{-(y-\eta)^2/(4\tau) + (\kappa-1)\eta/2} f(e^\eta)\, d\eta. \tag{8.41}$$

Combining this formula with (8.40) produces an explicit solution formula for the general final value problem for the Black–Scholes equation. In particular, for the European call option (8.33), the initial condition is

$$z(0, y) = h(y) = e^{(\kappa-1)y/2} \max\{\, e^y - p,\, 0\,\},$$

and so

$$z(\tau, y) = \frac{1}{2\sqrt{\pi\,\tau}} \int_{\log p}^{\infty} e^{-(y-\eta)^2/(4\tau) + (\kappa-1)\eta/2} (e^\eta - p)\, d\eta.$$

The integral can evaluated by completing the square inside the exponential, producing

$$z(\tau, y) = \frac{1}{2}\left[e^{(\kappa+1)^2\tau/4 + (\kappa+1)y/2}\, \operatorname{erfc}\left(\frac{\log p - (\kappa+1)\tau - y}{2\sqrt{\tau}}\right) - p\, e^{(\kappa-1)^2\tau/4 + (\kappa-1)y/2}\, \operatorname{erfc}\left(\frac{\log p - (\kappa-1)\tau - y}{2\sqrt{\tau}}\right)\right], \tag{8.42}$$

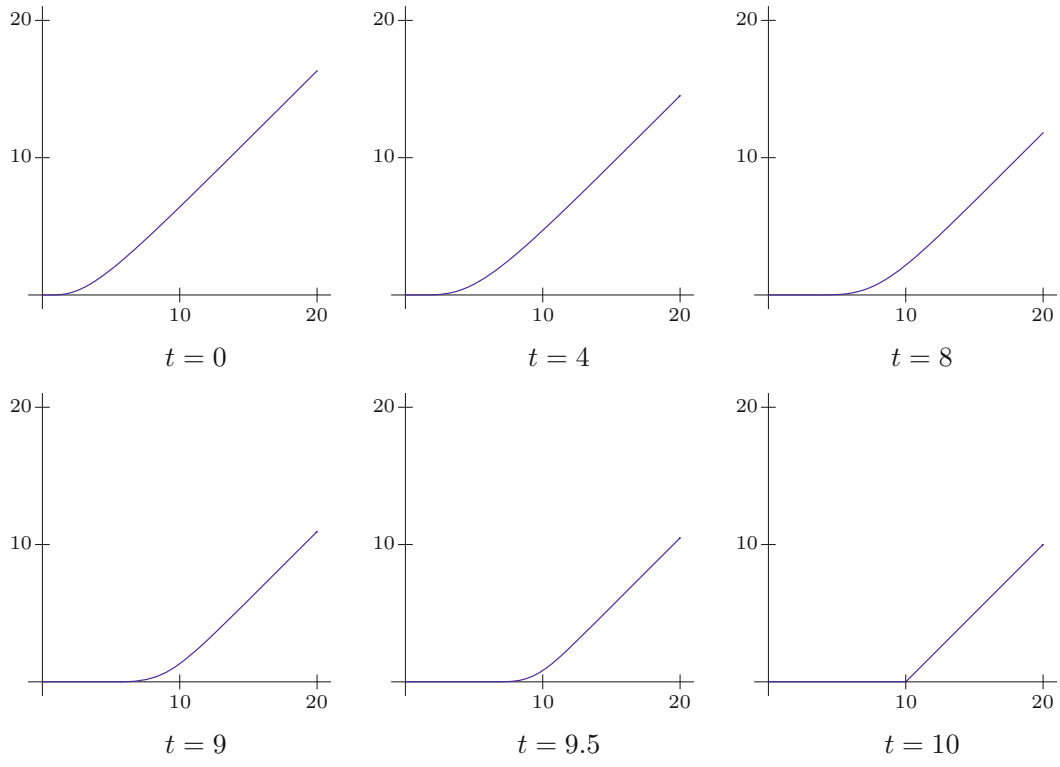


Figure 8.4. Solution to the Black–Scholes equation.

where

$$\operatorname{erfc} x = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-z^2}\, dz = 1 - \operatorname{erf} x \tag{8.43}$$

is the *complementary error function*, cf. (2.87). Substituting (8.42) into (8.40) results in the celebrated *Black–Scholes formula* for a European call option:

$$u(t,x) = \frac{1}{2}\left[x \operatorname{erfc}\left(-\frac{\left(r + \frac{1}{2}\sigma^2\right)(t_\star - t) + \log(x/p)}{\sqrt{2\sigma^2(t_\star - t)}} \right) - p\, e^{-r(t_\star - t)} \operatorname{erfc}\left(-\frac{\left(r - \frac{1}{2}\sigma^2\right)(t_\star - t) + \log(x/p)}{\sqrt{2\sigma^2(t_\star - t)}} \right) \right]. \tag{8.44}$$

A graph of the solution for the specific values $t_\star = 10$, $r = .1$, $\sigma = .2$, $p = 10$ appears in Figure 8.4. Observe that the option's value slowly decreases as the time gets closer and closer to the exercise time $t_\star$, thereby lessening any chances of further profit stemming from the option's underlying price volatility.

Exercises

8.1.1. Find the solution to the heat equation $u_t = u_{xx}$ on the real line having the following initial condition at time $t = 0$. Then sketch graphs of the resulting temperature distribution at times $t = 0, 1$, and 5.

(a) e^{-x^2}, (b) the step function $\sigma(x)$, (c) $e^{-|x|}$, (d) $\begin{cases} 1 - |x|, & |x| < 1, \\ 0, & \text{otherwise.} \end{cases}$

8.1.2. On an infinite bar with unit thermal diffusivity, a concentrated unit heat source is instantaneously applied at the origin at time $t = 0$. A heat sensor measures the resulting temperature in the bar at position $x = 1$. Determine the maximum temperature measured by the sensor. At what time is the maximum achieved?

8.1.3. (a) Find the solution to the heat equation (8.6) whose initial data corresponds to a pair of unit heat sources placed at positions $x = \pm 1$. (b) Graph the solution at times $t = .1, .25, .5, 1$. (c) At what time(s) does the origin experience its maximum overall temperature? What is the maximum temperature at the origin?

8.1.4. (a) Use the Fourier transform to solve the initial value problem

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \qquad u(0, x) = \delta'(x - \xi), \qquad -\infty < x < \infty, \quad t > 0,$$

whose initial data is the derivative of the delta function at a fixed position ξ.

(b) Show that your solution can be written as the derivative $\partial F/\partial x$ of the fundamental solution $F(t, x; \xi)$. Explain why this observation should be valid.

8.1.5. Suppose that the initial data $u(0, x) = f(x)$ is real. Explain why the Fourier transform solution formula (8.13) defines a real function $u(t, x)$ for all $t > 0$.

8.1.6. (a) What is the maximum value of the fundamental solution at time t?

(b) Can you justify the claim that its width is proportional to $\sqrt{t}$?

8.1.7. Prove directly that (8.5) is indeed a solution to the heat equation, and, moreover, has the correct initial and boundary conditions.

8.1.8. Show, by a direct computation, that the final formula in (8.14) is a solution to the heat equation for all $t > 0$.

♢ 8.1.9. Justify formula (8.15).

8.1.10. According to Exercises 4.1.11–12, both the t and x partial derivatives of the fundamental solution solve the heat equation. (a) Write down the initial value problem satisfied by these two solutions. (b) Set $\xi = 0$ and then sketch graphs of each solution at several selected times. (c) Reconstruct each solution as a Fourier integral.

8.1.11. Let $u(t, x) = \dfrac{\partial F}{\partial x}(t, x; 0)$ denote the x derivative of the fundamental solution (8.14).

(a) Prove that $u(t, x)$ is a solution to the heat equation $u_t = u_{xx}$ on the domain $\{-\infty < x < \infty,\ t > 0\}$. (b) For fixed x, prove that $\lim_{t \to 0^+} u(t, x) = 0$. (c) Explain why, despite the results in parts (a) and (b), $u(t, x)$ is *not* a classical solution to the initial value problem $u_t = u_{xx}$, $u(0, x) = 0$. What is the classical solution? (d) What initial value problem does $u(t, x)$ satisfy?

8.1.12. Justify all the statements in Example 8.2.

♡ 8.1.13. (a) Solve the heat equation on an infinite bar when the initial temperature is equal to 1 for $|x| < 1$ and 0 elsewhere, while a unit heat source is applied to the same part of the bar $|x| < 1$ for a unit time period $0 < t < 1$. (b) At what time and what location is the bar the hottest? (c) What is the final equilibrium temperature of the bar?

8.1.14. An insulated bar 1 meter long, with constant diffusivity $\gamma = 1$, is taken from a freezer that is kept at -10° C, and then has its ends kept at room temperature of 20° C. A soldering iron with temperature 350° C is continually held at the midpoint of the bar.
(*a*) Set up an initial value problem modeling the temperature distribution in the bar.
(*b*) Find the corresponding equilibrium temperature distribution.

♡ 8.1.15. Consider the heat equation with unit thermal diffusivity on the interval $0 < x < 1$ subject to homogeneous Dirichlet boundary conditions.
(*a*) Find a Fourier series representation for the fundamental solution $\widehat{F}(t, x; \xi)$ that solves the initial-boundary value problem
$$u_t = u_{xx}, \quad t > 0, \quad 0 < x < 1, \qquad u(0, x) = \delta(x - \xi), \quad u(t, 0) = 0 = u(t, 1).$$
Your solution should depend on t, x and the point ξ where the initial delta impulse is applied.
(*b*) For the value $\xi = .3$, use a computer program to sum the first few terms in the series and graph the result at times $t = .0001, .001, .01$, and $.1$. Make sure you have included enough terms to obtain a reasonably accurate graph.
(*c*) Compare your graphs with those of the fundamental solution $F(t, x; .3)$ on an infinite interval at the same times. What is the maximum deviation between the two solutions on the entire interval $0 \le x \le 1$?
(*d*) Use your fundamental solution $\widehat{F}(t, x; \xi)$ to construct a series solution to the general initial value problem $u(0, x) = f(x)$. Is your series the same as the usual Fourier series solution? If not, explain any discrepancy.

8.1.16. *True or false*: Periodic forcing of the heat equation at a particular frequency can produce resonance. Justify your answer.

8.1.17. Find the fundamental solution for the *cable equation* $v_t = \gamma\, v_{xx} - \alpha\, v$ on the real line. *Hint*: See Exercise 4.1.16.

8.1.18. The partial differential equation $u_t + c\, u_x = \gamma\, u_{xx}$ models transport of a diffusing pollutant in a fluid flow. Assuming that the speed c is constant, write down a solution to the initial value problem $u(0, x) = f(x)$ for $-\infty < x < \infty$. *Hint*: Look at Exercise 4.1.17.

♢ 8.1.19. Use the Fourier transform to solve the initial value problem $\mathrm{i}\, u_t = u_{xx}$, $u(0, x) = f(x)$, for the one-dimensional Schrödinger equation on the real line $-\infty < x < \infty$.

♢ 8.1.20. Let $u(t, x)$ be a solution to the heat equation having finite thermal energy, $E(t) = \displaystyle\int_{-\infty}^{\infty} u(t, x)\, dx < \infty$, and satisfying $u_x(t, x) \to 0$ as $x \to \pm\infty$, for all $t \ge 0$. Prove the law of *conservation of thermal energy*: $E(t) =$ constant.

8.1.21. Explain in your own words how a function $u(t, x)$ can satisfy $u(t, x) \to 0$ uniformly as $t \to \infty$ while maintaining the constancy of $\displaystyle\int_{-\infty}^{\infty} u(t, x)\, dx = 1$ for all t. Discuss what this signifies regarding the interchange of limits and integrals.

8.1.22. (*a*) Prove that if $\widehat{f}(k) \in \mathrm{L}^2$ is square-integrable, then so is $e^{-a\,k^2}\, \widehat{f}(k)$ for any $a > 0$.
(*b*) Prove that when the initial data $f(x) \in \mathrm{L}^2$ is square integrable, so is the Fourier integral solution (8.13) for all $t \ge 0$.

8.1.23. Find the solution to the Black–Scholes equation for a put option (8.34).

8.1.24. (*a*) If we increase the interest rate r, does the value of a call option (*i*) increase; (*ii*) decrease; (*iii*) stay the same; (*iv*) could do any of the above? Justify your answer.
(*b*) Answer the same question when rate stays fixed, but the volatility σ is increased.

♢ 8.1.25. Justify formula (8.42).

8.2 Symmetry and Similarity

The geometric approach to partial differential equations enables one to exploit their symmetry properties to construct explicit solutions of both mathematical and physical interest. Unlike separation of variables, which is restricted to special types of linear partial differential equations,† symmetry methods can also be successfully applied to a broad range of nonlinear partial differential equations. While we do not have the mathematical tools to develop the full range of symmetry techniques, we will learn how to exploit some of the most basic symmetry properties: translations, leading to traveling wave solutions; scalings, leading to similarity solutions; and, in subsequent chapters, rotational symmetries.

In general, by a *symmetry* of an equation, we mean a transformation that takes solutions to solutions. Thus, knowing a symmetry transformation, if we are in possession of one solution, then we can construct a second solution by applying the symmetry. And, possibly, a third solution by applying the symmetry yet again. And so on. If we know lots of symmetries, then we can produce lots of solutions by this simple device.

Remark: General symmetry techniques are founded on the theory of *Lie groups*, named after the influential nineteenth-century Norwegian mathematician Sophus Lie (pronounced "Lee"). Lie's theory is a profound synthesis of group theory and differential geometry, and provides an algorithm for completely determining all the (continuous) symmetries of a given differential equation. Although the theory lies beyond the scope of this introductory text, direct inspection and/or physical intuition will often produce the most important symmetries of the system, which can then be directly exploited. Modern applications of Lie's symmetry methods to partial differential equations arising in physics and engineering can be traced back to an influential book on hydrodynamics by the author's thesis advisor, Garrett Birkhoff, [**17**]. A complete and comprehensive treatment of Lie symmetry methods can be found in the author's first book [**87**], and, at a more introductory level, in the recent books [**27**, **58**], the first having a particular emphasis on applications in fluid mechanics.

The heat equation serves as an excellent testing ground for the general methodology, since it admits a rich variety of symmetry transformations that take solutions to solutions. The simplest are the translations. Moving the space and time coordinates by a fixed amount,

$$t \longmapsto t + a, \qquad x \longmapsto x + b, \tag{8.45}$$

where a, b are constants, changes the function $u(t, x)$ into the translated function‡

$$U(t, x) = u(t - a, x - b). \tag{8.46}$$

A simple application of the chain rule proves that the partial derivatives of U with respect to t and x agree with the corresponding partial derivatives of u, so

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

† This is not entirely fair: separation of variables can also be applied to certain nonlinear partial differential equations such as Hamilton–Jacobi equations, [**73**].

‡ The minus signs arise because when we set $\widehat{t} = t + a$, $\widehat{x} = x + b$, then the translated function is $U(\widehat{t}, \widehat{x}) = u(t, x) = u(\widehat{t} - a, \widehat{x} - b)$. Dropping the hats produces the stated formula.

and so on. In particular, the function $U(t,x)$ is a solution to the heat equation $U_t = \gamma\, U_{xx}$ whenever $u(t,x)$ also solves $u_t = \gamma\, u_{xx}$. Physically, translation symmetry formalizes the property that the heat equation models a homogeneous medium, and hence the solution does not depend on the choice of reference point or origin of our coordinate system.

As a consequence, each solution to the heat equation will produce an infinite family of translated solutions. For example, starting with the separable solution

$$u(t,x) = e^{-\gamma t}\, \sin x,$$

we immediately produce the additional translated solutions

$$U(t,x) = e^{-\gamma (t-a)}\, \sin(x-b),$$

valid for any choice of constants a, b.

Warning: Typically, the symmetries of a differential equation do not respect initial or boundary conditions. For instance, if $u(t,x)$ is defined for $t \geq 0$ and in the domain $0 \leq x \leq \ell$, then its translated version (8.46) is defined for $t \geq a$ and in the translated domain $b \leq x \leq \ell + b$, and so will solve a translated initial-boundary value problem.

A second important class of symmetries consists of the scaling invariances. We already know that if $u(t,x)$ is a solution, then so is the scalar multiple $c\, u(t,x)$ for any constant c; this is a simple consequence of linearity of the heat equation. We can also add an arbitrary constant to the temperature, noting that

$$U(t,x) = c\, u(t,x) + k \tag{8.47}$$

is a solution for any choice of constants c, k. Physically, the transformation (8.47) amounts to a change in the scale used to measure temperature. For instance, if u is measured in degrees Celsius, and we set $c = \frac{9}{5}$ and $k = 32$, then $U = \frac{9}{5}\, u + 32$ will be measured in degrees Fahrenheit. Thus, reassuringly, the physical processes described by the heat equation do not depend on our choice of thermometer.

More interestingly, suppose we rescale the space and time variables:

$$t \;\longmapsto\; \alpha\, t, \qquad\qquad x \;\longmapsto\; \beta\, x, \tag{8.48}$$

where $\alpha, \beta \neq 0$ are nonzero constants. The effect of such a scaling transformation is to convert $u(t,x)$ into a rescaled function[†]

$$U(t,x) = u(\alpha^{-1} t, \beta^{-1} x). \tag{8.49}$$

The derivatives of U are related to those of u according to the formulas

$$\frac{\partial U}{\partial t} = \frac{1}{\alpha}\, \frac{\partial u}{\partial t}, \qquad\qquad \frac{\partial U}{\partial x} = \frac{1}{\beta}\, \frac{\partial u}{\partial x}, \qquad\qquad \frac{\partial^2 U}{\partial x^2} = \frac{1}{\beta^2}\, \frac{\partial^2 u}{\partial x^2}.$$

Therefore, if u satisfies the heat equation $u_t = \gamma\, u_{xx}$, then U satisfies the rescaled heat equation

$$U_t = \frac{1}{\alpha}\, u_t = \frac{\gamma}{\alpha}\, u_{xx} = \frac{\beta^2\, \gamma}{\alpha}\, U_{xx},$$

[†] As before, setting $\widehat{t} = \alpha\, t$, $\widehat{x} = \beta\, x$, produces the rescaled function $U(\widehat{t}, \widehat{x}) = u(t,x) = u(\alpha^{-1}\, \widehat{t}, \beta^{-1}\, \widehat{x})$, and we then drop the hats.

which we rewrite as

$$U_t = \Gamma\, U_{xx}, \qquad \text{where} \qquad \Gamma = \frac{\beta^2\,\gamma}{\alpha}\,. \tag{8.50}$$

Thus, the net effect of scaling space and time is merely to rescale the diffusion coefficient. Physically, the scaling symmetry (8.48) corresponds to a change in the physical units used to measure time and distance. For instance, to change from minutes to seconds, set $\alpha = 60$, and from yards to meters, set $\beta = .9144$. The net effect (8.50) on the diffusion coefficient γ is a reflection of its physical units, namely distance^2/time.

In particular, if we choose

$$\alpha = \gamma, \qquad \beta = 1,$$

then the rescaled diffusion coefficient becomes $\Gamma = 1$. This observation has the following important consequence. If $U(t,x)$ solves the heat equation for a unit diffusivity, $\Gamma = 1$, then

$$u(t,x) = U(\gamma\, t, x) \tag{8.51}$$

solves the heat equation for the diffusivity $\gamma > 0$. Thus, the only effect of the diffusion coefficient is to speed up or slow down time. A body with diffusivity $\gamma = 2$ will cool down twice as fast as a body (of the same shape subject to similar boundary conditions and initial conditions) with diffusivity $\gamma = 1$. Note that this particular rescaling has not altered the space coordinates, and so $U(t,x)$ is defined on the same spatial domain as $u(t,x)$.

On the other hand, if we set $\alpha = \beta^2$, then the rescaled diffusion coefficient is exactly the same as the original: $\Gamma = \gamma$. Thus, the transformation

$$t \;\longmapsto\; \beta^2\, t, \qquad\qquad x \;\longmapsto\; \beta\, x, \tag{8.52}$$

does not alter the equation, and hence defines a *scaling symmetry* — also known as a *similarity transformation* — for the heat equation. Combining (8.52) with the linear rescaling $u \mapsto c\, u$, we make the elementary, but important, observation that if $u(t,x)$ is any solution to the heat equation, then so is the function

$$U(t,x) = c\, u(\beta^{-2}\, t, \beta^{-1}\, x), \tag{8.53}$$

for the *same* diffusion coefficient γ. For example, rescaling the solution

$$u(t,x) = e^{-\gamma\, t} \cos x \qquad \text{leads to the solution} \qquad U(t,x) = c\, e^{-\gamma\, t/\beta^2} \cos \frac{x}{\beta}\,.$$

Warning: As in the case of translations, rescaling space by a factor $\beta \neq 1$ will alter the domain of definition of the solution. If $u(t,x)$ is defined for $a \leq x \leq b$, then $U(t,x)$, as given in (8.53), is defined for $\beta\, a \leq x \leq \beta\, b$ (or, when $\beta < 0$, for $\beta\, b \leq x \leq \beta\, a$).

For example, suppose that we have solved the heat equation for the temperature $u(t,x)$ on a bar of length 1, subject to certain initial and boundary conditions. We are then given a bar composed of the same material of length 2. Since the diffusivity coefficient has not changed, we can directly construct the new solution $U(t,x)$ by rescaling. Setting $\beta = 2$ will serve to double the length. If we also rescale time by a factor $\alpha = \beta^2 = 4$, then the rescaled function $U(t,x) = u\big(\frac{1}{4}\, t, \frac{1}{2}\, x\big)$ will be a solution of the heat equation on the longer bar with the same diffusivity constant. The net effect is that the rescaled solution will be evolving four times as slowly as the original, and hence it effectively takes a bar that is twice the length four times as long to cool down.

Similarity Solutions

A *similarity solution* of a partial differential equation is one that remains unchanged (invariant) under a one-parameter family† of scaling symmetryscaling symmetries. For a partial differential equation in two variables — say t and x — the similarity solutions can be found by solving an *ordinary differential equation*.

Suppose our partial differential equation admits the scaling symmetries

$$t \longmapsto \beta^a t, \qquad x \longmapsto \beta^b x, \qquad u \longmapsto \beta^c u, \qquad \beta \neq 0, \tag{8.54}$$

where a, b, c are fixed constants with a, b not both zero. As above, this means that if $u(t, x)$ is a solution to the differential equation, so is the rescaled function

$$U(t, x) = \beta^c u(\beta^{-a} t, \beta^{-b} x) \tag{8.55}$$

for all values of $\beta \neq 0$. Checking that this indeed defines a symmetry is a simple matter of applying the chain rule, which implies that the derivatives scale according to

$$u_t \longmapsto \beta^{c-a} u_t, \quad u_x \longmapsto \beta^{c-b} u_x, \quad u_{tt} \longmapsto \beta^{c-2a} u_{tt}, \quad u_{xt} \longmapsto \beta^{c-a-b} u_{xt}, \tag{8.56}$$

and so on. Products of derivatives scale multiplicatively, e.g., $x^4 u u_{xt} \mapsto \beta^{2c-a+3b} x^4 u u_{xt}$. In order that a (polynomial) differential equation admit such a scaling symmetry, each of its terms must scale by the *same* overall power of β.

By definition, $u(t, x)$ is called a *similarity solution* if it remains unchanged (invariant) under the scaling symmetries (8.54), so that

$$u(t, x) = \beta^c u(\beta^{-a} t, \beta^{-b} x) \tag{8.57}$$

for all $\beta > 0$. Let us, for specificity, assume that $a \neq 0$, leaving the case $a = 0$, $b \neq 0$, for the reader to complete in Exercise 8.2.13. Since the left-hand side of (8.57) does not depend on β, we can fix its value to be‡ $\beta = t^{1/a}$, and conclude that the similarity solution must have the form

$$u(t, x) = t^{c/a} v(\xi), \qquad \text{where} \qquad \xi = x\, t^{-b/a} \qquad \text{and} \qquad v(\xi) = u(1, \xi), \tag{8.58}$$

are referred to as the *similarity variables*, since they remain invariant when subjected to the scaling transformations (8.54). We then use the chain rule to find the formulas for the partial derivatives of u in terms of the ordinary derivatives of v with respect to ξ. Substituting these expressions into the scale-invariant partial differential equation for $u(t, x)$, and then canceling a common factor of t, will effectively reduce it to an *ordinary differential equation* for the function $v(\xi)$. Each solution to the resulting ordinary differential equation then gives rise to a scale-invariant solution to the original partial differential equation through the similarity ansatz (8.58).

Example 8.4. As a first example, let us return to the nonlinear transport equation

$$u_t + u u_x = 0, \tag{8.59}$$

† Or, more accurately, a one-parameter group, [**87**].

‡ This assumes $t > 0$; for $t < 0$, just replace t by $-t$.

which we studied in Section 2.3. Under (8.54, 56), the equation rescales to

$$\beta^{c-a} u_t + \beta^{2c-b} u\, u_x = 0,$$

which is unchanged, provided $c - a = 2c - b$, and hence $c = b - a$. Setting $a = 1$, $c = b - 1$, we conclude that if $u(t, x)$ is any solution, then so is the rescaled function

$$U(t, x) = \beta^{b-1}\, u(\beta^{-1} t, \beta^{-b} x)$$

for any b and any $\beta \neq 0$.

To find the associated similarity solutions, we use (8.58) to introduce the ansatz

$$u(t, x) = t^{b-1}\, v(\xi), \qquad \text{where} \qquad \xi = x\, t^{-b}. \tag{8.60}$$

Differentiating, we obtain

$$u_t = -b\, x\, t^{-2}\, v'(\xi) + (b-1)\, t^{b-2}\, v(\xi) = t^{b-2}\big[-b\,\xi\, v'(\xi) + (b-1)\, v(\xi)\big], \qquad u_x = t^{-1}\, v'(\xi).$$

Substituting these expressions into the transport equation (8.59) yields

$$0 = u_t + u\, u_x = t^{b-2}\big[(v - b\,\xi)\, v' + (b-1)\, v\big],$$

and so

$$(v - b\,\xi)\, \frac{dv}{d\xi} + (b-1)\, v = 0. \tag{8.61}$$

Any solution to this nonlinear first-order ordinary differential equation will, when substituted into (8.60), produce a similarity solution to the nonlinear transport equation.

If $b = 1$, then either $v = b\,\xi$, producing the particular similarity solution $u(t, x) = x/t$ that we earlier used to construct the rarefaction wave (2.54), or v is constant, and so is u. Otherwise, we can, in fact, linearize (8.61) by treating ξ as a function of v, whence

$$(b-1)\, v\, \frac{d\xi}{dv} - b\,\xi = -v.$$

The general solution to such a linear first-order ordinary differential equation is found by the standard method, [**18**, **23**], resulting in

$$\xi = v + k\, v^{b/(b-1)},$$

where k is the constant of integration. Recalling (8.60), we find that the similarity solutions $u(t, x)$ are defined by an implicit equation

$$x = k\, u^{b/(b-1)} + t\, u.$$

For example, if $b = 2$, the (multi-valued) solution is a sideways-moving parabola:

$$x = k\, u^2 + t\, u, \qquad \text{so that} \qquad u = \frac{-t \pm \sqrt{t^2 + 4\, k\, x}}{2\, k}.$$

Example 8.5. Consider the linear heat equation

$$u_t = u_{xx}. \tag{8.62}$$

Under the rescaling (8.54), the equation becomes $\beta^{c-a} u_t = \beta^{c-2b} u_{xx}$, and thus (8.54) represents a symmetry if and only if $a = 2b$. Therefore, if $u(t, x)$ is any solution, so is the rescaled function

$$U(t, x) = \beta^{c}\, u(\beta^{-2} t, \beta^{-1} x).$$

Of course, the initial scaling factor stems from the linearity of the equation.

The scale-invariant solutions are constructed through the similarity ansatz

$$u(t,x) = t^{c/2}\, v(\xi), \qquad \text{where} \qquad \xi = x/\sqrt{t}\,.$$

Differentiation yields

$$u_t = -\tfrac{1}{2}\, x\, t^{c/2-3/2}\, v'(\xi) + \tfrac{1}{2}\, c\, t^{c/2-1}\, v(\xi) = t^{c/2-1}\big[-\tfrac{1}{2}\,\xi\, v'(\xi) + \tfrac{1}{2}\, c\, v(\xi)\big],$$
$$u_{xx} = t^{c/2-1}\, v''(\xi).$$

Substituting these expressions into the heat equation and canceling a common power of t, we find that v must satisfy the linear ordinary differential equation

$$v'' + \tfrac{1}{2}\,\xi\, v' - \tfrac{1}{2}\, c\, v = 0. \tag{8.63}$$

If $c = 0$, then (8.63) is effectively a linear first-order ordinary differential equation for $v'(\xi)$, which can be readily solved by the usual method, thereby producing the solution

$$v(\xi) = c_1 + c_2 \operatorname{erf}\big(\tfrac{1}{2}\,\xi\big),$$

where c_1, c_2 are arbitrary constants and erf is the error function (2.87). The corresponding similarity solution to the heat equation is

$$u(t,x) = c_1 + c_2 \operatorname{erf}\left(\frac{x}{\sqrt{t}}\right).$$

The error function solutions that we encountered in (8.17) can be built up as a linear combination of translations of this similarity solution.

If $c \neq 0$, most solutions to the ordinary differential equation (8.63) are not elementary functions.† One is in need of more sophisticated techniques, e.g., the method of power series to be developed in Section 11.3, to understand its solutions, and hence the resulting similarity solutions to the heat equation.

Exercises

8.2.1. If it takes a 2 cm long insulated bar 23 minutes to cool down to room temperature, how long does it take a 4 cm bar?

8.2.2. If it takes a 5 centimeter long insulated iron bar 10 minutes to cool down so as not to burn your hand, how long does it take a 20 centimeter bar made out of the same material to cool down to the same temperature?

♢ 8.2.3. (a) Given $\gamma > 0$, use a scaling transformation to write down the formula for the fundamental solution for the general heat equation $u_t = \gamma\, u_{xx}$ for $x \in \mathbb{R}$. (b) Write down the corresponding integral formula for the solution to the initial value problem.

† According to [**87**; Example 3.3], the general solution can be written in terms of parabolic cylinder functions, [**86**].

8.2.4. Use scaling to construct the series solution for a heated circular ring of radius r and thermal diffusivity γ. Does scaling also give the correct formulas for the Fourier coefficients in terms of the initial temperature distribution?

8.2.5. A solution $u(t,x)$ to the heat equation is measured in degrees Fahrenheit. What is the corresponding temperature in degrees Kelvin? Which symmetry transformation takes the first solution to the second solution, and how does it affect the diffusion coefficient?

8.2.6. Is time reversal, $t \mapsto -t$, a symmetry of the heat equation? Write down a physical explanation, and then a mathematical justification.

8.2.7. According to Exercise 4.1.17, the partial differential equation $u_t + c\,u_x = \gamma\,u_{xx}$ models diffusion in a convective flow. Show how to use scaling to place the differential equation in the form $u_t + u_x = P^{-1}\,u_{xx}$, where P is called the *Péclet number*, and controls the rate of mixing. Is there a scaling that will reduce the problem to the case $P = 1$?

8.2.8. Suppose you know a solution $u^\star(t,x)$ to the heat equation that satisfies $u^\star(1,x) = f(x)$. Explain how to solve the initial value problem with $u(0,x) = f(x)$.

8.2.9. Solve the following initial value problems for the heat equation $u_t = u_{xx}$ for $x \in \mathbb{R}$:
(a) $u(0,x) = e^{-x^2/4}$. *Hint*: Use Exercise 8.2.8. (b) $u(0,x) = e^{-4x^2}$.
(c) $u(0,x) = x^2\,e^{-x^2/4}$. *Hint*: Use Exercise 4.1.12.

8.2.10. Define the functions $H_n(x)$ for $n = 0, 1, 2, \ldots$, by the formula

$$\frac{d^n}{dx^n}\,e^{-x^2} = (-1)^n H_n(x)\,e^{-x^2}. \tag{8.64}$$

(a) Prove that $H_n(x)$ is a polynomial of degree n, known as the n^{th} *Hermite polynomial*.
(b) Calculate the first four Hermite polynomials.
(c) Assuming $\gamma = 1$, find the solution to the heat equation for $-\infty < x < \infty$ and $t > 0$, given the initial data $u(0,x) = H_n(x)\,e^{-x^2}$. *Hint*: Combine Exercises 4.1.11, 8.2.8.

8.2.11. Find the scaling symmetries and corresponding similarity solutions of the following partial differential equations: (a) $u_t = x^2\,u_x$, (b) $u_t + u^2 u_x = 0$, (c) $u_{tt} = u_{xx}$.

8.2.12. Show that the wave equation $u_{tt} = c^2 u_{xx}$ has the following invariance properties: if $u(t,x)$ is a solution, so is (a) any time translate: $u(t-a,x)$, where a is fixed; (b) any space translate: $u(t,x-b)$, where b is fixed; (c) the dilated function $u(\beta t, \beta x)$ for $\beta \neq 0$; (d) any derivative: say $\partial u/\partial x$ or $\partial^2 u/\partial t^2$, provided u is sufficiently smooth.

♢ 8.2.13. Suppose $a = 0$, $b \neq 0$ in the scaling transformation (8.57).
(a) Discuss how to reduce the partial differential equation to an ordinary differential equation for the corresponding similarity solutions.
(b) Illustrate your method with the partial differential equation $t\,u_t = u\,u_{xx}$.

8.2.14. *True or false*: (a) A homogeneous polynomial solution to a partial differential equation is always a similarity solution. (b) An inhomogeneous polynomial solution to a partial differential equation can never be a similarity solution.

8.2.15. (a) Find all scaling symmetries of the two-dimensional Laplace equation $u_{xx} + u_{yy} = 0$.
(b) Write down the ordinary differential equation for the similarity solutions. (c) Can you find an explicit formula for the similarity solutions? *Hint*: Look at Exercise 8.2.14(a).

♡ 8.2.16. Besides the translations and scalings, Lie symmetry methods, [**87**], produce two other classes of symmetry transformations for the heat equation $u_t = u_{xx}$. Given that $u(t,x)$ is a solution to the heat equation:
(a) Prove that $U(t,x) = e^{c^2 t - c x}\,u(t, x - 2ct)$ is also a solution to the heat equation for any $c \in \mathbb{R}$. What solution do you obtain if $u(t,x) = a$ is a constant solution? *Remark*: This transformation can be interpreted as the effect of a *Galilean boost* to a coordinate frame that is moving with speed c.

(*b*) Prove that $U(t,x) = \dfrac{e^{-cx^2/(4(1+ct))}}{\sqrt{1+ct}}\, u\left(\dfrac{t}{1+ct}, \dfrac{x}{1+ct}\right)$ is a solution to the heat equation for any $c \in \mathbb{R}$. What solution do you obtain if $u(t,x) = a$ is a constant?

8.3 The Maximum Principle

We have already noted the temporal decay of temperature, as governed by the heat equation, to thermal equilibrium. While the temperature at any individual point in a physical medium can fluctuate — depending on what is happening elsewhere, thermodynamics tells us that the overall heat content of an isolated body must continually decrease. The *Maximum Principle* is the mathematical formulation of this physical law, and states that the temperature of a body cannot, in the absence of external heat sources, ever become larger than its initial or boundary values. This can be viewed as a dynamical counterpart to the Maximum Principle for the Laplace equation, as formulated in Theorem 4.9, stating that the maximum temperature of a body in equilibrium is achieved only on its boundary.

The proof of the Maximum Principle will be facilitated if we analyze the more general situation in which heat energy is being continually extracted throughout the body.

Theorem 8.6. *Let $\gamma > 0$. Suppose $u(t,x)$ is a solution to the forced heat equation*

$$\frac{\partial u}{\partial t} = \gamma\, \frac{\partial^2 u}{\partial x^2} + F(t,x) \tag{8.65}$$

on the rectangular domain

$$R = \{\, a < x < b,\ 0 < t < c \,\}.$$

Assume that the forcing term is nowhere positive: $F(t,x) \le 0$ *for all* $(t,x) \in R$. *Then the maximum of* $u(t,x)$ *on the closed rectangle* $\overline{R}$ *is attained at* $t = 0$ *or* $x = a$ *or* $x = b$.

In other words, if no new heat is being introduced, the maximum overall temperature occurs either at the initial time or on the body's boundary. In particular, in the fully insulated case $F(t,x) \equiv 0$, (8.65) reduces to the heat equation, and Theorem 8.6 applies as stated.

Proof: First let us first prove the result under the stronger assumption $F(t,x) < 0$, which implies that

$$\frac{\partial u}{\partial t} < \gamma\, \frac{\partial^2 u}{\partial x^2} \tag{8.66}$$

everywhere in the rectangle R. Suppose first that $u(t,x)$ has a (local) maximum at a point $(t^\star, x^\star)$ in the interior of R. Then, by multivariable calculus, [**8**, **108**], its gradient must vanish there, $\nabla u(t^\star, x^\star) = \mathbf{0}$, and hence

$$u_t(t^\star, x^\star) = u_x(t^\star, x^\star) = 0. \tag{8.67}$$

Our assumption implies that the scalar function $h(x) = u(t^\star, x)$ has a maximum at $x = x^\star$. Thus, by the second derivative test for functions of a single variable,

$$h''(x^\star) = u_{xx}(t^\star, x^\star) \le 0. \tag{8.68}$$

But the requirements (8.67–68) are clearly incompatible with the initial inequality (8.66). We conclude that the solution $u(t,x)$ cannot have a local maximum at any point in the interior of R.

We still need to exclude the possibility of a maximum occurring at a non-corner point $(t^\star, x^\star) = (c, x^\star)$, $a < x^\star < b$, on the right-hand edge of the rectangle. If such were to occur, then the function $g(t) = u(t, x^\star)$ would be nondecreasing at $t = c$, and hence $g'(t) = u_t(c, x^\star) \geq 0$ there. The preceding argument also implies that $u_{xx}(c, x^\star) \leq 0$, and again these two requirements are incompatible with (8.66). We conclude that any (local) maximum must occur on one of the other three sides of the rectangle, in accordance with the statement of the theorem.

To generalize the argument to the case $F(t,x) \leq 0$ — which includes the heat equation — requires a little trick. Starting with the solution $u(t,x)$ to (8.65), we set

$$v(t,x) = u(t,x) + \varepsilon\, x^2, \qquad \text{where} \qquad \varepsilon > 0.$$

Then,

$$\frac{\partial v}{\partial t} = \frac{\partial u}{\partial t} = \gamma\, \frac{\partial^2 u}{\partial x^2} + F(t,x) = \gamma\, \frac{\partial^2 v}{\partial x^2} - 2\gamma\, \varepsilon + F(t,x) = \gamma\, \frac{\partial^2 v}{\partial x^2} + \widetilde{F}(t,x),$$

where, by our original assumption on $F(t,x)$,

$$\widetilde{F}(t,x) = F(t,x) - 2\gamma\, \varepsilon < 0$$

everywhere in R. Thus, by the previous argument, a local maximum of $v(t,x)$ can occur only when $t = 0$ or $x = a$ or $x = b$. Now we let $\varepsilon \to 0$ and conclude the same for u. More rigorously, let M denote the maximum value of $u(t,x)$ on the indicated three sides of the rectangle. Then

$$v(t,x) \leq M + \varepsilon \max\{\, a^2, b^2 \,\}$$

there, and hence, by the preceding argument,

$$u(t,x) \leq v(t,x) \leq M + \varepsilon \max\{\, a^2, b^2 \,\} \qquad \text{for all} \qquad (t,x) \in R.$$

Now, letting $\varepsilon \to 0^+$ proves that $u(t,x) \leq M$ everywhere in R. *Q.E.D.*

For the unforced heat equation, we can bound the solution from both above and below by its boundary and initial temperatures:

Corollary 8.7. *Suppose $u(t,x)$ solves the heat equation $u_t = \gamma\, u_{xx}$, with $\gamma > 0$, for $a < x < b$, $0 < t < c$. Set*

$$B = \{\, (0,x) \mid\ a \leq x \leq b \,\} \cup \{\, (t,a) \mid\ 0 \leq t \leq c \,\} \cup \{\, (t,b) \mid\ 0 \leq t \leq c \,\},$$

and let

$$M = \max\{\, u(t,x) \mid\ (t,x) \in B \,\}, \qquad m = \min\{\, u(t,x) \mid\ (t,x) \in B \,\}, \tag{8.69}$$

be, respectively, the maximum and minimum values for the initial and boundary temperatures. Then $m \leq u(t,x) \leq M$ for all $a \leq x \leq b$, $0 \leq t \leq c$.

Proof: The upper bound $u(t,x) \leq M$ follows from the Maximum Principle of Theorem 8.6. To establish the lower bound, we note that $\widetilde{u}(t,x) = -\,u(t,x)$ also solves the heat equation, satisfying $\widetilde{u}(t,x) \leq -m$ on B, and hence, by the Maximum Principle, everywhere in the rectangle. But this implies $u(t,x) = -\,\widetilde{u}(t,x) \geq m$. *Q.E.D.*

Remark: Theorem 8.6 is sometimes referred to as the *Weak Maximum Principle* for the heat equation. The *Strong Maximum Principle* states that, provided the solution $u(t,x)$ is not constant, its value at any non-initial, non-boundary point $(t,x) \in \widehat{R} = \{\, a < x < b,\ 0 < t \leq c \,\}$ is *strictly* less than its maximum initial and boundary values; in other words, $u(t,x) < M$ for $(t,x) \in \widehat{R}$, where M is given in (8.69). Similarly, the Strong Maximum Principle implies that, for nonconstant solutions to the heat equation, the inequalities in Corollary 8.7 are strict: $m < u(t,x) < M$ for all $(t,x) \in \widehat{R}$. Proofs of the Strong Maximum Principle are more delicate, and can be found in [**38**, **61**].

One immediate application of the Maximum Principle is to prove uniqueness of solutions to the heat equation.

Theorem 8.8. *There is at most one solution to the Dirichlet initial-boundary value problem for the forced heat equation.*

Proof: Suppose u and $\widetilde{u}$ are any two solutions with the same initial and boundary values. Then their difference $v = u - \widetilde{u}$ solves the homogeneous initial-boundary value problem for the unforced heat equation, with minimum and maximum boundary values $m = 0 \leq v(t,x) \leq 0 = M$ for $t = 0$, $a \leq x \leq b$, and also $x = a$ or b, $0 \leq t \leq c$. But then Corollary 8.7 implies that $0 \leq v(t,x) \leq 0$ everywhere, which implies that $u \equiv \widetilde{u}$, thereby establishing uniqueness. *Q.E.D.*

Remark: Existence of the solution follows from the convergence of our Fourier series — assuming that the initial and boundary data and the forcing function are sufficiently nice.

Exercises

8.3.1. *True or false*: Assuming no external heat source, if the initial and boundary temperatures of a one-dimensional body are always positive, the temperature within the body is necessarily positive.

8.3.2. Suppose $u(t,x)$ and $v(t,x)$ are two solutions to the heat equation such that $u \leq v$ when $t = 0$ and when $x = a$ or $x = b$. Prove that $u(t,x) \leq v(t,x)$ for all $a \leq x \leq b$ and all $t \geq 0$. Provide a physical interpretion of this result.

8.3.3. For $t > 0$, let $u(t,x)$ be a solution to the unforced heat equation on an interval $a < x < b$, subject to homogeneous Dirichlet boundary conditions. Prove that $M(t) = \max\{\, u(t,x) \,|\, a \leq x \leq b \,\}$ is a nonincreasing function of t.

8.3.4. (a) State and prove a Maximum Principle for the *convection-diffusion equation* $u_t = u_{xx} + u_x$. (b) Does the equation $u_t = u_{xx} - u_x$ also admit a Maximum Principle?

8.3.5. Consider the parabolic equation $\dfrac{\partial u}{\partial t} = x\,\dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial u}{\partial x}$ on the interval $1 < x < 2$, with initial and boundary conditions $u(0,x) = f(x)$, $u(t,1) = \alpha(t)$, $u(t,2) = \beta(t)$.
(a) State and prove a version of the Maximum Principle for this problem.
(b) Establish uniqueness of the solution to this initial-boundary value problem.

8.3.6. (a) Show that $u(t,x) = -x^2 - 2xt$ is a solution to the diffusion equation $u_t = x\,u_{xx}$.
(b) Explain why this differential equation does not admit a Maximum Principle.

8.3.7. Suppose that $u(t,x)$ is a nonconstant solution to the heat equation on the interval $0 < x < \ell$, with homogeneous (*a*) Dirichlet, (*b*) Neumann, or (*c*) mixed boundary conditions. Prove that the function $E(t) = \int_0^\ell u(t,x)^2\,dx$ is everywhere decreasing: $E(t_1) > E(t_2)$ whenever $t_1 < t_2$.

8.3.8. *True or false*: The wave equation $u_{tt} = c^2 u_{xx}$ satisfies a Maximum Principle. If true, clearly state the principle; if false, explain why not.

8.4 Nonlinear Diffusion

First-order partial differential equations serve to model conservative wave motion, beginning with the basic one-dimensional scalar transport equations that we studied in Chapter 2, and progressing on to higher-dimensional systems, the equations of gas dynamics, the full-blown Euler equations of fluid mechanics, and yet more complicated systems of partial differential equations modeling plasmas, magneto-hydrodynamics, etc. However, such systems fail to account for frictional and viscous effects, which are typically modeled by parabolic diffusion equations such as the heat equation and its generalizations, both linear and nonlinear. In this section, we investigate the consequences of combining nonlinear wave motion with linear diffusion by analyzing the simplest such model. As we will see, the dissipative term has the effect of smoothing out abrupt shock discontinuities, and the result is a well-determined, smooth dynamical process with classical solutions. Moreover, in the inviscid limit, the smooth solutions converge (nonuniformly) to a discontinuous shock wave, leading to the method of viscosity solutions that has been successfully employed to analyze such nonlinear dynamical processes.

Burgers' Equation

The simplest nonlinear diffusion equation is known as[†] *Burgers' equation*

$$u_t + u\,u_x = \gamma\,u_{xx}, \tag{8.70}$$

which is obtained by appending a simple linear diffusion term to the nonlinear transport equation (2.31). As with the heat equation, the diffusion coefficient $\gamma \geq 0$ must be nonnegative in order that the initial value problem be well-posed in forwards time. In fluid and gas dynamics, one interprets the right-hand side as modeling the effect of viscosity, and so Burgers' equation represents a very simplified version of the equations of viscous fluid flows, including the celebrated and widely applied Navier–Stokes equations (1.4), [**122**]. When the viscosity coefficient vanishes, $\gamma = 0$, Burgers' equation reduces to the nonlinear transport equation (2.31), which, as a consequence, is often referred to as the *inviscid Burgers' equation*.

[†] The equation is named after the Dutch physicist Johannes Martinus Burgers, [**26**], and so the apostrophe goes after the "s". Burgers' equation was apparently first studied as a physical model by the British (later American) applied mathematician Harry Bateman, [**13**], in the early twentieth century.

Since Burgers' equation is of first order in t, we expect that its solutions will be uniquely prescribed by their initial values

$$u(0,x) = f(x), \qquad -\infty < x < \infty. \tag{8.71}$$

(For simplicity, we will ignore boundary effects here.) Small, slowly varying solutions — more specifically, those for which both $|\, u(t,x)\,|$ and $|\, u_x(t,x)\,|$ are small — tend to act like solutions to the heat equation, smoothing out and decaying to 0 as time progresses. On the other hand, when the solution is large or rapidly varying, the nonlinear term tends to play the dominant role, and we might expect the solution to behave like nonlinear transport waves, perhaps steepening into some sort of shock. But, as we will learn, the smoothing effect of the diffusion term, no matter how small, ultimately prevents the appearance of a discontinuous shock wave. Indeed, it can be proved that, under rather mild assumptions on the initial data, the solution to the initial value problem (8.70–71) remains smooth and well defined for all subsequent times, [**122**].

The simplest explicit solutions are the *traveling waves*, for which

$$u(t,x) = v(\xi) = v(x - ct), \qquad \text{where} \qquad \xi = x - ct, \tag{8.72}$$

indicates a fixed profile, moving to the right with constant speed c. By the chain rule,

$$\frac{\partial u}{\partial t} = -c\, v'(\xi), \qquad \frac{\partial u}{\partial x} = v'(\xi), \qquad \frac{\partial^2 u}{\partial x^2} = v''(\xi).$$

Substituting these expressions into Burgers' equation (8.70), we conclude that $v(\xi)$ must satisfy the nonlinear second-order ordinary differential equation

$$-c\, v' + v\, v' = \gamma\, v''.$$

This equation can be solved by first integrating both sides with respect to ξ, and so

$$\gamma\, v' = k - c\, v + \tfrac{1}{2}\, v^2,$$

where k is a constant of integration. Following the analysis after Proposition 2.3, as $\xi \to \pm\infty$, the bounded solutions to such an autonomous first-order ordinary differential equation tend to one of the fixed points provided by the roots of the quadratic polynomial on the right-hand side. Therefore, for there to be a *bounded* traveling-wave solution $v(\xi)$, the quadratic polynomial must have two real roots, which requires $k < \frac{1}{2}\, c^2$. Assuming this holds, we rewrite the equation in the form

$$2\gamma\, \frac{dv}{d\xi} = (v-a)(v-b), \qquad \text{where} \qquad c = \tfrac{1}{2}\,(a+b), \qquad k = \tfrac{1}{2}\, a\, b. \tag{8.73}$$

To obtain bounded solutions, we must require $a < v < b$. Integrating (8.73) by the usual method, cf. (2.19), we find

$$\int \frac{2\gamma\, dv}{(v-a)(v-b)} \;=\; \frac{2\gamma}{b-a}\, \log\left(\frac{b-v}{v-a}\right) \;=\; \xi - \delta,$$

where δ is another constant of integration. Solving for

$$v(\xi) = \frac{a\, e^{(b-a)(\xi-\delta)/(2\gamma)} + b}{e^{(b-a)(\xi-\delta)/(2\gamma)} + 1},$$

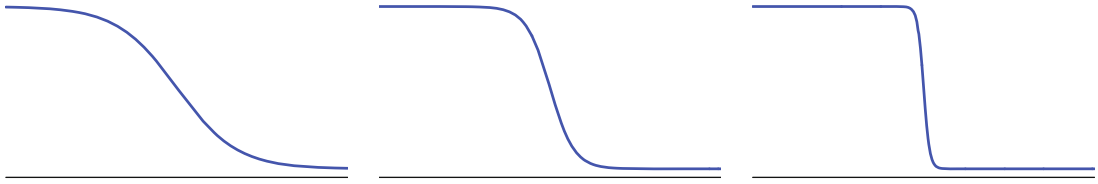


Figure 8.5. Traveling-wave solutions to Burgers' equation.

and recalling (8.73), we conclude that the bounded traveling-wave solutions to Burgers' equation all have the explicit form

$$u(t,x) = \frac{a\, e^{(b-a)(x-ct-\delta)/(2\gamma)} + b}{e^{(b-a)(x-ct-\delta)/(2\gamma)} + 1}\,, \tag{8.74}$$

where $a < b$ and δ are arbitrary constants. Observe that our solution is a monotonically decreasing function of x, with asymptotic values

$$\lim_{x \to -\infty} u(t,x) = b, \qquad \lim_{x \to \infty} u(t,x) = a,$$

at large distances. The wave travels to the right, unchanged in form, with speed $c = \frac{1}{2}(a+b)$ equal to the average of its asymptotic values. In particular, if $a = -b$, the result is a stationary-wave solution. In Figure 8.5 we graph sample profiles, corresponding to $a = .1$, $b = 1$, for three different values of the diffusion coefficient. Note that the smaller γ is, the sharper the transition layer between the two asymptotic values of the solution.

In the *inviscid limit* as the diffusion becomes vanishingly small, $\gamma \to 0$, the traveling-wave solutions (8.74) converge to the step shock-wave solutions (2.51) of the nonlinear transport equation. Indeed, this can be proved to hold in general: as $\gamma \to 0$, solutions to Burgers' equation (8.70) converge to the corresponding solutions to the nonlinear transport equation (2.31) that are subject to the Rankine–Hugoniot and entropy conditions (2.53, 55). Thus, the method of vanishing viscosity allows one to monitor solutions to the nonlinear transport equation as they evolve into regimes where multiple shocks interact and merge. This approach also reconfirms our physical intuition, in that most physical systems retain a very small dissipative component that serves to mollify abrupt discontinuities that might appear in a theoretical model that fails to take friction or viscous effects into account. In the modern theory of partial differential equations, the resulting *viscosity solution method* has been successfully used to characterize the discontinuous solutions to a broad range of inviscid nonlinear wave equations as limits of classical solutions to a viscously regularized system. We refer the interested reader to [**64**, **107**, **122**] for further details.

The Hopf–Cole Transformation

By a remarkable stroke of good fortune, the nonlinear Burgers' equation can be converted into the linear heat equation and thereby explicitly solved. The transformation that *linearizes* the nonlinear Burgers' equation first appeared in an obscure exercise in a nineteenth-century differential equations textbook, [**41**; vol. 6, p. 102]. Its rediscovery by

the applied mathematicians Eberhard Hopf, [**56**], and Julian Cole, [**32**], was a milestone in the modern era of nonlinear partial differential equations, and it is now named the Hopf–Cole transformation in their honor.

In general, *linearization* — that is, converting a given nonlinear differential equation into a linear equation — is extremely challenging, and, in most instances, impossible. On the other hand, the reverse process — "nonlinearizing" a linear equation — is trivial: any nonlinear change of dependent variables will do the trick! However, the resulting nonlinear equation, while evidently linearizable by inverting the change of variables, is rarely of independent interest. But sometimes there is a lucky accident, and the resulting linearization of a physically relevant nonlinear differential equation can have a profound impact on our understanding of more complicated nonlinear systems.

In the present context, our starting point is the linear heat equation

$$v_t = \gamma\, v_{xx}. \tag{8.75}$$

Among all possible nonlinear changes of dependent variable, one of the simplest that might spring to mind is an exponential function. Let us, therefore, investigate the effect of an exponential change of variables

$$v(t,x) = e^{\alpha\,\varphi(t,x)}, \qquad \text{so} \qquad \varphi(t,x) = \frac{1}{\alpha}\,\log v(t,x), \tag{8.76}$$

where α is a nonzero constant. The function $\varphi(t,x)$ is real, provided $v(t,x)$ is a *positive* solution to the heat equation. Fortunately, this is not hard to arrange: if the initial data $v(0,x) > 0$ is strictly positive, then, as a consequence of the Maximum Principle in Corollary 8.7, the resulting solution $v(t,x) > 0$ is positive for all $t > 0$.

To determine the differential equation satisfied by the function φ, we invoke the chain and product rules to differentiate (8.76):

$$v_t = \alpha\,\varphi_t\, e^{\alpha\,\varphi}, \qquad v_x = \alpha\,\varphi_x\, e^{\alpha\,\varphi}, \qquad v_{xx} = (\alpha\,\varphi_{xx} + \alpha^2\,\varphi_x^2)\, e^{\alpha\,\varphi}.$$

Substituting the first and last formulas into the heat equation (8.75) and canceling a common exponential factor, we conclude that $\varphi(t,x)$ satisfies the nonlinear partial differential equation

$$\varphi_t = \gamma\,\varphi_{xx} + \gamma\,\alpha\,\varphi_x^2, \tag{8.77}$$

known as the *potential Burgers' equation*, for reasons that will soon become apparent.

The second step in the process is to differentiate the potential Burgers' equation with respect to x; the result is

$$\varphi_{tx} = \gamma\,\varphi_{xxx} + 2\,\gamma\,\alpha\,\varphi_x\,\varphi_{xx}. \tag{8.78}$$

If we now set

$$\frac{\partial \varphi}{\partial x} = u, \tag{8.79}$$

so that φ acquires the status of a *potential function*, then the resulting partial differential equation

$$u_t = \gamma\, u_{xx} + 2\,\gamma\,\alpha\, u\, u_x$$

coincides with Burgers' equation (8.70) when $\alpha = -1/(2\gamma)$. In this manner, we have arrived at the famous *Hopf–Cole transformation*.

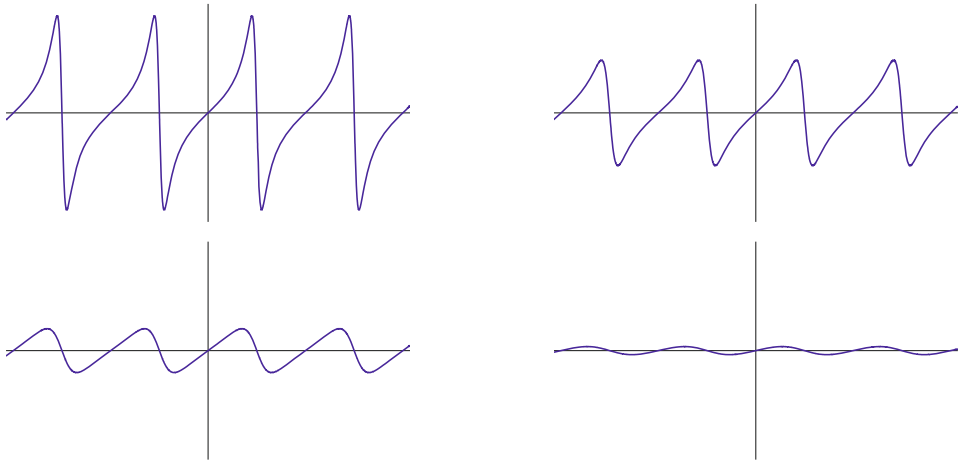


Figure 8.6. Trignometric solution to Burgers' equation.

Theorem 8.9. *If $v(t,x) > 0$ is any positive solution to the linear heat equation $v_t = \gamma\, v_{xx}$, then*

$$u(t,x) = \frac{\partial}{\partial x}\big[\, -2\gamma \log v(t,x)\,\big] = -2\gamma\,\frac{v_x}{v} \tag{8.80}$$

solves Burgers' equation $u_t + u\,u_x = \gamma\, u_{xx}$.

Do all solutions to Burgers' equation arise in this way? In order to answer this question, we run the argument in reverse. First, choose a potential function $\widetilde{\varphi}(t,x)$ that satisfies (8.79); for example,

$$\widetilde{\varphi}(t,x) = \int_0^x u(t,y)\,dy.$$

If $u(t,x)$ is any solution to Burgers' equation, then $\widetilde{\varphi}(t,x)$ satisfies (8.78). Integrating both sides of the latter equation with respect to x, we conclude that

$$\widetilde{\varphi}_t = \gamma\,\widetilde{\varphi}_{xx} + \gamma\,\alpha\,\widetilde{\varphi}_x^2 + g(t),$$

for some integration "constant" $g(t)$. Thus, unless $g(t) \equiv 0$, our potential function $\widetilde{\varphi}$ doesn't satisfy the potential Burgers' equation (8.77), but that is because we chose the "wrong" potential. Indeed, if we define

$$\varphi(t,x) = \widetilde{\varphi}(t,x) - G(t), \qquad \text{where} \qquad G'(t) = g(t),$$

then

$$\varphi_t = \widetilde{\varphi}_t - g(t) = \gamma\,\widetilde{\varphi}_{xx} + \gamma\,\alpha\,\widetilde{\varphi}_x^2 = \gamma\,\varphi_{xx} + \gamma\,\alpha\,\varphi_x^2,$$

and hence the modified potential $\varphi(t,x)$ *is* a solution to the potential Burgers' equation (8.77). From this it easily follows that

$$v(t,x) = e^{-\varphi(t,x)/(2\gamma)} \tag{8.81}$$

is a positive solution to the heat equation, from which the Burgers' solution $u(t,x)$ can be recovered through (8.80). We conclude that *every* solution to Burgers' equation comes from a positive solution to the heat equation via the Hopf–Cole transformation.

Example 8.10. As a simple example, the separable solution

$$v(t,x) = a + b\, e^{-\gamma \omega^2 t} \cos \omega x$$

to the heat equation leads to the following solution to Burgers' equation:

$$u(t,x) = \frac{2\gamma\, b\, \omega\, \sin \omega x}{a\, e^{\gamma \omega^2 t} + b \cos \omega x}\,. \tag{8.82}$$

A representative example is plotted in Figure 8.6. We should require that $a > |\, b\,|$ in order that $v(t,x) > 0$ be a positive solution to the heat equation for $t \geq 0$; otherwise the resulting solution to Burgers' equation will have singularities at the roots of u — as in the first graph in Figure 8.6. This family of solutions is primarily affected by the viscosity term, and rapidly decays to zero.

To solve the initial value problem (8.70–71) for Burgers' equation, we note that, under the Hopf–Cole transformation (8.80),

$$v(0,x) = \exp\left(-\frac{\varphi(0,x)}{2\gamma}\right) = \exp\left(-\frac{1}{2\gamma}\int_0^x f(y)\,dy\right) \equiv h(x). \tag{8.83}$$

Remark: The lower limit of the integral can be changed from 0 to any other convenient value. The only effect is to multiply $v(t,x)$ by an overall constant, which does not change the final form of $u(t,x)$ in (8.80).

According to formula (8.16) (adapted to general diffusivity, as in Exercise 8.2.3), the solution to the initial value problem (8.75, 83) for the heat equation can be expressed as a convolution integral with the fundamental solution

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

Therefore, setting $\widehat{v}(t,x) = 2\sqrt{\pi\gamma\, t}\, v(t,x)$, the solution to the Burgers' initial value problem (8.70–71), valid for $t > 0$, is given by

$$u(t,x) = -\frac{2\gamma}{\widehat{v}(t,x)}\,\frac{\partial \widehat{v}}{\partial x}\,, \qquad \text{where} \qquad \begin{cases} \widehat{v}(t,x) = \displaystyle\int_{-\infty}^{\infty} e^{-H(t,x;\xi)}\, d\xi, \\[2ex] H(t,x;\xi) = \dfrac{(x-\xi)^2}{4\gamma\, t} + \dfrac{1}{2\gamma}\displaystyle\int_0^{\xi} f(\eta)\, d\eta. \end{cases} \tag{8.84}$$

Example 8.11. To demonstrate the smoothing effect of the diffusion terms, let us see what happens to the initial data

$$u(0,x) = \begin{cases} a, & x < 0, \\ b, & x > 0, \end{cases} \tag{8.85}$$

in the form of a step function. We assume that $a > b$, which corresponds to a shock wave in the inviscid limit $\gamma = 0$. (In Exercise 8.4.4, the reader is asked to analyze the case $a < b$, which corresponds to a rarefaction wave.) In this case,

$$H(t,x;\xi) = \frac{(x-\xi)^2}{4\gamma\, t} + \begin{cases} \dfrac{a\xi}{2\gamma}\,, & \xi < 0, \\[2ex] \dfrac{b\xi}{2\gamma}\,, & \xi > 0. \end{cases} \tag{8.86}$$

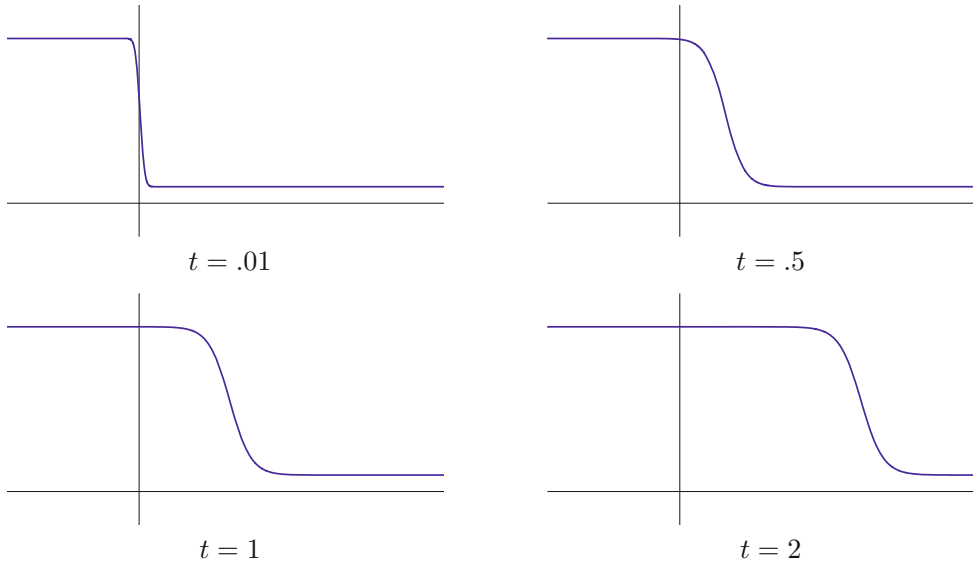


Figure 8.7. Shock-wave solution to Burgers' equation.

After some algebraic manipulations, the solution (8.84) is found to have the explicit form

$$u(t,x) = a \,+\, \frac{b-a}{1+\exp\left(\dfrac{b-a}{2\gamma}\,(x-ct)\right)\operatorname{erfc}\left(\dfrac{x-at}{2\sqrt{\gamma\, t}}\right)\Big/\operatorname{erfc}\left(\dfrac{bt-x}{2\sqrt{\gamma\, t}}\right)}, \tag{8.87}$$

with $c = \frac{1}{2}(a+b)$, where $\operatorname{erfc} z = 1 - \operatorname{erf} z$ denotes the complementary error function (8.43). The solution, for $a = 1$, $b = .1$, and $\gamma = .03$, is plotted at various times in Figure 8.7. Observe that, as with the heat equation, the jump discontinuity is immediately smoothed out, and the solution soon assumes the form of a smoothly varying transition between its two original heights. The larger the diffusion coefficient in relation to the jump magnitude, the more pronounced the smoothing effect. Moreover, as $\gamma \to 0$, the solution $u(t,x)$ converges to the shock-wave solution (2.51) to the transport equation, in which the speed of the shock is c, the average of the step heights — in accordance with the Rankine–Hugoniot shock rule. Indeed, in view of (2.88),

$$\lim_{z \to \infty} \operatorname{erfc} z = 0, \qquad \lim_{z \to -\infty} \operatorname{erfc} z = 2. \tag{8.88}$$

Thus, for $t > 0$, as $\gamma \to 0$, the ratio of the two complementary error functions in (8.87) tends to ∞ when $x < bt$, to 1 when $bt < x < at$, and to 0 when $x > at$. On the other hand, since $a > b$, the exponential term tends to ∞ when $x < ct$, and to 0 when $x > ct$. Put together, these imply that the solution $u(t,x) \to a$ when $x < ct$, while $u(t,x) \to b$, when $x > ct$, thus proving convergence to the shock-wave solution.

Example 8.12. Consider the case in which the initial data $u(0,x) = \delta(x)$ is a concentrated delta function impulse at the origin. In the solution formula (8.84), starting the integral for $H(t,x;\xi)$ at 0 is problematic, but as noted earlier, we are free to select any

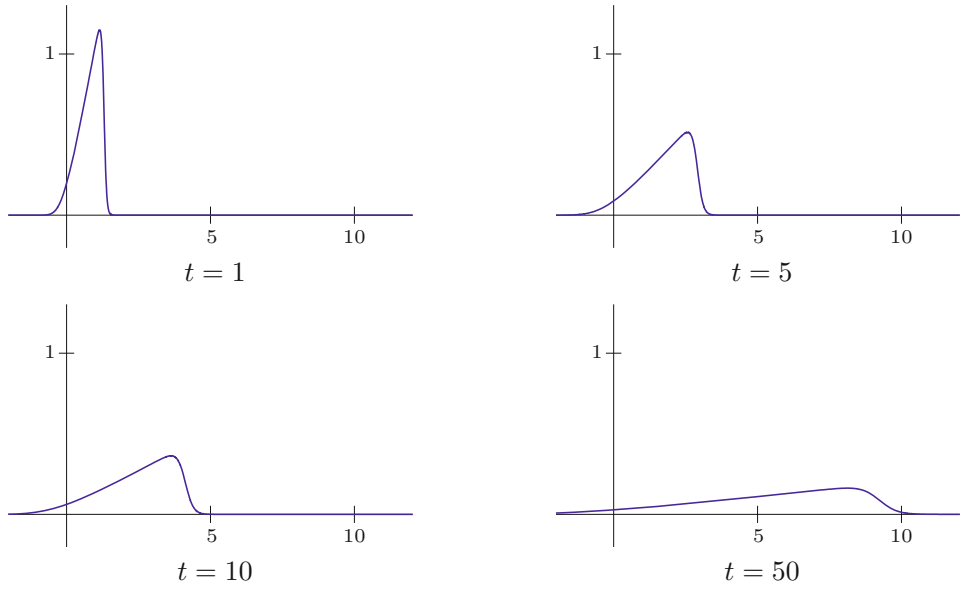


Figure 8.8. Triangular-wave solution to Burgers' equation.

other starting point, e.g., $-\infty$. Thus, we take

$$H(t,x;\xi) = \frac{(x-\xi)^2}{4\gamma t} + \frac{1}{2\gamma}\int_{-\infty}^{\xi} \delta(\eta)\, d\eta = \begin{cases} \dfrac{(x-\xi)^2}{4\gamma t}, & \xi < 0, \\ \dfrac{1}{2\gamma} + \dfrac{(x-\xi)^2}{4\gamma t}, & \xi > 0. \end{cases}$$

We then evaluate

$$\widehat{v}(t,x) = \int_{-\infty}^{\infty} e^{-H(t,x;\xi)}\, d\xi = \sqrt{\pi\gamma t}\left[1 - \operatorname{erf}\left(\frac{x}{2\sqrt{\gamma t}}\right) + e^{-1/(2\gamma)}\left\{1 + \operatorname{erf}\left(\frac{x}{2\sqrt{\gamma t}}\right)\right\}\right].$$

Therefore, the solution to the initial value problem is

$$u(t,x) = -\frac{2\gamma}{\widehat{v}(t,x)}\frac{\partial \widehat{v}}{\partial x} = 2\sqrt{\frac{\gamma}{\pi t}}\,\frac{e^{-x^2/(4\gamma t)}}{\coth\left(\dfrac{1}{4\gamma}\right) - \operatorname{erf}\left(\dfrac{x}{2\sqrt{\gamma t}}\right)}, \tag{8.89}$$

where

$$\coth z = \frac{\cosh z}{\sinh z} = \frac{e^z + e^{-z}}{e^z - e^{-z}} = \frac{e^{2z}+1}{e^{2z}-1}$$

is the hyperbolic cotangent function. A graph of this solution when $\gamma = .02$ and $a = 1$ appears in Figure 8.8. As you can see, the initial concentration diffuses out, but, in contrast to the heat equation, does not remain symmetric, since the nonlinear advection term causes the wave to steepen in front. Eventually, as the effect of the diffusion accumulates, the propagating triangular wave becomes vanishingly small.

Exercises

8.4.1. Find the solution to Burgers' equation that has the following initial data:

$$u(0,x) = \quad (a)\ \sigma(x), \qquad (b)\ \sigma(-x), \qquad (c)\ \begin{cases} 1, & 0 < x < 1, \\ 0, & \text{otherwise.} \end{cases}$$

8.4.2. Starting with the heat equation solution $v(t,x) = 1 + t^{-1/2}\, e^{-x^2/(4\gamma t)}$, find the corresponding solution to Burgers' equation and discuss its behavior.

8.4.3. Justify the solution formula (8.87).

♢ 8.4.4. (a) Prove that $\lim_{z \to \infty} z e^{z^2} \operatorname{erfc} z = 1/\sqrt{\pi}$. (b) Show that when $a < b$, the Burgers' solution (8.87) converges to the rarefaction wave (2.54) in the inviscid limit $\gamma \to 0^+$.

8.4.5. *True or false*: If $u(t,x)$ solves Burgers' equation for the step function initial condition $u(0,x) = \sigma(x)$, then $v(t,x) = u_x(t,x)$ solves the initial value problem with $v(0,x) = \delta(x)$.

8.4.6. *True or false*: If $\widehat{v}(t,x)$ is as given in (8.84), then

$$\frac{\partial \widehat{v}}{\partial x} = \int_{-\infty}^{\infty} \frac{\xi - x}{2\gamma t}\, e^{-H(t,x;\xi)}\, d\xi,$$

and hence the solution to the Burgers' initial value problem (8.70–71) can be written as

$$u(t,x) = \frac{\displaystyle\int_{-\infty}^{\infty} \frac{x-\xi}{t}\, e^{-H(t,x;\xi)}\, d\xi}{\displaystyle\int_{-\infty}^{\infty} e^{-H(t,x;\xi)}\, d\xi}, \qquad \text{where} \qquad H(t,x;\xi) = \frac{(x-\xi)^2}{4\gamma t} + \frac{1}{2\gamma}\int_0^{\xi} f(\eta)\, d\eta.$$

8.4.7. Show that if $u(t,x)$ solves Burgers' equation, then $U(t,x) = u(t, x - ct) + c$ is also a solution. What is the physical interpretation of this symmetry?

8.4.8. (a) What is the effect of a scaling transformation $(t,x,u) \longmapsto (\alpha t, \beta x, \lambda u)$ on Burgers' equation? (b) Use your result to solve the initial value problem for the rescaled Burgers' equation $U_t + \rho U U_x = \sigma U_{xx}$, $U(0,x) = F(x)$.

♡ 8.4.9. (a) Find all scaling symmetries of Burgers' equation. (b) Determine the ordinary differential equation satisfied by the similarity solutions. (c) *True or false*: The Hopf–Cole transformation maps similarity solutions of the heat equation to similarity solutions of Burgers' equation.

8.4.10. What happens if you nonlinearize the heat equation (8.75) using the change of variables

$$(a)\ v = \varphi^2; \qquad (b)\ v = \sqrt{\varphi}\,; \qquad (c)\ v = \log \varphi\,?$$

8.4.11. What partial differential equation results from applying the exponential change of variables (8.76) to:

(a) the wave equation $v_{tt} = c^2 v_{xx}$? (b) the Laplace equation $v_{xx} + v_{yy} = 0$?

8.5 Dispersion and Solitons

In this section, we finally venture beyond the by now familiar terrain of second-order partial differential equations. While considerably less common than those of first and second order, higher-order equations arise in certain applications, particularly third-order

dispersive models for wave motion, [**2**, **122**], and fourth-order systems modeling elastic plates and shells, [**7**]. We will focus our attention on two basic third-order evolution equations. The first is a simple linear equation with a third derivative term. It arises as a simplified model for unidirectional wave motion, and thus has more in common with first-order transport equations than with the second-order dissipative heat equation. The third-order derivative induces a process of *dispersion*, in which waves of different frequencies propagate at different speeds. Thus, unlike the first- and second-order wave equations, in which waves maintain their initial profile as they move, dispersive waves will spread out and decay even while conserving energy. Waves on the surface of a liquid are familiar examples of dispersive waves — an initially concentrated disturbance, caused by, say, throwing a rock in a pond, spreads out over the surface as its different vibrational components move off at different speeds.

Our second example is a remarkable nonlinear third-order evolution equation known as the Korteweg–de Vries equation, which combines dispersive effects with nonlinear transport. As with Burgers' equation (but for very different mathematical reasons), the dispersive term thwarts the tendency for solutions to break into shock waves, and, in fact, classical solutions exist for all time. Moreover, a general localized initial disturbance will break up into a finite number of solitary waves; the taller the wave, the faster it moves. Even more remarkable are the interactive properties of these solitary waves. One ordinarily expects nonlinearity to induce very complicated and not easily predictable behavior. However, when two solitary-wave solutions to the Korteweg–de Vries equation collide, they eventually emerge from the interaction unchanged, save for a phase shift. This unexpected and remarkable phenomenon was first detected through numerical simulations in the 1960s and distinguished with the neologism *soliton*. It was then found that solitons appear in a surprising number of basic nonlinear physical models. The investigation of their mathematical properties has had deep ramifications, not just within partial differential equations and fluid mechanics, but throughout applied mathematics and theoretical physics; it has even contributed to the solution of long-outstanding problems in complex function theory. Further development of the modern theory and amazing properties of integrable soliton equations can be found in [**2**, **36**].

Linear Dispersion

The simplest nontrivial third-order partial differential equation is the linear equation

$$u_t + u_{xxx} = 0, \tag{8.90}$$

which models the unidirectional† propagation of linear dispersive waves. To avoid complications engendered by boundary conditions, we shall initially look only at solutions on the entire line, so $-\infty < x < \infty$. Since the equation involves only a first-order time derivative, one expects its solutions to be uniquely specified by a single initial condition

$$u(0, x) = f(x), \qquad -\infty < x < \infty. \tag{8.91}$$

† Bidirectional propagation, as we saw in the wave equation, requires a second-order time derivative. As in the d'Alembert solution to the second-order wave equation, the reduction to a unidirectional model is based on an (approximate) factorization of the bidirectional operator.

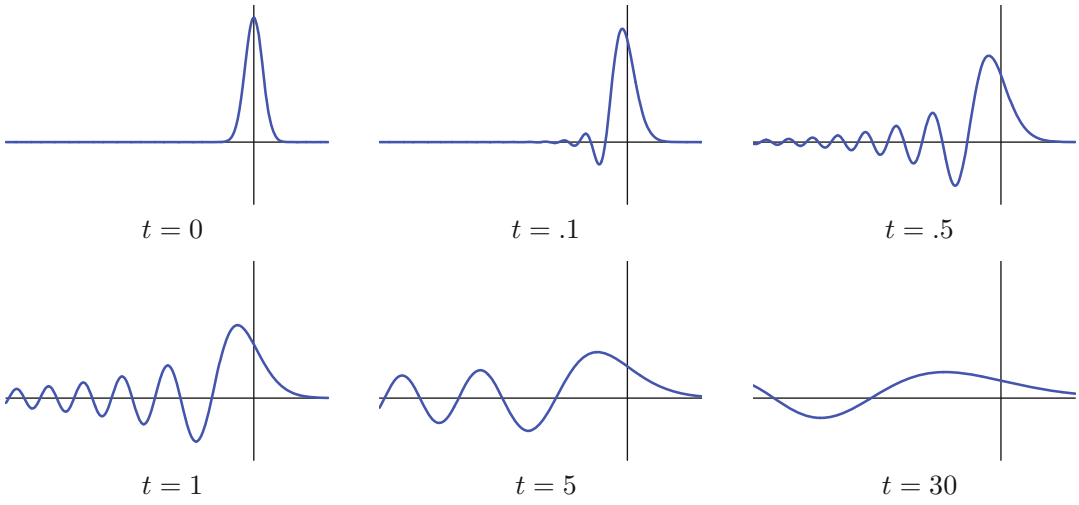


Figure 8.9. Gaussian solution to the dispersive wave equation.

In wave mechanics, $u(t,x)$ represents the height of the fluid at time t and position x, and the initial condition (8.91) specifies the initial disturbance.

As with the heat equation (and, indeed, any linear constant-coefficient evolution equation), the Fourier transform is an effective tool for solving the initial value problem on the real line. Assuming that the solution $u(t,\cdot) \in \mathrm{L}^2(\mathbb{R})$ remains square integrable at all times t (a fact that can be justified a priori — see Exercise 8.5.18(b)), let

$$\widehat{u}(t,k) = \frac{1}{\sqrt{2\,\pi}} \int_{-\infty}^{\infty} u(t,x)\, e^{-\mathrm{i}\,k\,x}\, dx$$

be its spatial Fourier transform. Owing to its effect on derivatives, the Fourier transform converts the partial differential equation (8.90) into a first-order linear ordinary differential equation:

$$\frac{\partial\, \widehat{u}}{\partial t} + (\mathrm{i}\,k)^3\, \widehat{u} = \frac{\partial\, \widehat{u}}{\partial t} - \mathrm{i}\,k^3\, \widehat{u} = 0, \tag{8.92}$$

in which the frequency variable k appears as a parameter. The corresponding initial conditions

$$\widehat{u}(0,k) = \widehat{f}(k) = \frac{1}{\sqrt{2\,\pi}} \int_{-\infty}^{\infty} f(x)\, e^{-\mathrm{i}\,k\,x}\, dx \tag{8.93}$$

are provided by the Fourier transform of (8.91). The solution to the initial value problem (8.92–93) is

$$\widehat{u}(t,k) = \widehat{f}(k)\, e^{\,\mathrm{i}\,k^3\,t}.$$

Inverting the Fourier transform yields the explicit formula for the solution

$$u(t,x) = \frac{1}{\sqrt{2\,\pi}} \int_{-\infty}^{\infty} \widehat{f}(k)\, e^{\,\mathrm{i}\,(k\,x + k^3\,t)}\, dk \tag{8.94}$$

to the initial value problem (8.90–91) for the dispersive wave equation.

Example 8.13. Suppose that the initial profile

$$u(0,x) = f(x) = e^{-x^2}$$

is a Gaussian. According to our table of Fourier transforms (see page 272),

$$\widehat{f}(k) = \frac{e^{-k^2/4}}{\sqrt{2}},$$

and hence the corresponding solution to the dispersive wave equation (8.90) is

$$u(t,x) = \frac{1}{2\sqrt{\pi}} \int_{-\infty}^{\infty} e^{\mathrm{i}\,(k\,x+k^3\,t)-k^2/4}\,dk = \frac{1}{2\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-k^2/4}\cos(k\,x+k^3\,t)\,dk;$$

the imaginary part vanishes thanks to the oddness of the integrand. (Indeed, the solution must be real, since the initial data is real.) A plot of the solution at various times appears in Figure 8.9. Note the propagation of initially rapid oscillations to the rear (negative x) of the initial disturbance. The dispersion causes the oscillations to gradually spread out and decrease in amplitude, with the effect that $u(t,x) \to 0$ uniformly as $t \to \infty$, even though, according to Exercise 8.5.7, both the mass $M = \displaystyle\int_{-\infty}^{\infty} u(t,x)\,dx$ and the energy $E = \displaystyle\int_{-\infty}^{\infty} u(t,x)^2\,dx$ of the wave are conserved, i.e., are both constant in time.

Example 8.14. The *fundamental solution* to the dispersive wave equation is generated by a concentrated initial disturbance:

$$u(0,x) = \delta(x).$$

The Fourier transform of the delta function is just $\widehat{\delta}(k) = 1/\sqrt{2\,\pi}$. Therefore, the corresponding solution (8.94) is

$$u(t,x) = \frac{1}{2\,\pi} \int_{-\infty}^{\infty} e^{\mathrm{i}\,(k\,x+k^3\,t)}\,dk = \frac{1}{\pi} \int_{0}^{\infty} \cos(k\,x+k^3\,t)\,dk, \tag{8.95}$$

since the solution is real (or, equivalently, the imaginary part of the integrand is odd), while the real part of the integrand is even.

A priori, it appears that the integral (8.95) does not converge, because the integrand does not go to zero as $|\,k\,| \to \infty$. However, the increasingly rapid oscillations induced by the cubic term tend to cancel each other out and allow convergence. To prove this, given $l > 0$, we perform a (non-obvious) integration by parts:

$$\begin{aligned}
\int_0^l \cos(k\,x+k^3\,t)\,dk &= \int_0^l \frac{1}{x+3\,k^2\,t}\,\frac{d}{dk}\sin(k\,x+k^3\,t)\,dk \qquad (8.96)\\
&= \frac{\sin(k\,x+k^3\,t)}{x+3\,k^2\,t}\,\Big|_{k=0}^{l} - \int_0^l \frac{d}{dk}\left(\frac{1}{x+3\,k^2\,t}\right)\sin(k\,x+k^3\,t)\,dk\\
&= \frac{\sin(l\,x+l^3\,t)}{x+3\,l^2\,t} + \int_0^l \frac{6\,k\,t\sin(k\,x+k^3\,t)}{(x+3\,k^2\,t)^2}\,dk.
\end{aligned}$$

Provided $t \neq 0$, as $l \to \infty$, the first term on the right goes to zero, while the final integral converges absolutely due to the rapid decay of the integrand.

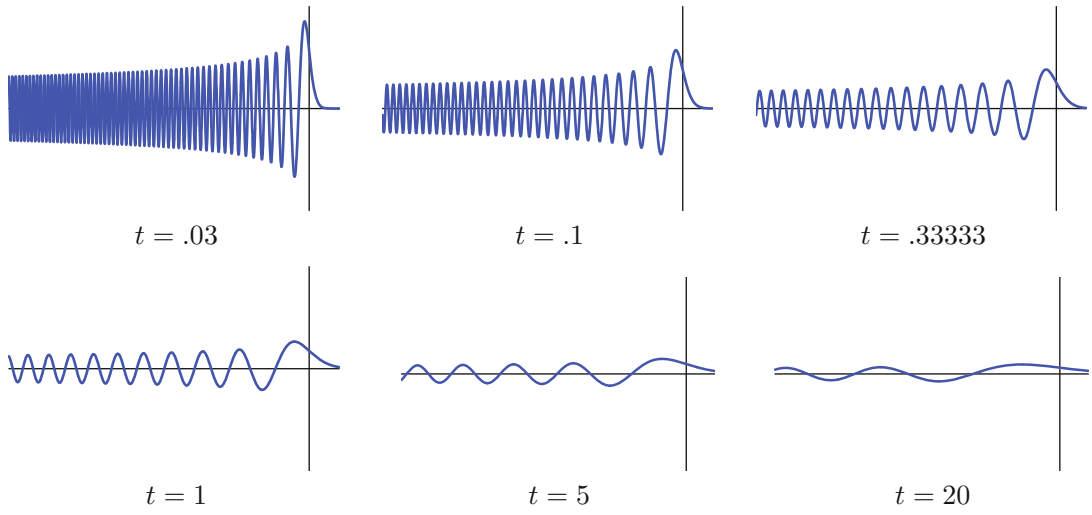


Figure 8.10. Fundamental solution to the dispersive wave equation.

While the integral in the solution formula (8.95) cannot be evaluated in terms of elementary functions, it is related to the integral defining the *Airy function*

$$\mathrm{Ai}(z) = \frac{1}{\pi} \int_0^\infty \cos\left(s\, z + \tfrac{1}{3}\, s^3\right) ds, \tag{8.97}$$

an important special function, [**86**], that was first employed by the nineteenth-century British applied mathematician George Airy in his studies of optical caustics (the focusing of light waves through a lens, e.g., a magnifying glass) and rainbows, [**4**]. Indeed, applying the change of variables

$$s = k\, \sqrt[3]{3\,t}\,, \qquad z = \frac{x}{\sqrt[3]{3\,t}}\,,$$

to the Airy function integral (8.97), we deduce that the fundamental solution to the dispersive wave equation (8.90) can be written as

$$u(t,x) = \frac{1}{\sqrt[3]{3\,t}}\ \mathrm{Ai}\left(\frac{x}{\sqrt[3]{3\,t}}\right). \tag{8.98}$$

See Figure 8.10 for a graph of the solution at several times; in particular, at $t = 1/3$ the solution is exactly the Airy function. We see that the immediate effect of the initial delta impluse is to spawn a highly oscillatory wave trailing off to $-\infty$. (As with the heat equation, signals propagate with infinite speed.) As time progresses, the dispersive effects cause the oscillations to spread out, with their overall amplitude decaying in proportion to $t^{-1/3}$. On the other hand, as $t \to 0^+$, the solution becomes more and more oscillatory for negative x, and so converges *weakly* to the initial delta function. We also note that (8.98) has the form of a similarity solution, since it is invariant under the scaling symmetry

$$(t, x, u) \longmapsto (\lambda^{-3} t, \lambda^{-1}\, x, \lambda\, u).$$

Equation (8.98) gives the response to an initial delta function concentrated at the

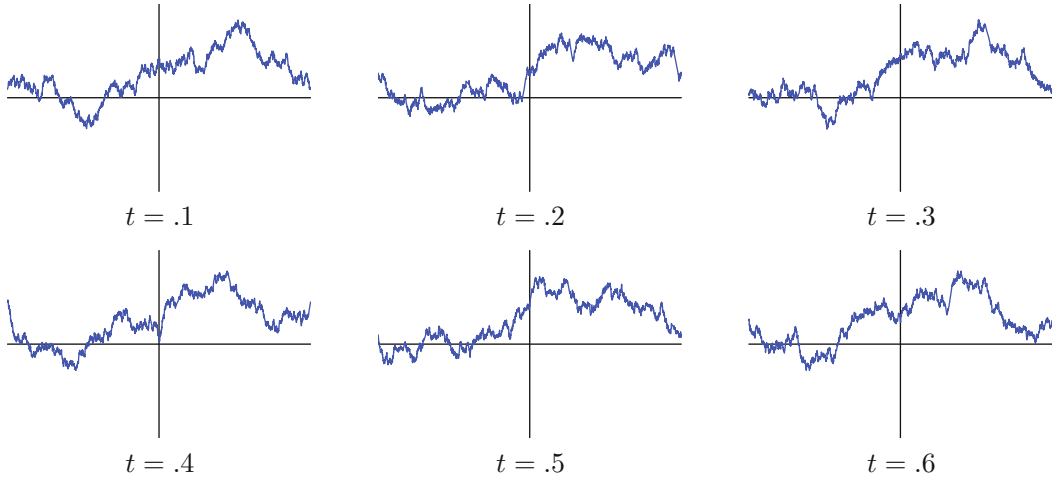


Figure 8.11. Periodic dispersion at irrational (with respect to π) times.

origin. By translation invariance, we immediately deduce that

$$F(t,x;\xi) = \frac{1}{\sqrt[3]{3t}}\ \mathrm{Ai}\left(\frac{x-\xi}{\sqrt[3]{3t}}\right)$$

is the *fundamental solution* corresponding to an initial delta impulse at $x = \xi$. Therefore, we can use linear superposition to find an explicit formula for the solution to the initial value problem that bypasses the Fourier transform. Namely, writing the general initial data as a superposition of delta functions,

$$u(0,x) = f(x) = \int_{-\infty}^{\infty} f(\xi)\,\delta(x-\xi)\,d\xi,$$

we conclude that the resulting solution is the selfsame combination of fundamental solutions:

$$u(t,x) = \frac{1}{\sqrt[3]{3t}}\int_{-\infty}^{\infty} f(\xi)\ \mathrm{Ai}\left(\frac{x-\xi}{\sqrt[3]{3t}}\right)d\xi. \tag{8.99}$$

Example 8.15. *Dispersive Quantization.* Let us investigate the periodic initial-boundary value problem for our basic linear dispersive equation on the interval $-\pi \le x \le \pi$:

$$u_t + u_{xxx} = 0, \quad u(t,-\pi) = u(t,\pi), \quad u_x(t,-\pi) = u_x(t,\pi), \quad u_{xx}(t,-\pi) = u_{xx}(t,\pi), \tag{8.100}$$

with initial data $u(0,x) = f(x)$. The Fourier series formula for the resulting solution is straightforwardly constructed:

$$u(t,x) = \sum_{k=-\infty}^{\infty} c_k\, e^{\,\mathrm{i}\,(k\,x + k^3 t)}, \tag{8.101}$$

where c_k are the usual (complex) Fourier coefficients (3.65) of the initial data $f(x)$.

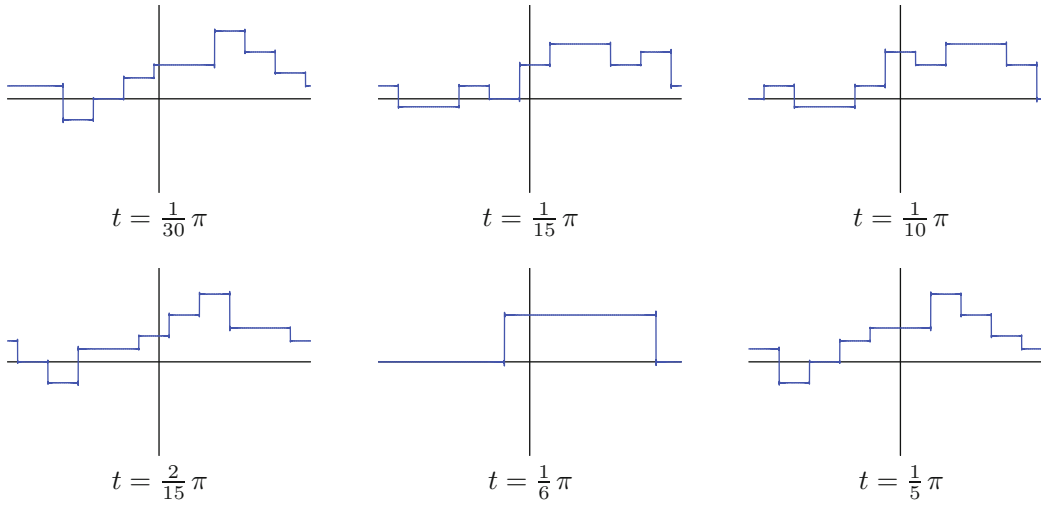


Figure 8.12. Periodic dispersion at rational (with respect to π) times.

Let us take the initial data to be the unit step function: $u(0, x) = \sigma(x)$. In view of its Fourier series (3.67), the resulting solution formula (8.101) becomes

$$\begin{aligned} u(t,x) &= \frac{1}{2} - \frac{\mathrm{i}}{\pi} \sum_{l=-\infty}^{\infty} \frac{e^{\,\mathrm{i}\,[(2l+1)x + (2l+1)^3 t]}}{2l+1} \\ &= \frac{1}{2} + \frac{2}{\pi} \sum_{l=0}^{\infty} \frac{\sin\big[\,(2l+1)\,x + (2l+1)^3 t\,\big]}{2l+1}\,. \end{aligned} \tag{8.102}$$

Let us graph this solution. At times uniformly spaced by $\Delta t = .1$, the resulting solution profiles are plotted in Figure 8.11. The solution appears to have a continuous but fractal-like structure, reminiscent of Weierstrass' continuous but nowhere differentiable function, [**55**; pp. 401–421]. The temporal evolution continues in this fashion until the initial data are formed again at $t = 2\pi$, after which the process periodically repeats.

However, when the times are spaced by $\Delta t = \frac{1}{30}\pi \approx .10472$, the resulting solution profiles, as plotted in Figure 8.12, are strikingly different! Indeed, as you are asked to prove in Exercise 8.5.8, at each rational time $t = 2\pi p/q$, where p, q are integers, the solution (8.102) to the initial-boundary value problem is discontinuous but constant on subintervals of length $2\pi/q$. This remarkable behavior, in which the solution profiles of linearly dispersive periodic boundary value problems have markedly different behaviors at rational and irrational times (with respect to π), was first observed, in the 1990's, in optics and quantum mechanics by the British physicist Michael Berry, [**16**, **115**], and named the *Talbot effect*, after an optical experiment conducted by the inventor of the photographic negative, William Henry Fox Talbot. While writing this book, I rediscovered the effect, which I like to call *dispersive quantization*, [**88**], and found that it arises in a wide range of linearly dispersive periodic initial-boundary value problems, [**30**].

The Dispersion Relation

As noted earlier, a key feature of the third-order wave equation (8.90) is that waves disperse, in the sense that those of different frequencies move at different speeds. Our goal now is to better understand the dispersion process. To this end, consider a solution whose initial profile

$$u(0,x) = e^{\,\mathrm{i}\,k\,x}$$

is a complex oscillatory function. Since the initial data does not decay as $|\,x\,| \to \infty$, we cannot use the Fourier integral solution formula (8.94) directly. Instead, anticipating the induced wave to exhibit temporal oscillations, let us try an exponential solution ansatz

$$u(t,x) = e^{\,\mathrm{i}\,(k\,x - \omega\,t)} \tag{8.103}$$

representing a complex oscillatory wave of temporal *frequency* ω and *wave number* (spatial frequency) k. Since

$$\frac{\partial u}{\partial t} = -\,\mathrm{i}\,\omega\,e^{\,\mathrm{i}\,(k\,x-\omega\,t)}, \qquad \frac{\partial^3 u}{\partial x^3} = -\,\mathrm{i}\,k^3\,e^{\,\mathrm{i}\,(k\,x-\omega\,t)},$$

(8.103) satisfies the partial differential equation (8.90) if and only if its frequency and wave number satisfy the *dispersion relation*

$$\omega = -k^3. \tag{8.104}$$

Therefore, the exponential solution (8.103) of wave number k takes the form

$$u(t,x) = e^{\,\mathrm{i}\,(k\,x+k^3\,t)}. \tag{8.105}$$

Our Fourier transform formula (8.94) for the solution can thus be viewed as a (continuous) linear superposition of these elementary exponential solutions. In general, to find the dispersion relation for a linear constant-coefficient partial differential equation, one substitutes the exponential ansatz (8.103). On cancellation of the common exponential factors, the result is an equation expressing the frequency ω as a function of the wave number k.

Any exponential solution (8.103) is automatically in the form of a traveling wave, since we can write

$$u(t,x) = e^{\,\mathrm{i}\,(k\,x-\omega\,t)} = e^{\,\mathrm{i}\,k\,(x-c_p\,t)}, \qquad \text{where} \qquad c_p = \frac{\omega}{k} \tag{8.106}$$

is the *wave speed* or, as it is more usually called, the *phase velocity*. If the dispersion relation is linear in the wave number, $\omega = c\,k$, as occurs in the linear transport equation $u_t + c\,u_x = 0$, then all waves move at an identical speed $c_p = c$, and hence localized disturbances stay localized as they propagate through the medium. In the dispersive case, ω is no longer a linear function of k, and so waves of different spatial frequencies move at different speeds. In the particular case (8.90), those with wave number k move at speed $c_p = \omega/k = -k^2$, and so the higher the wave number, the faster the wave propagates to the left. As the individual exponential constituents separate, the overall effect is the dispersive decay of an initially localized wave, with slowly diminishing amplitude and increasingly rapid oscillation as $x \to -\infty$.

The general solution to the linear partial differential equation under consideration is then built up by linear superposition of the exponential solutions,

$$u(t,x) = \int_{-\infty}^{\infty} e^{\,\mathrm{i}\,(k\,x-\omega\,t)}\,g(k)\,dk, \tag{8.107}$$

where $\omega = \omega(k)$ is determined by the relevant dispersion relation. While the evolution of the individual waves is an immediate consequence of the dispersion relation, the evolution of the localized wave packet represented by (8.107) is less evident. To determine its speed of propagation, let us switch to a moving coordinate frame of speed c by setting $x = c\,t + \xi$. The solution formula (8.107) then becomes

$$u(t, c\,t + \xi) = \int_{-\infty}^{\infty} e^{\,\mathrm{i}\,(c\,k - \omega)\,t} e^{\,\mathrm{i}\,k\,\xi}\, g(k)\, dk. \tag{8.108}$$

For a fixed value of ξ, the integral is of the general oscillatory form

$$H(t) = \int_{-\infty}^{\infty} e^{\,\mathrm{i}\,\varphi(k)\,t}\, h(k)\, dk, \tag{8.109}$$

where, in our case, $\varphi(k) = c\,k - \omega(k)$ and $h(k) = e^{\,\mathrm{i}\,k\,\xi}\, g(k)$. We are interested in understanding the behavior of such an oscillatory integral as $t \to \infty$. Now, if $\varphi(k) = k$, then (8.109) is just a Fourier integral, (7.9), and, as we learned in Chapter 7, $H(t) \to 0$ as $t \to \infty$, for any reasonable function $h(k)$. Intuitively, the increasingly rapid oscillations of the exponential factor tend to cancel each other out in the high-frequency limit. A similar result holds wherever $\varphi(k)$ has no stationary points, i.e., $\varphi'(k) \neq 0$, since one can then perform a local change of variables $\widetilde{k} = \varphi(k)$ to convert that part of the oscillatory integral to Fourier form, and again the increasingly rapid oscillations cause the limit to vanish. In this fashion, we arrive at the key insight of Stokes and Kelvin that produced the powerful *Method of Stationary Phase*. Namely, for large $t \gg 0$, the primary contribution to the highly oscillatory integral (8.109) occurs at the *stationary points* of the phase function, that is, where $\varphi'(k) = 0$. A rigorous justification of the method, along with precise error bounds, can be found in [**85**].

In the present context, the Method of Stationary Phase implies that the most significant contribution to the integral (8.108) occurs when

$$0 = \frac{d}{dk}\,(\omega - c\,k) = \frac{d\omega}{dk} - c. \tag{8.110}$$

Thus, surprisingly, the principal contribution of the components at wave number k is felt when moving at the *group velocity*

$$c_{\!g} = \frac{d\omega}{dk}\,. \tag{8.111}$$

Interestingly, unless the dispersion relation is linear in the wave number, the group velocity (8.111), which determines the speed of propagation of the energy, is *not the same* as the phase velocity (8.106), which governs the speed of propagation of an individual oscillatory wave. For example, in the case of the dispersive wave equation (8.90), $\omega = -k^3$, and so $c_{\!g} = -3\,k^2$, which is three times as fast as the phase velocity, $c_{\!p} = \omega/k = -k^2$. Thus, the energy propagates faster than the individual waves. This can be observed in Figure 8.9: while the bulk of the disturbance is spreading out rather rapidly to the left, the individual wave crests are moving slower.

On the other hand, the dispersion relation associated with deep water waves is (ignoring physical constants) $\omega = \sqrt{k}\,$, [**122**]. Now, the phase velocity is $c_{\!p} = \omega/k = 1/\sqrt{k}\,$, whereas the group velocity is $c_{\!g} = d\omega/dk = 1/(2\sqrt{k}\,) = \frac{1}{2}\,c_{\!p}$, and so the individual waves move twice as fast as the speed of propagation of the underlying wave energy. For an experimental verification, just throw a stone in a still pond. An individual wave crest emerges

in back and then steadily grows as it moves through the disturbance, eventually subsiding and disappearing into the still water ahead of the expanding wave packet triggered by the stone. The distinction between group velocity and phase velocity is also well understood by surfers, who know that the largest waves seen out to sea are not the largest when they break upon the shore.

Exercises

8.5.1. Sketch a picture of the solution for the initial value problem in Example 8.13 at times $t = -.1, -.5$, and -1.

♠ 8.5.2. (*a*) Write down an integral formula for the solution to the dispersive wave equation (8.90) with initial data $u(0,x) = \begin{cases} 1, & 0 < x < 1, \\ 0, & \text{otherwise.} \end{cases}$ (*b*) Use a computer package to plot your solution at several times and discuss what you observe.

8.5.3. (*a*) Write down an integral formula for the solution to the initial value problem

$$u_t + u_x + u_{xxx} = 0, \qquad u(0,x) = f(x).$$

(*b*) Based on the results in Example 8.13, discuss the behavior of the solution to the initial value problem $u(0,x) = e^{-x^2}$ as t increases.

8.5.4. Find the (*i*) dispersion relation, (*ii*) phase velocity, and (*iii*) group velocity for the following partial differential equations. Which are dispersive? (*a*) $u_t + u_x + u_{xxx} = 0$, (*b*) $u_t = u_{xxxxx}$, (*c*) $u_t + u_x - u_{xxt} = 0$, (*d*) $u_{tt} = c^2 u_{xx}$, (*e*) $u_{tt} = u_{xx} - u_{xxxx}$.

8.5.5. Find all linear evolution equations for which the group velocity equals the phase velocity. Justify your answer.

8.5.6. Show that the phase velocity is greater than the group velocity if and only if the phase velocity is a decreasing function of k for $k > 0$ and an increasing function of k for $k < 0$. How would you observe this in a physical system?

♢ 8.5.7. (*a*) *Conservation of Mass*: Prove that $T = u$ is a density associated with a conservation law of the dispersive wave equation (8.90). What is the corresponding flux? Under what conditions is total mass conserved? (*b*) *Conservation of Energy*: Establish the same result for the energy density $T = u^2$. (*c*) Is u^3 the density of a conservation law?

♢ 8.5.8. Prove that when $t = \pi p/q$, where p, q are integers, the solution (8.102) is constant on each interval $\pi j/q < x < \pi(j+1)/q$ for integers $j \in \mathbb{Z}$. *Hint*: Use Exercise 6.1.29(*d*). *Remark*: The proof that the solution is continuous and fractal at irrational times is considerably more difficult, [**90**].

♢ 8.5.9. (*a*) Find the complex Fourier series representing the fundamental solution $F(t,x;\xi)$ to the periodic initial-boundary value problem (8.100). (*b*) Prove that at time $t = 2\pi p/q$, where p, q are relatively prime integers, $F(t,x;\xi)$ is a linear combination of delta functions based at the points $\xi + 2\pi j/q$. *Hint*: Use Exercise 6.1.29(*c*). (*c*) Let $u(t,x)$ be any solution to (8.100). Prove that $u(2\pi p/q, x)$ is a linear combination of a finite number of translates, $f(x - x_j)$, of the initial data.

The Korteweg–de Vries Equation

The simplest wave model that combines dispersion with nonlinearity is the celebrated *Korteweg–de Vries equation*

$$u_t + u_{xxx} + u\,u_x = 0. \tag{8.112}$$

It was first derived, in 1872, by the French applied mathematician Joseph Boussinesq, [**21**; eq. (30)], [**22**; eqs. (283, 291)], as a model for surface waves on shallow water. Two decades later, it was rediscovered by the Dutch applied mathematician Diederik Korteweg and his student Gustav de Vries, [**65**], and, despite Boussinesq's priority, it is nowadays named after them. In the early 1960s, the American mathematical physicists Martin Kruskal and Norman Zabusky, [**125**], used the Korteweg–de Vries equation as a continuum model for a one-dimensional chain of masses interconnected by nonlinear springs: the Fermi–Pasta–Ulam problem, [**40**]. Numerical experimentation revealed its many remarkable properties, which were soon rigorously established. Their work sparked the rapid development of one of the most remarkable and far-reaching discoveries of the modern era: integrable nonlinear partial differential equations, [**2**, **36**].

The most important special solutions to the Korteweg–de Vries equation are the *traveling waves*. We seek solutions

$$u = v(\xi) = v(x - ct), \qquad \text{where} \qquad \xi = x - ct,$$

that have a fixed profile while moving with speed c. By the chain rule,

$$\frac{\partial u}{\partial t} = -c\,v'(\xi), \qquad \frac{\partial u}{\partial x} = v'(\xi), \qquad \frac{\partial^3 u}{\partial x^3} = v'''(\xi).$$

Substituting these expressions into the Korteweg–de Vries equation (8.112), we conclude that $v(\xi)$ must satisfy the nonlinear third-order ordinary differential equation

$$v''' + v\,v' - c\,v' = 0. \tag{8.113}$$

Let us further assume that the traveling wave is *localized*, meaning that the solution and its derivatives are vanishingly small at large distances:

$$\lim_{x \to \pm\infty} u(t,x) \;=\; \lim_{x \to \pm\infty} \frac{\partial u}{\partial x}\,(t,x) \;=\; \lim_{x \to \pm\infty} \frac{\partial^2 u}{\partial x^2}\,(t,x) = 0. \tag{8.114}$$

This implies that we should impose the boundary conditions

$$\lim_{\xi \to \pm\infty} v(\xi) \;=\; \lim_{\xi \to \pm\infty} v'(\xi) \;=\; \lim_{\xi \to \pm\infty} v''(\xi) = 0. \tag{8.115}$$

The ordinary differential equation (8.113) can, in fact, be solved in closed form. First, note that it has the form

$$\frac{d}{d\xi}\big(v'' + \tfrac{1}{2}\,v^2 - c\,v\big) = 0, \qquad \text{and hence} \qquad v'' + \tfrac{1}{2}\,v^2 - c\,v = a,$$

where a indicates the constant of integration. The localizing boundary conditions (8.115) imply that $a = 0$. Multiplying the resulting equation by v' allows us to integrate a second time:

$$0 = v'\big(v'' + \tfrac{1}{2}\,v^2 - c\,v\big) = \frac{d}{d\xi}\Big[\,\tfrac{1}{2}\,(v')^2 + \tfrac{1}{6}\,v^3 - \tfrac{1}{2}\,c\,v^2\,\Big] = 0.$$

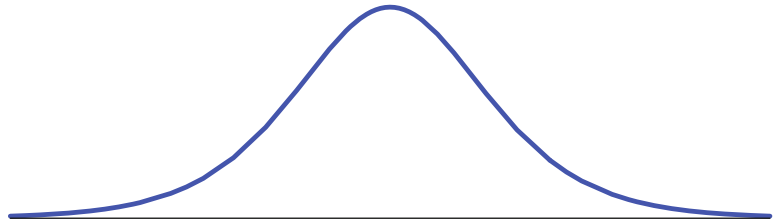


Figure 8.13. Solitary wave/soliton.

Thus,

$$\tfrac{1}{2}(v')^2 + \tfrac{1}{6}v^3 - \tfrac{1}{2}cv^2 = b,$$

where b is a second constant of integration, which, again by the boundary conditions (8.115), is also zero. Setting $b = 0$, and solving for v', we conclude that $v(\xi)$ satisfies the autonomous first-order ordinary differential equation

$$\frac{dv}{d\xi} = v\sqrt{c - \tfrac{1}{3}v}\,,$$

which is integrated by the standard method:

$$\int \frac{dv}{v\sqrt{c - \frac{1}{3}v}} = \xi + \delta,$$

where δ is constant. Consulting a table of integrals, e.g., [**48**], and then solving for v, we conclude that the solution has the form

$$v(\xi) = 3c\,\operatorname{sech}^2\big(\tfrac{1}{2}\sqrt{c}\,\xi + \delta\big), \tag{8.116}$$

where

$$\operatorname{sech} y = \frac{1}{\cosh y} = \frac{2}{e^y + e^{-y}}$$

is the *hyperbolic secant function*. The solution has the form graphed in Figure 8.13. It is a symmetric, monotone, exponentially decreasing function on either side of its maximum height of $3c$. (Despite its suggestive profile, it is *not* a Gaussian.) The resulting localized traveling-wave solutions to the Korteweg–de Vries equation are thus

$$u(t,x) = 3c\,\operatorname{sech}^2\big[\tfrac{1}{2}\sqrt{c}\,(x - ct) + \delta\big], \tag{8.117}$$

where $c > 0$ represents the wave speed — which is necessarily positive, and so all such solutions move to the right — while δ represents an overall phase shift. The amplitude of the wave is three times its speed, while its width is proportional to $1/\sqrt{c}$. Thus, the taller (and narrower) the wave, the faster it moves.

Localized traveling waves are commonly known as *solitary waves*. They were first observed in nature by the British engineer J. Scott Russell, [**104**], who recounts how one was triggered by the sudden motion of a barge along an Edinburgh canal. Scott Russell ended up chasing the propagating wave on horseback for several miles — a physical indication of its stability. Russell's observations were dismissed by his contemporary Airy, who, relying on his linearly dispersive model for surface waves (8.90), claimed that such localized

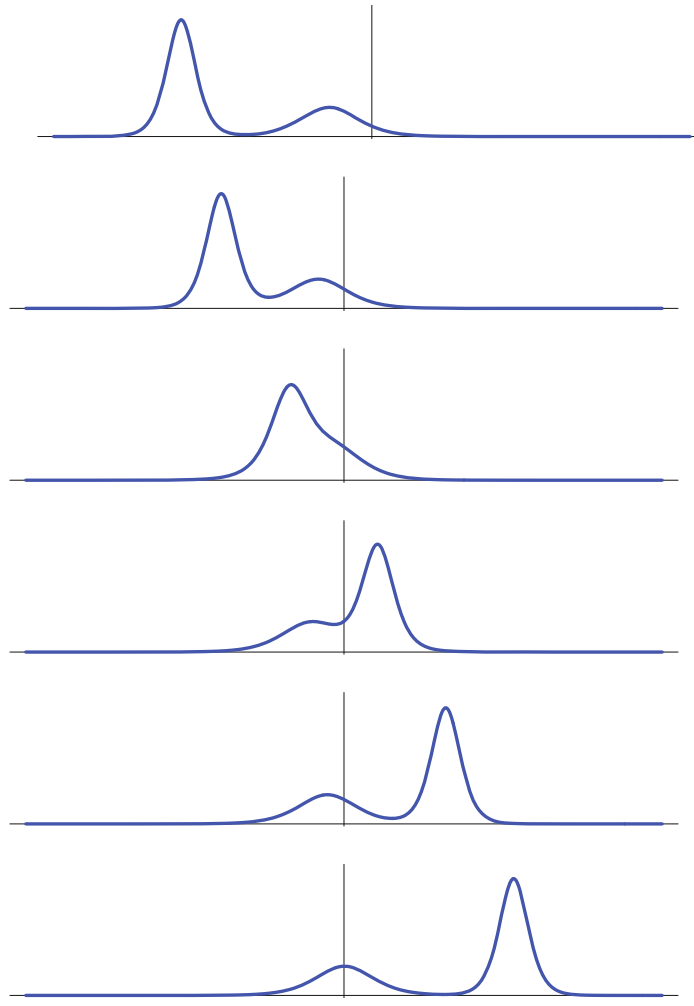


Figure 8.14. Interaction of two solitons.

disturbances could not exist. Much later, Boussinesq derived the proper nonlinear surface wave model (8.112), valid for long waves in shallow water, along with its solitary wave solutions (8.117), thereby fully exonerating Russell's physical observations and insight.

It took almost a century before all the remarkable properties of these solutions came to light. The most striking is how two such solitary waves interact. While linear equations always admit a superposition principle, one cannot naïvely combine two solutions to a nonlinear equation. However, in the case of the Korteweg–de Vries equation, suppose the initial data represent a taller solitary wave to the left of a shorter one. As time evolves, the taller wave will move faster, and eventually catch up to the shorter one. They then experience a complicated nonlinear interaction, as expected. But, remarkably, after a while, they emerge from the interaction unscathed! The smaller wave is now in back and the larger one in front, and both unchanged in speed, amplitude, and profile. They then

proceed independently, with the smaller solitary wave lagging farther and farther behind the faster, taller wave. The only effect of their encounter is an overall phase shift, so that the taller wave is a bit behind where it would be if it had not encountered the shorter wave, while the shorter wave is a little ahead of its unhindered position. Figure 8.14 plots a typical such interaction.

Owing to this "particle-like" behavior under interaction, these solutions were given a special name: *soliton*. An explicit formula for a *two-soliton solution* to the Korteweg–de Vries equation can be written in the following form:

$$u(t,x) = 12\,\frac{\partial^2}{\partial x^2}\log \Delta(t,x), \tag{8.118}$$

where

$$\Delta(t,x) = \det\begin{pmatrix} 1+\varepsilon_1(t,x) & \dfrac{2\,b_1}{b_1+b_2}\,\varepsilon_2(t,x) \\[2ex] \dfrac{2\,b_2}{b_1+b_2}\,\varepsilon_1(t,x) & 1+\varepsilon_2(t,x) \end{pmatrix}, \tag{8.119}$$

where $0 < b_1 < b_2$, and

$$\varepsilon_j(t,x) = \exp\big[\,b_j(x - b_j^2\,t) + d_j\,\big], \qquad j = 1,2. \tag{8.120}$$

The constants $c_j = b_j^2$ represent the wave speeds, while the d_j correspond to phase shifts of the individual solitons. Proving that (8.118) is indeed a solution to the Korteweg–de Vries equation is a straightforward, albeit tedious, exercise in differentiation. In Exercise 8.5.14, the reader is asked to investigate its asymptotic behavior, as $t \to \pm\infty$, and prove that the solution does, indeed, break up into two solitons, having the same profiles, speeds, and amplitudes in both the distant past and future.

A similar dynamic occurs when there are multiple collisions among solitons. Faster solitons catch up to slower ones moving to their right. After the various solitons finish colliding and interacting, they emerge in order, from smallest to largest, each moving at its characteristic speed and becoming more and more separated from its peers. An explicit formula for the n–soliton solution is provided by the same logarithmic derivative (8.118) in which $\Delta(t,x)$ now represents the determinant of an $n \times n$ matrix whose i^{th} diagonal entry is $1 + \varepsilon_i(t,x)$, while the off-diagonal (i,j) entry, $i \neq j$, is $\dfrac{2\,b_i}{b_i+b_j}\,\varepsilon_j(t,x)$, using the same formula (8.120) for the ε_j's, and where $0 < b_1 < \cdots < b_n$ correspond to the n different soliton wave speeds $c_j = b_j^2$. Furthermore, it can be shown that, starting with an *arbitrary* localized initial disturbance $u(0,x) = f(x)$ that decays sufficiently rapidly as $|\,x\,| \to \infty$, the resulting solution eventually emits a finite number of solitons of different heights, moving off at their respective speeds to the right, and so arranged in order from smallest to largest, followed by a small, asymptotically self-similar dispersive tail that gradually disappears.

The source of these highly non-obvious facts and formulas lies beyond the scope of this introductory text. Soon after the initial numerical studies, Gardner, Green, Kruskal, and Miura, [**45**], discovered a profound connection between the solutions to the Korteweg–de Vries equation and the eigenvalues λ of the Sturm–Liouville boundary value problem

$$-\frac{d^2\psi}{dx^2} + 6\,u(t,x)\,\psi = \lambda\,\psi, \quad -\infty < x < \infty, \qquad \text{with} \qquad \psi(t,x) \longrightarrow 0 \quad \text{as} \quad |\,x\,| \longrightarrow \infty. \tag{8.121}$$

Their remarkable result is that whenever $u(t,x)$ is a localized solution to the Korteweg–de Vries equation (8.112), the eigenvalues of (8.121) are constant, meaning that they do not

vary with the time t, while the continuous spectrum has a very simple temporal evolution. In physical applications of the stationary Schrödinger equation (8.121), in which $u(t,x)$ represents a quantum-mechanical potential, the eigenvalues correspond to bound states, while the continuous spectrum governs its scattering behavior. The solution to the so-called *inverse scattering problem* reconstructs the potential $u(t,x)$ from its spectrum, and can be viewed as a nonlinear version of the Fourier transform, in that it effectively linearizes the Korteweg–de Vries equation and thereby reveals its many remarkable properties. In particular, the eigenvalues are responsible for the preceding determinantal formulae for the multi-soliton solutions, while, when present, the continuous spectrum governs the dispersive tail. See [**2**, **36**] for additional details.

Exercises

8.5.10. Justify the statement that the width of a soliton is proportional to the inverse of the square root of its speed.

8.5.11. Prove that the function (8.116) is a symmetric, monotone, exponentially decreasing function on either side of its maximum height of $3c$.

8.5.12. Let $u(t,x)$ solve the Korteweg–de Vries equation.
(*a*) Show that $U(t,x) = u(t, x - ct) + c$ is also a solution.
(*b*) Give a physical interpretation of this symmetry.

8.5.13. (*a*) Find all scaling symmetries of the Korteweg–de Vries equation.
(*b*) Write down an ansatz for the similarity solutions, and then find the corresponding reduced ordinary differential equation. (Unfortunately, the similarity solutions cannot be written in terms of elementary functions, [**2**].)

♡ 8.5.14. (*a*) Let $u(t,x)$ be the two-soliton solution defined in (8.118). Let $\widetilde{u}(t,\xi) = u(t, \xi + ct)$ represent the solution as viewed in a coordinate frame moving with speed c. Prove that

$$\lim_{t \to \infty} \widetilde{u}(t,\xi) = \begin{cases} 3c_1 \operatorname{sech}^2\left[\, \tfrac{1}{2}\sqrt{c_1}\,\xi + \delta_1 \,\right], & c = c_1, \\ 3c_2 \operatorname{sech}^2\left[\, \tfrac{1}{2}\sqrt{c_2}\,\xi + \delta_2 \,\right], & c = c_2, \\ 0, & \text{otherwise}, \end{cases}$$

for suitable constants δ_1, δ_2. Explain why this justifies the statement that the solution indeed breaks up into two individual solitons as $t \to \infty$. (*b*) Explain why $\widetilde{u}(t,\xi)$ has a similar limiting behavior as $t \to -\infty$, but with possibly different constants $\widehat{\delta}_1, \widehat{\delta}_2$. (*c*) Use your formulas to discuss how the solitons are affected by the collision.

8.5.15. Let $\alpha, \beta \neq 0$. Find the soliton solutions to the rescaled Korteweg–de Vries equation $u_t + \alpha\, u_{xxx} + \beta\, u\, u_x = 0$. How are their speed, amplitude, and width interrelated?

8.5.16. (*a*) Find the solitary wave solutions to the *modified Korteweg–de Vries equation* $u_t + u_{xxx} + u^2 u_x = 0$. (*b*) Discuss how the amplitude and width of the solitary waves are related to their speeeds. *Note*: The modified Korteweg–de Vries equation is also integrable, and its solitary wave solutions are solitons, cf. [**36**].

8.5.17. Answer Exercise 8.5.16 for the *Benjamin–Bona–Mahony equation* $u_t - u_{xxt} + u\, u_x = 0$, [**14**]. *Note*: The BBM equation is *not* integrable, and collisions between its solitary waves produce a small, but measurable, inelastic effect, [**1**].

♢ 8.5.18. (*a*) Show that $T_1 = u$ is the density for a conservation law for the Korteweg–de Vries equation. (*b*) Show that $T_2 = u^2$ is also a conserved density. (*c*) Find a conserved density of the form $T_3 = u_x^2 + \mu u^3$ for a suitable constant μ. *Remark*: The Korteweg–de Vries

equation in fact has *infinitely many* conservation laws, whose densities depend on higher and higher-order derivatives of the solution, [**76**, **87**]. It was this discovery that unlocked the door to all its remarkable integrability properties, [**2**, **36**].

8.5.19. Find two conservation laws of
(*a*) the modified Korteweg–de Vries equation $u_t + u_{xxx} + u^2 u_x = 0$;
(*b*) the Benjamin–Bona–Mahony equation $u_t - u_{xxt} + u\, u_x = 0$.

Chapter 11
Dynamics of Planar Media

In previous chapters, we studied the equilibrium configurations of planar media — plates and membranes — governed by the two-dimensional Laplace and Poisson equations. In this chapter, we analyze their dynamics, modeled by the two-dimensional heat and wave equations. The heat equation describes diffusion of, say, heat energy in a thin metal plate, an animal population dispersing over a region, or a pollutant spreading out into a shallow lake. The wave equation models small vibrations of a two-dimensional membrane such as a drum. Since both equations fit into the general framework for dynamics that we established in Section 9.5, their solutions share many of the general qualitative and analytic properties possessed by their respective one-dimensional counterparts.

Although the increase in dimension may tax our analytical prowess, we have, in fact, already mastered the principal solution techniques: separation of variables, eigenfunction series, and fundamental solutions. When applied to partial differential equations in higher dimensions, separation of variables in curvilinear coordinates often leads to new linear, but non-constant-coefficient, ordinary differential equations, whose solutions are no longer elementary functions. Rather, they are expressed in terms of a variety of important *special functions*, which include the error and Airy functions we encountered earlier; the Bessel functions, which play a starring role in the present chapter; and the Legendre and Ferrers functions, spherical harmonics, and spherical Bessel functions arising in three-dimensional problems. Special functions are ubiquitous in more advanced applications in physics, chemistry, mechanics, and mathematics, and, over the last two hundred and fifty years, many prominent mathematicians have devoted significant effort to establishing their fundamental properties, to the extent that they are now, by and large, well understood, [**86**]. To acquire the requisite familiarity with special functions, in preparation for employing them to solve higher-dimensional partial differential equations, we must first learn basic series solution techniques for linear second-order ordinary differential equations.

11.1 Diffusion in Planar Media

As we learned in Chapter 4, the equilibrium temperature $u(x, y)$ of a thin, uniform, isotropic plate is governed by the two-dimensional Laplace equation

$$\Delta u = u_{xx} + u_{yy} = 0.$$

Working by analogy, the dynamical diffusion of the plate's temperature should be modeled by the two-dimensional heat equation

$$u_t = \gamma\, \Delta u = \gamma\, (u_{xx} + u_{yy}). \tag{11.1}$$

P.J. Olver, *Introduction to Partial Differential Equations*, Undergraduate Texts in Mathematics,
DOI 10.1007/978-3-319-02099-0_11,

The coefficient $\gamma > 0$, assumed constant, measures the relative speed of diffusion of heat energy throughout the plate; its positivity is required on physical grounds, and also serves to avoid ill-posedness inherent in running diffusion processes backwards in time. In this model, we are assuming that the plate is uniform and isotropic, and experiences no loss of heat or external heat sources other than at its edge — which can be arranged by covering its top and bottom with insulation.

The solution $u(t, \mathbf{x}) = u(t, x, y)$ to the heat equation measures the temperature, at time t, at each point $\mathbf{x} = (x, y)$ in the (bounded) domain $\Omega \subset \mathbb{R}^2$ occupied by the plate. To uniquely specify the solution $u(t, x, y)$, we must impose suitable initial and boundary conditions. The initial data is the temperature of the plate

$$u(0, x, y) = f(x, y), \qquad (x, y) \in \Omega, \tag{11.2}$$

at an initial time, which for simplicity, we take to be $t_0 = 0$. The most important boundary conditions are as follows:

- *Dirichlet boundary conditions*: Specifying

$$u = h \qquad \text{on} \qquad \partial\Omega \tag{11.3}$$

  fixes the temperature along the edge of the plate.
- *Neumann boundary conditions*: Let $\mathbf{n}$ be the unit outwards normal on the boundary of the domain. Specifying the normal derivative of the temperature,

$$\frac{\partial u}{\partial \mathbf{n}} = k \qquad \text{on} \qquad \partial\Omega, \tag{11.4}$$

  effectively prescribes the heat flux along the boundary. Setting $k = 0$ corresponds to an insulated boundary.
- *Mixed boundary conditions*: More generally, we can impose Dirichlet conditions on part of the boundary $D \subsetneq \partial\Omega$ and Neumann conditions on its complement $N = \partial\Omega \setminus D$. For instance, homogeneous mixed boundary conditions

$$u = 0 \qquad \text{on} \qquad D, \qquad\qquad \frac{\partial u}{\partial \mathbf{n}} = 0 \qquad \text{on} \qquad N, \tag{11.5}$$

  correspond to freezing a portion of the boundary and insulating the remainder.
- *Robin boundary conditions*:

$$\frac{\partial u}{\partial \mathbf{n}} + \beta\, u = \tau \qquad \text{on} \qquad \partial\Omega, \tag{11.6}$$

  where the edge of the plate sits in a heat bath at temperature τ.

Under reasonable assumptions on the domain, the initial data, and the boundary data, a general theorem, [**34**, **38**, **99**], guarantees the existence of a unique solution $u(t, x, y)$ to any of these initial-boundary value problems for all subsequent times $t > 0$. Our practical goal is to both compute and understand the behavior of the solution in specific situations.

Derivation of the Diffusion and Heat Equations

The physical derivation of the two-dimensional (and three-dimensional) heat equation relies on the same basic thermodynamic laws that were used, in Section 4.1, to establish the one-dimensional version. The first principle is that heat energy flows from hot to cold as

rapidly as possible. According to multivariable calculus, [**8**, **108**], the negative temperature gradient $-\nabla u$ points in the direction of the steepest decrease in the temperature function u at a point, and so heat energy will flow in that direction. Therefore, the heat flux vector $\mathbf{w}$, which measures the magnitude and direction of the flow of heat energy, should be proportional to the temperature gradient:

$$\mathbf{w}(t, x, y) = -\kappa(x, y)\, \nabla u(t, x, y). \tag{11.7}$$

The scalar quantity $\kappa(x, y) > 0$ measures the *thermal conductivity* of the material, so (11.7) is the multi-dimensional form of *Fourier's Law of Cooling* (4.5). We are assuming that the thermal conductivity depends only on the position $(x, y) \in \Omega$, which means that the material in the plate

(a) is not changing in time;

(b) is *isotropic*, meaning that its thermal conductivity is the same in all directions;

(c) and, moreover, its thermal conductivity is not affected by any change in temperature.

Dropping either assumption (b) or (c) would result in a considerably more challenging nonlinear diffusion equation.

The second thermodynamic principle is that, in the absence of external heat sources, heat can enter any subregion $R \subset \Omega$ only through its boundary ∂R. (Keep in mind that the plate is insulated from above and below.) Let $\varepsilon(t, x, y)$ denote the heat energy density at each time and point in the domain, so that

$$H_R(t) = \iint_R \varepsilon(t, x, y)\, dx\, dy$$

represents the total heat energy contained within the subregion R at time t. The amount of additional heat energy entering R at a boundary point $\mathbf{x} \in \partial R$ is given by the normal component of the heat flux vector, namely $-\mathbf{w} \cdot \mathbf{n}$, where, as always, $\mathbf{n}$ denotes the *outward* unit normal to the boundary ∂R. Thus, the total heat flux entering the region R is obtained by integration along the boundary of R, resulting in the line integral $-\oint_{\partial R} \mathbf{w} \cdot \mathbf{n}\; ds$. Equating the rate of change of heat energy to the heat flux yields

$$\frac{dH_R}{dt} = \iint_R \frac{\partial \varepsilon}{\partial t}\,(t, x, y)\, dx\, dy = -\oint_{\partial R} \mathbf{w} \cdot \mathbf{n}\; ds = -\iint_R \nabla \cdot \mathbf{w}\; dx\, dy,$$

where we applied the divergence form of Green's Theorem, (6.80), to convert the flux line integral into a double integral. Thus,

$$\iint_R \left(\frac{\partial \varepsilon}{\partial t} + \nabla \cdot \mathbf{w} \right) dx\, dy = 0. \tag{11.8}$$

Keep in mind that this result must hold for *any* subdomain $R \subset \Omega$. Now, according to Exercise 11.1.13, the only way in which an integral of a continuous function can vanish for *all* subdomains is if the integrand is identically zero, and so

$$\frac{\partial \varepsilon}{\partial t} + \nabla \cdot \mathbf{w} = 0. \tag{11.9}$$

In this manner, we arrive at the basic *conservation law* relating the heat energy density ε and the heat flux vector $\mathbf{w}$.

As in our one-dimensional model, cf. (4.3), the heat energy density $\varepsilon(t,x,y)$ is proportional to the temperature, so

$$\varepsilon(t,x,y) = \sigma(x,y)\,u(t,x,y), \qquad \text{where} \qquad \sigma(x,y) = \rho(x,y)\,\chi(x,y) \tag{11.10}$$

is the product of the *density* ρ and the *specific heat capacity* χ of the material at the point $(x,y) \in \Omega$. Combining this with the Fourier Law (11.7) and the energy balance equation (11.10) leads to the general two-dimensional *diffusion equation*

$$\frac{\partial u}{\partial t} = \frac{1}{\sigma}\,\nabla \cdot \big(\,\kappa\,\nabla u\,\big) \tag{11.11}$$

governing the thermodynamics of an isotropic medium in the absence of external heat sources or sinks. In full detail, this second-order partial differential equation is

$$\frac{\partial u}{\partial t} = \frac{1}{\sigma(x,y)}\,\left[\,\frac{\partial}{\partial x}\left(\,\kappa(x,y)\,\frac{\partial u}{\partial x}\,\right) + \frac{\partial}{\partial y}\left(\,\kappa(x,y)\,\frac{\partial u}{\partial y}\,\right)\,\right]. \tag{11.12}$$

Such diffusion equations are also used to model movements of populations, e.g., bacteria in a petri dish or wolves in the Canadian Rockies, [**81**, **84**]. Here the solution $u(t,x,y)$ represents the population density at position (x,y) at time t, which diffuses over the domain due to random motions of the individuals. Similar diffusion processes model the mixing of solutes in liquids, with the diffusion induced by the random Brownian motion from molecular collisions. More generally, diffusion processes in the presence of chemical reactions and convection due to fluid motion are modeled by the more general class of *reaction-diffusion* and *convection-diffusion equations*, [**107**].

In particular, if the body (or the environment or the solvent) is uniform, then both σ and κ are constant, and so (11.11) reduces to the heat equation (11.1) with *thermal diffusivity*

$$\gamma = \frac{\kappa}{\sigma} = \frac{\kappa}{\rho\,\chi}\,. \tag{11.13}$$

Both the heat and more general diffusion equations are examples of *parabolic* partial differential equations, the terminology being adapted from Definition 4.12 to apply to partial differential equations in more than two variables. As we will see, all the basic qualitative features of solutions to the one-dimensional heat equation carry over to parabolic partial differential equations in higher dimensions.

Indeed, the general diffusion equation (11.12) can be readily fit into the self-adjoint dynamical framework of Section 9.5, taking the form

$$u_t = -\,\nabla^* \circ \nabla u. \tag{11.14}$$

The gradient operator ∇ maps scalar fields u to vector fields $\mathbf{v} = \nabla u$; its adjoint ∇^*, which goes in the reverse direction, is taken with respect to the weighted inner products

$$\langle\, u\,,\widetilde{u}\,\rangle = \iint_\Omega u(x,y)\,\widetilde{u}(x,y)\,\sigma(x,y)\,dx\,dy, \quad \langle\!\langle\, \mathbf{v}\,,\widetilde{\mathbf{v}}\,\rangle\!\rangle = \iint_\Omega \mathbf{v}(x,y)\cdot\widetilde{\mathbf{v}}(x,y)\,\kappa(x,y)\,dx\,dy, \tag{11.15}$$

between, respectively, scalar and vector fields. As in (9.33), a straightforward integration by parts tells us that

$$\nabla^*\mathbf{v} = -\,\frac{1}{\sigma}\,\nabla\cdot(\kappa\,\mathbf{v}) \;=\; -\,\frac{1}{\sigma}\,\left[\,\frac{\partial(\kappa\,v_1)}{\partial x} + \frac{\partial(\kappa\,v_2)}{\partial y}\,\right], \qquad \text{when} \qquad \mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}. \tag{11.16}$$

Therefore, the right-hand side of (11.14) equals

$$-\nabla^* \circ \nabla u = \frac{1}{\sigma}\,\nabla \cdot (\kappa\,\nabla u), \tag{11.17}$$

which thereby recovers the general diffusion equation (11.11). As always, the validity of the adjoint formula (11.16) rests on the imposition of suitable homogeneous boundary conditions: Dirichlet, Neumann, mixed, or Robin.

In particular, to obtain the heat equation, we take σ and κ to be constant, and so the inner products (11.15) reduce, up to a constant factor, to the usual L^2 inner products between scalar and vector fields. In this case, the adjoint of the gradient is, up to a scale factor, minus the divergence: $\nabla^* = -\gamma\,\nabla\cdot$, where $\gamma = \kappa/\sigma$. In this scenario, (11.14) reduces to the two-dimensional heat equation (11.1).

Separation of Variables

Let us now discuss analytical solution techniques. According to Section 9.5, the separable solutions to any linear evolution equation

$$u_t = -\,S[u] \tag{11.18}$$

are of exponential form

$$u(t,x,y) = e^{-\lambda t}\,v(x,y). \tag{11.19}$$

Since the linear operator S involves differentiation with respect to only the spatial variables x, y, we obtain

$$\frac{\partial u}{\partial t} = -\,\lambda\,e^{-\lambda t}\,v(x,y), \qquad \text{while} \qquad S[u] = e^{-\lambda t}\,S[v].$$

Substituting back into the diffusion equation (11.18) and canceling the exponentials, we conclude that

$$S[v] = \lambda\,v. \tag{11.20}$$

Thus, $v(x,y)$ must be an eigenfunction for the linear operator S, subject to the relevant homogeneous boundary conditions.

In the case of the heat equation (11.1),

$$S[u] = -\gamma\,\Delta u,$$

and hence, as in Example 9.40, the eigenvalue equation (11.20) is the two-dimensional *Helmholtz equation*

$$\gamma\,\Delta v + \lambda\,v = 0, \qquad \text{or, in detail,} \qquad \gamma\left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2}\right) + \lambda\,v = 0. \tag{11.21}$$

According to Theorem 9.34, self-adjointness implies that the eigenvalues are all real and nonnegative: $\lambda \geq 0$. In the positive definite cases — Dirichlet and mixed boundary conditions — they are strictly positive, while the Neumann boundary value problem admits a zero eigenvalue $\lambda_0 = 0$ corresponding to the constant eigenfunction $v_0(x,y) \equiv 1$.

Let us index the eigenvalues in increasing order:

$$0 < \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \cdots, \tag{11.22}$$

repeated according to their multiplicities, where $\lambda_0 = 0$ is an eigenvalue only in the Neumann case, and $\lambda_k \to \infty$ as $k \to \infty$. For each eigenvalue λ_k, let $v_k(x, y)$ be an independent eigenfunction. The corresponding separable solution is

$$u_k(t, x, y) = e^{-\lambda_k t}\, v_k(x, y).$$

Those corresponding to positive eigenvalues are exponentially decaying in time, while a zero eigenvalue produces a constant solution $u_0(t, x, y) \equiv 1$. The general solution to the homogeneous boundary value problem can then be built up as an infinite series in these basic eigensolutions

$$u(t, x, y) = \sum_{k=1}^{\infty} c_k\, u_k(t, x, y) = \sum_{k=1}^{\infty} c_k\, e^{-\lambda_k t}\, v_k(x, y). \tag{11.23}$$

The coefficients c_k are prescribed by the initial conditions, which require

$$\sum_{k=1}^{\infty} c_k\, v_k(x, y) = f(x, y). \tag{11.24}$$

Since S is self-adjoint, Theorem 9.33 guarantees orthogonality† of the eigenfunctions under the L^2 inner product on the domain Ω:

$$\langle\, v_j\,, v_k\,\rangle = \iint_{\Omega} v_j(x, y)\, v_k(x, y)\, dx\, dy = 0, \qquad j \neq k. \tag{11.25}$$

As a consequence, the coefficients in (11.24) are given by the standard orthogonality formula (9.104), namely

$$c_k = \frac{\langle\, f\,, v_k\,\rangle}{\|\, v_k\,\|^2} = \frac{\displaystyle\iint_{\Omega} f(x, y)\, v_k(x, y)\, dx\, dy}{\displaystyle\iint_{\Omega} v_k(x, y)^2\, dx\, dy}\,. \tag{11.26}$$

(For the more general diffusion equation (11.11), one uses the appropriately weighted inner product.) The exponential decay of the eigenfunction coefficients implies that the resulting eigensolution series (11.23) converges and thus produces the solution to the initial-boundary value problem for the diffusion equation. See [**34**; p. 369] for a precise statement and proof of the general theorem.

Qualitative Properties

Before tackling examples in which we are able to construct explicit formulas for the eigenfunctions and eigenvalues, let us see what the eigenfunction series solution (11.23) can tell us about general diffusion processes. Based on our experience with the case of a one-dimensional bar, the final conclusions will not be especially surprising. Indeed, they also apply, word for word, to diffusion processes in three-dimensional solid bodies. A reader who is impatient to see the explicit formulas may wish to skip ahead to the following section, returning here as needed.

† As usual, in the case of a repeated eigenvalue, one chooses an orthogonal basis of the associated eigenspace to ensure orthogonality of all the basis eigenfunctions.

Keep in mind that we are still dealing with the solution to the homogeneous boundary value problem. The first observation is that all terms in the series solution (11.23), with the possible exception of a null eigenfunction term that appears in the semi-definite Neumann case, are tending to zero exponentially fast. Since most eigenvalues are large, all the higher-order terms in the series become almost instantaneously negligible, and hence the solution can be accurately approximated by a finite sum over the first few eigenfunction modes. As time goes on, more and more of the modes can be neglected, and the solution decays to thermal equilibrium at an exponentially fast rate. The rate of convergence to thermal equilibrium is, for most initial data, governed by the smallest positive eigenvalue $\lambda_1 > 0$ for the Helmholtz boundary value problem on the domain.

In the positive definite cases of homogeneous Dirichlet or mixed boundary conditions, thermal equilibrium is $u(t, x, y) \to u_\star(x, y) \equiv 0$. Here, the equilibrium temperature is equal to the zero boundary temperature — even if this temperature is fixed on only a small part of the boundary. The initial heat is eventually dissipated away through the uninsulated part of the boundary. In the semi-definite Neumann case, corresponding to a completely insulated plate, the general solution has the form

$$u(t, x, y) = c_0 + \sum_{k=1}^{\infty} c_k \, e^{-\lambda_k t} \, v_k(x, y), \tag{11.27}$$

where the sum is over the positive eigenmodes, $\lambda_k > 0$. Since all the summands are exponentially decaying, the final equilibrium temperature $u_\star = c_0$ is the same as the constant term in the eigenfunction expansion. We evaluate this term using the orthogonality formula (11.26), and so, as $t \to \infty$,

$$u(t, x, y) \;\longrightarrow\; c_0 = \frac{\langle\, f \,,\, 1 \,\rangle}{\| \, 1 \, \|^2} \;=\; \frac{\displaystyle\iint_\Omega f(x, y)\, dx\, dy}{\displaystyle\iint_\Omega dx\, dy} \;=\; \frac{1}{\text{area } \Omega} \iint_\Omega f(x, y)\, dx\, dy. \tag{11.28}$$

We conclude that the equilibrium temperature is equal to the average initial temperature distribution. Thus, when the plate is fully insulated, the heat energy cannot escape, and instead redistributes itself in a uniform manner over the domain.

Diffusion has a smoothing effect on the initial temperature distribution $f(x, y)$. Assume that the eigenfunction coefficients are uniformly bounded, so $| \, c_k \, | \le M$ for some constant M. This will certainly be the case if $f(x, y)$ is piecewise continuous or, more generally, belongs to L^2, since Bessel's inequality, (3.117), which holds for general orthogonal systems, implies that $c_k \to 0$ as $k \to \infty$. Many distributions, including delta functions, also have bounded Fourier coefficients. Then, at any time $t > 0$ after the initial instant, the coefficients $c_k \, e^{-\lambda_k t}$ in the eigenfunction series solution (11.23) are exponentially small as $k \to \infty$, which is enough to ensure smoothness of the solution $u(t, x, y)$ for each $t > 0$. Therefore, the diffusion process serves to immediately smooth out jumps, corners, and other discontinuities in the initial data. As time progresses, the local variations in the solution become less and less pronounced, as it asymptotically reaches a constant equilibrium state.

As a result, diffusion processes can be effectively applied to smooth and denoise planar images. The initial data $u(0, x, y) = f(x, y)$ represents the gray-scale value of the image at position (x, y), so that $0 \le f(x, y) \le 1$, with 0 representing black and 1 representing white. As time progresses, the solution $u(t, x, y)$ represents a more and more smoothed version

Figure 11.1. Smoothing a gray scale image.

of the image. Although this has the effect of removing unwanted high-frequency noise, there is also a gradual blurring of the actual features. Thus, the "time" or "multiscale" parameter t needs to be chosen to optimally balance between the two effects — the larger t is the more noise is removed, but the more noticeable the blurring. A representative illustration appears in Figure 11.1. The blurring affects small-scale features first, then, gradually, those at larger and larger scales, until eventually the entire image is blurred to a uniform gray. To further suppress undesirable blurring effects, modern image-processing filters are based on anisotropic (and thus *nonlinear*) diffusion equations; see [**100**] for a survey of recent progress in this active field.

Since the forward heat equation effectively blurs the features in an image, we might be tempted to reverse "time" in order to sharpen the image. However, the argument presented in Section 4.1 tells us that the backwards heat equation is ill-posed, and hence cannot be used directly for this purpose. Various "regularization" strategies have been devised to circumvent this mathematical barrier, and thereby design effective image enhancement algorithms, [**46**].

Inhomogeneous Boundary Conditions and Forcing

Let us next briefly discuss how to incorporate inhomogeneous boundary conditions and external heat sources into the general solution framework. Consider, as a specific example, the forced heat equation

$$u_t = \gamma\, \Delta u + F(x,y) \qquad \text{for} \qquad (x,y) \in \Omega, \tag{11.29}$$

where $F(x,y)$ represents an unvarying external heat source or sink, subject to inhomogeneous Dirichlet boundary conditions

$$u(x,y) = h(x,y) \qquad \text{for} \qquad (x,y) \in \partial\Omega, \tag{11.30}$$

that fixes the temperature of the plate on its boundary. When the external forcing does not vary in time, we expect the solution to eventually settle down to an equilibrium configuration: $u(t,x,y) \to u_\star(x,y)$ as $t \to \infty$. This will be justified below.

The time-independent equilibrium temperature $u_\star(x,y)$ satisfies the equation obtained by setting $u_t = 0$ in the evolution equation (11.29), which reduces it to the Poisson equation

$$-\gamma\, \Delta u_\star = F \qquad \text{for} \qquad (x,y) \in \Omega. \tag{11.31}$$

The equilibrium solution is subject to the same inhomogeneous Dirichlet boundary conditions (11.30). Positive definiteness of the Dirichlet boundary value problem implies that

there is a unique equilibrium solution, which can be characterized as the sole minimizer of the associated Dirichlet principle; for details see Section 9.3.

With the equilibrium solution in hand, we let

$$v(t,x,y) = u(t,x,y) - u_\star(x,y)$$

measure the deviation of the dynamical solution u from its eventual equilibrium. By linearity $v(t,x,y)$ satisfies the unforced heat equation subject to homogeneous boundary conditions:

$$v_t = \gamma\,\Delta v, \qquad (x,y) \in \Omega, \qquad\qquad v = 0, \qquad (x,y) \in \partial\Omega. \tag{11.32}$$

Therefore, v can be expanded in an eigenfunction series (11.23), and will decay to zero, $v(t,x,y) \to 0$, at an exponentially fast rate prescribed by the smallest eigenvalue λ_1 of the associated homogeneous Helmholtz boundary value problem. (Special initial data can decay at a faster rate, prescribed by a larger eigenvalue.) Consequently, the solution to the forced inhomogeneous problem (11.29–30) will approach thermal equilibrium,

$$u(t,x,y) = v(t,x,y) + u_\star(x,y) \;\longrightarrow\; u_\star(x,y),$$

at exactly the same exponential rate as its homogeneous counterpart.

The Maximum Principle

Finally, let us state and prove the (Weak) Maximum Principle for the two-dimensional heat equation. As in the one-dimensional situation described in Section 8.3, it states that the maximum temperature in a body that is either insulated or having heat removed from its interior must occur either at the initial time or on its boundary. Observe that there are no conditions imposed on the boundary temperatures.

Theorem 11.1. *Suppose $u(t,x,y)$ is a solution to the forced heat equation*

$$u_t = \gamma\,\Delta u + F(t,x,y), \qquad \text{for} \qquad (x,y) \in \Omega, \quad 0 < t < c,$$

where Ω is a bounded domain, and $\gamma > 0$. Suppose $F(t,x,y) \le 0$ for all $(x,y) \in \overline{\Omega}$ and $0 \le t \le c$. Then the global maximum of u on the set $\{\,(t,x,y)\,|\,(x,y) \in \overline{\Omega},\ 0 \le t \le c\,\}$ occurs either when $t = 0$ or at a boundary point $(x,y) \in \partial\Omega$.

Proof: First, let us prove the result under the assumption that $F(t,x,y) < 0$ everywhere. At a local interior maximum, $u_t = 0$, and, since its Hessian matrix $\nabla^2 u = \begin{pmatrix} u_{xx} & u_{xy} \\ u_{xy} & u_{yy} \end{pmatrix}$ must be negative semi-definite, both diagonal entries $u_{xx}, u_{yy} \le 0$ there. This would imply that $u_t - \gamma\,\Delta u \ge 0$, resulting in a contradiction. If the maximum were to occur when $t = c$, then $u_t \ge 0$ there, and also $u_{xx}, u_{yy} \le 0$, leading again to a contradiction.

To generalize to the case $F(t,x,y) \le 0$, which includes the heat equation when $F(t,x,y) \equiv 0$, set

$$v(t,x,y) = u(t,x,y) + \varepsilon\,(x^2 + y^2), \qquad \text{where} \qquad \varepsilon > 0.$$

Then,

$$\frac{\partial v}{\partial t} = \gamma\,\Delta v - 4\gamma\,\varepsilon + F(t,x,y) = \gamma\,\Delta v + \widetilde{F}(t,x,y),$$

where

$$\widetilde{F}(t,x,y) = F(t,x,y) - 4\gamma\varepsilon < 0.$$

Thus, by the previous paragraph, the maximum of v occurs either when $t = 0$ or at a boundary point $(x,y) \in \partial\Omega$. We then let $\varepsilon \to 0$ and conclude the same for u. More precisely, let $u(t,x,y) \le M$ on $t = 0$ or $(x,y) \in \partial\Omega$. Then

$$v(t,x,y) \le M + C\varepsilon, \qquad \text{where} \qquad C = \max\big\{\, x^2 + y^2 \;\big|\; (x,y) \in \partial\Omega \,\big\} < \infty,$$

since Ω is a bounded domain. Thus,

$$u(t,x,y) \le v(t,x,y) \le M + C\varepsilon.$$

Letting $\varepsilon \to 0$ proves that $u(t,x,y) \le M$ at all $(x,y) \in \overline{\Omega},\ 0 \le t \le c$, which completes the proof. *Q.E.D.*

Remark: The preceding proof can be readily adapted to general diffusion equations (11.12) — assuming that the coefficients σ, κ remain strictly positive throughout the domain.

Exercises

11.1.1. A homogeneous, isotropic circular metal disk of radius 1 meter has its entire boundary insulated. The initial temperature at a point is equal to the distance of the point from the center. Formulate an initial-boundary value problem governing the disk's subsequent temperature dynamics. What is the eventual equilibrium temperature of the disk?

11.1.2. A homogeneous, isotropic, circular metal disk of radius 2 cm has half its boundary fixed at 100° and the other half insulated. Given a prescribed initial temperature distribution, set up the initial-boundary value problem governing its subsequent temperature profile. What is the eventual equilibrium temperature of the disk? Does your answer depend on the initial temperature?

11.1.3. Given the initial temperature distribution $f(x,y) = x\,y\,(1-x)(1-y)$ on the unit square $\Omega = \{\, 0 \le x, y \le 1 \,\}$, determine the equilibrium temperature when subject to homogeneous (*a*) Dirichlet boundary conditions; (*b*) Neumann boundary conditions.

11.1.4. A square plate with side lengths 1 meter has its right and left edges insulated, its top edge held at 100°, and its bottom edge held at 0°. Assuming that the plate is made out of a homogeneous, isotropic material, formulate an appropriate initial-boundary value problem describing the temperature dynamics of the plate. Then find its eventual equilibrium temperature.

11.1.5. A square plate with side lengths 1 meter has initial temperature 5° throughout, and evolves subject to the Neumann boundary conditions $\partial u/\partial \mathbf{n} = 1$ on its entire boundary. What is the eventual equilibrium temperature?

♡ 11.1.6. Let $u(t,x,y)$ be a solution to the heat equation on a bounded domain Ω subject to homogeneous Neumann conditions on its boundary $\partial\Omega$. (*a*) Prove that the total heat $H(t) = \iint_\Omega u(t,x,y)\,dx\,dy$ is conserved, i.e., is constant in time. (*b*) Use part (*a*) to prove that the eventual equilibrium solution is everywhere equal to the average of the initial temperature $u(0,x,y)$. (*c*) What can you say about the behavior of the total heat for the homogeneous Dirichlet boundary value problem? (*d*) What about an inhomogeneous Dirichlet boundary value problem?

11.1.7. Let $u(t, x, y)$ be a nonconstant solution to the heat equation on a connected, bounded domain Ω subject to homogeneous Dirichlet boundary conditions on $\partial\Omega$. (*a*) Prove that its L^2 norm $N(t) = \sqrt{\iint_\Omega u(t,x,y)^2\, dx\, dy}$ is a strictly decreasing function of t. (*b*) Is this also true for mixed boundary conditions? (*c*) For Neumann boundary conditions?

11.1.8. Are the conclusions in Exercises 11.1.6 and 11.1.7 valid for the general diffusion equation (11.12)?

♢ 11.1.9. Write out the eigenvalue equation governing the separable solutions to the general diffusion equation (11.11), subject to appropriate boundary conditions. Given a complete system of eigenfunctions, write down the eigenfunction series solution to the initial value problem $u(0, x, y) = f(x, y)$, including the formulas for the coefficients.

11.1.10. *True or false*: The equilibrium temperature of a fully insulated nonuniform plate whose thermodynamics are governed by the general diffusion equation (11.12) equals the average initial temperature.

11.1.11. Let $\alpha > 0$, and consider the initial-boundary value problem $u_t = \Delta u - \alpha\, u$, $u(0, x, y) = f(x, y)$ on a bounded domain $\Omega \subset \mathbb{R}^2$, with boundary conditions $\partial u/\partial \mathbf{n} = 0$ on $\partial\Omega$.
(*a*) Write the equation in self-adjoint form (9.122). *Hint*: Look at Exercise 9.3.26.
(*b*) Prove that the problem has a unique equilibrium solution.

11.1.12. Write each of the following linear evolution equations in the self-adjoint form (9.122) by choosing suitable inner products and a suitable set of homogeneous boundary conditions. Is the operator you construct positive definite?

$$(a)\ u_t = u_{xx} + u_{yy} - u, \quad (b)\ u_t = y\, u_{xx} + x\, u_{yy}, \quad (c)\ u_t = \Delta^2 u.$$

♢ 11.1.13. Prove that if $f(x, y)$ is continuous and $\iint_R f(x, y)\, dx\, dy = 0$ for all $R \subset \Omega$, then $f(x, y) \equiv 0$ for $(x, y) \in \Omega$. *Hint*: Adapt the method in Exercise 6.1.23.

11.2 Explicit Solutions of the Heat Equation

Solving the two-dimensional heat equation in series form requires knowing the eigenfunctions for the associated Helmholtz boundary value problem. Unfortunately, as with the vast majority of partial differential equations, explicit solution formulas are few and far between. In this section, we discuss two specific cases in which the required eigenfunctions can be found in closed form. The calculations rely on a further separation of variables, which, as we know, works in only a very limited class of domains. Nevertheless, interesting solution features can be gleaned from these particular geometries.

The first example is a rectangular domain, and the eigensolutions can be expressed in terms of elementary functions — trigonometric functions and exponentials. We then study the heating of a circular disk. In this case, the eigenfunctions are no longer elementary functions, but, rather, are expressed in terms of Bessel functions. Understanding their basic properties will require us to take a detour to develop the fundamentals of power series solutions to ordinary differential equations.

Heating of a Rectangle

A homogeneous rectangular plate

$$R = \{\, 0 < x < a,\ 0 < y < b \,\}$$

is heated to a prescribed initial temperature,

$$u(0, x, y) = f(x, y), \qquad \text{for} \qquad (x, y) \in R. \tag{11.33}$$

Then its top and bottom are insulated, while its sides are held at zero temperature. Our task is to understand the thermodynamic evolution of the plate's temperature.

The temperature $u(t, x, y)$ evolves according to the two-dimensional heat equation

$$u_t = \gamma(u_{xx} + u_{yy}), \qquad \text{for} \qquad (x, y) \in R, \qquad t > 0, \tag{11.34}$$

where $\gamma > 0$ is the plate's thermal diffusivity, while subject to homogeneous Dirichlet conditions along the boundary of the rectangle at all subsequent times:

$$u(t, 0, y) = u(t, a, y) = u(t, x, 0) = u(t, x, b) = 0, \quad 0 < x < a, \quad 0 < y < b, \quad t > 0. \tag{11.35}$$

As in (11.19), the eigensolutions to the heat equation are obtained from the usual exponential ansatz $u(t, x, y) = e^{-\lambda t} v(x, y)$. Substituting this expression into the heat equation, we conclude that the function $v(x, y)$ solves the Helmholtz eigenvalue problem

$$\gamma(v_{xx} + v_{yy}) + \lambda v = 0, \qquad (x, y) \in R, \tag{11.36}$$

subject to the same homogeneous Dirichlet boundary conditions:

$$v(0, y) = v(a, y) = v(x, 0) = v(x, b) = 0, \qquad 0 < x < a, \quad 0 < y < b. \tag{11.37}$$

To tackle the rectangular Helmholtz eigenvalue problem (11.36–37), we shall, as in (4.89), introduce a further separation of variables, writing the solution

$$v(x, y) = p(x)\, q(y)$$

as the product of functions depending on the individual Cartesian coordinates. Substituting this expression into the Helmholtz equation (11.36), we find

$$\gamma\, p''(x)\, q(y) + \gamma\, p(x)\, q''(y) + \lambda\, p(x)\, q(y) = 0.$$

To effect the variable separation, we collect all terms involving x on one side and all terms involving y on the other side of the equation, which is accomplished by dividing by $v = pq$ and rearranging the terms:

$$\gamma\, \frac{p''(x)}{p(x)} = -\gamma\, \frac{q''(y)}{q(y)} - \lambda \equiv -\mu.$$

The left-hand side of this equation depends only on x, whereas the middle term depends only on y. As before, this requires that the expressions equal a common *separation constant*, denoted by $-\mu$. (The minus sign is for later convenience.) In this manner, we reduce our partial differential equation to a pair of one-dimensional eigenvalue problems

$$\gamma\, \frac{d^2 p}{dx^2} + \mu\, p = 0, \qquad \gamma\, \frac{d^2 q}{dy^2} + (\lambda - \mu)\, q = 0, \tag{11.38}$$

each of which is subject to homogeneous Dirichlet boundary conditions

$$p(0) = p(a) = 0, \qquad q(0) = q(b) = 0, \tag{11.39}$$

stemming from the boundary conditions (11.37). To obtain a nontrivial separable solution to the Helmholtz equation, we seek nonzero solutions to these two supplementary eigenvalue problems.

We have already solved these particular two boundary value problems (11.38–39) many times; see, for instance, (4.21). The eigenfunctions are, respectively,

$$p_m(x) = \sin \frac{m\,\pi\, x}{a}\,, \qquad m = 1, 2, 3, \ldots, \qquad\qquad q_n(y) = \sin \frac{n\,\pi\, y}{b}\,, \qquad n = 1, 2, 3, \ldots,$$

with

$$\mu = \frac{m^2\,\pi^2\,\gamma}{a^2}\,, \qquad \lambda - \mu = \frac{n^2\,\pi^2\,\gamma}{b^2}\,, \qquad \text{so that} \qquad \lambda = \frac{m^2\,\pi^2\,\gamma}{a^2} + \frac{n^2\,\pi^2\,\gamma}{b^2}\,.$$

Therefore, the separable eigenfunction solutions to the Helmholtz boundary value problem (11.35–36) have the doubly trigonometric form

$$v_{m,n}(x,y) = \sin \frac{m\,\pi\, x}{a}\ \sin \frac{n\,\pi\, y}{b}\,, \qquad \text{for} \qquad m, n = 1, 2, 3, \ \ldots\ , \tag{11.40}$$

with associated eigenvalues

$$\lambda_{m,n} = \frac{m^2\,\pi^2\,\gamma}{a^2} + \frac{n^2\,\pi^2\,\gamma}{b^2} = \left(\frac{m^2}{a^2} + \frac{n^2}{b^2}\right)\pi^2\,\gamma\,. \tag{11.41}$$

Each of these corresponds to an exponentially decaying eigensolution

$$u_{m,n}(t,x,y) = e^{-\,\lambda_{m,n}\,t}\,v_{m,n}(x,y) = \exp\left[-\left(\frac{m^2}{a^2} + \frac{n^2}{b^2}\right)\pi^2\,\gamma\,t\right]\ \sin \frac{m\,\pi\, x}{a}\ \sin \frac{n\,\pi\, y}{b} \tag{11.42}$$

to the original rectangular Dirichlet boundary value problem for the heat equation.

Using the fact that the univariate sine functions form a complete system, it is not hard to prove, [**120**], that the separable eigenfunction solutions (11.42) are complete, and so there are no non-separable eigenfunctions.† As a consequence, the general solution to the initial-boundary value problem can be expressed as a linear combination

$$u(t,x,y) = \sum_{m,n=1}^{\infty} c_{m,n}\,u_{m,n}(t,x,y) = \sum_{m,n=1}^{\infty} c_{m,n}\,e^{-\,\lambda_{m,n}\,t}\,v_{m,n}(x,y) \tag{11.43}$$

of the eigenmodes. The coefficients $c_{m,n}$ are prescribed by the initial conditions, which take the form of a double Fourier sine series

$$f(x,y) = u(0,x,y) = \sum_{m,n=1}^{\infty} c_{m,n}v_{m,n}(x,y) = \sum_{m,n=1}^{\infty} c_{m,n}\ \sin \frac{m\,\pi\, x}{a}\ \sin \frac{n\,\pi\, y}{b}\,.$$

Self-adjointness of the Laplacian operator coupled with the boundary conditions implies that‡ the eigenfunctions $v_{m,n}(x,y)$ are orthogonal with respect to the L^2 inner product

† This appears to be a general fact, true in all known examples, but I know of no general proof. Theorem 9.47 can be used to establish completeness of the eigenfunctions, but does not guarantee that they can all be constructed by separation of variables.

‡ Technically, orthogonality is guaranteed only when the eigenvalues are distinct: $\lambda_{m,n} \neq \lambda_{k,l}$. However, by a direct computation, one finds that orthogonality continues to hold even when the indicated eigenfunctions are associated with equal eigenvalues. See the final subsection of this chapter for a discussion of when such "accidental degeneracies" arise.

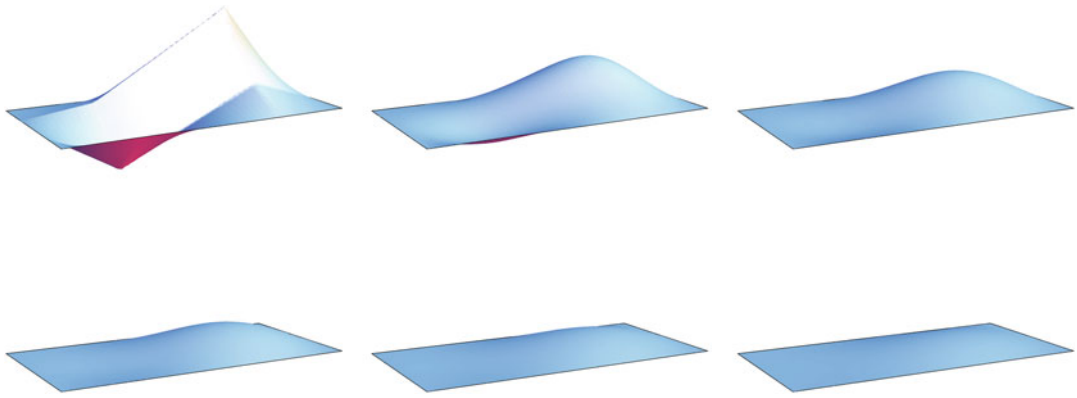


Figure 11.2. Heat diffusion in a rectangle.

on the rectangle:

$$\langle\, v_{k,l}\,, v_{m,n}\,\rangle = \int_0^b \int_0^a v_{k,l}(x,y)\, v_{m,n}(x,y)\, dx\, dy = 0 \qquad \text{unless} \qquad k = m \quad \text{and} \quad l = n.$$

(The skeptical reader can verify the orthogonality relations directly from the eigenfunction formulas (11.40).) Thus, we can appeal to our usual orthogonality formula (11.26) to evaluate the coefficients

$$c_{m,n} = \frac{\langle\, f\,, v_{m,n}\,\rangle}{\|\, v_{m,n}\,\|^2} = \frac{4}{a\,b} \int_0^b \int_0^a f(x,y)\ \sin\frac{m\,\pi\,x}{a}\ \sin\frac{n\,\pi\,y}{b}\, dx\, dy, \tag{11.44}$$

where the formula for the norms of the eigenfunctions

$$\|\, v_{m,n}\,\|^2 = \int_0^b \int_0^a v_{m,n}(x,y)^2\, dx\, dy = \int_0^b \int_0^a \sin^2\frac{m\,\pi\,x}{a}\ \sin^2\frac{n\,\pi\,y}{b}\, dx\, dy = \tfrac{1}{4}\,a\,b \tag{11.45}$$

follows from a direct evaluation of the double integral. Unfortunately, while orthogonality is (mostly) automatic, computation of the norms must inevitably be done "by hand".

For generic initial temperature distributions, the rectangle approaches thermal equilibrium at a rate equal to the smallest eigenvalue:

$$\lambda_{1,1} = \left(\frac{1}{a^2} + \frac{1}{b^2}\right)\pi^2\,\gamma, \tag{11.46}$$

i.e., the sum of the reciprocals of the squared lengths of its sides multiplied by the diffusion coefficient. The larger the rectangle, or the smaller the diffusion coefficient, the smaller the value of $\lambda_{1,1}$, and hence the slower the return to thermal equilibrium. The exponentially fast decay rate of the Fourier series implies that the solution immediately smooths out any discontinuites in the initial temperature profile. Indeed, the higher modes, with m and n large, decay to zero almost instantaneously, and so the solution quickly behaves like a finite sum over a few low-order modes. Assuming that $c_{1,1} \neq 0$, the slowest-decaying mode

in the Fourier series (11.43) is

$$c_{1,1}\, u_{1,1}(t,x,y) = c_{1,1} \exp\left[-\left(\frac{1}{a^2} + \frac{1}{b^2} \right) \pi^2 \gamma\, t \right] \sin \frac{\pi x}{a} \sin \frac{\pi y}{b}\,. \tag{11.47}$$

Thus, in the long run, the temperature becomes entirely of one sign — either positive or negative depending on the sign of $c_{1,1}$ — throughout the rectangle. This observation is, in fact, indicative of the general phenomenon that an eigenfunction associated with the smallest positive eigenvalue of a self-adjoint elliptic operator is necessarily of one sign throughout the domain, [**34**]. A typical solution is plotted at several times in Figure 11.2. Non-generic initial conditions, with $c_{1,1} = 0$, decay more rapidly, and their asymptotic temperature profiles are not of one sign.

Exercises

11.2.1. A rectangle of size 2 cm by 1 cm has initial temperature $f(x,y) = \sin \pi x \sin \pi y$ for $0 \le x \le 2$, $0 \le y \le 1$. All four sides of the rectangle are held at 0°. Assuming that the thermal diffusivity of the plate is $\gamma = 1$, write down a formula for its subsequent temperature $u(t,x,y)$. What is the rate of decay to thermal equilibrium?

11.2.2. Solve Exercise 11.2.1 when the initial temperature $f(x,y)$ is

(a) $x\,y$, (b) $\begin{cases} 1, & 0 < x < 1, \\ 0, & 1 < x < 2; \end{cases}$ (c) $\big(1 - |\,1 - x\,|\big)\big(\tfrac{1}{2} - \big|\,\tfrac{1}{2} - y\,\big|\big)$.

11.2.3. Solve the initial-boundary value problem for the heat equation $u_t = 2\,\Delta u$ on the rectangle $-1 < x < 1$, $0 < y < 1$ when the two short sides are kept at 0°, the two long sides are insulated, and the initial temperature distribution is $u(0,x,y) = \begin{cases} -1, & x < 0, \\ +1, & x > 0, \end{cases}$ $0 < y < 1$.

11.2.4. Answer Exercise 11.2.3 when the two long sides are kept at 0° and the two short sides are insulated.

♡ 11.2.5. A rectangular plate of size 1 meter by 3 meters is made out a metal with unit diffusivity. The plate is taken from a 0° freezer, and, from then on, one of its long sides is heated to 100°, the other is held at 0°, while its top, bottom, and both of the short sides are fully insulated. (*a*) Set up the initial-boundary value problem governing the time-dependent temperature of the plate. (*b*) What is the equilibrium temperature? (*c*) Use your answer from part (*b*) to construct an eigenfunction series for the solution. (*d*) How long until the temperature of the plate is everywhere within 1° of its eventual equilibrium?
Hint: Once t is no longer small, you can approximate the series solution by its first term.

11.2.6. Among all rectangular plates of a prescribed area, which one returns to thermal equilibrium the slowest when subject to Dirichlet boundary conditions? The fastest? Use your physical intuition to explain your answer, but justify it mathematically.

11.2.7. Answer Exercise 11.2.6 for a fully insulated rectangular plate, i.e., subject to Neumann boundary conditions.

♡ 11.2.8. A square metal plate is taken from an oven, and then set out to cool, with its top and bottom insulated. Find the rate of cooling, in terms of the side length and the thermal diffusivity, if (*a*) all four sides are held at 0°; (*b*) one side is insulated and the other three sides are held at 0°; (*c*) two adjacent sides are insulated and the other two are held at 0°; (*d*) two opposite sides are insulated and the other two are held at 0°; (*e*) three sides are insulated and the remaining side is held at 0°. Order the cooling rates of the plates from fastest to slowest. Do your results confirm your intuition?

♡ 11.2.9. Two square plates are made out of the same homogeneous material, and both are initially heated to 100°. All four sides of the first plate are held at 0°, whereas one of the sides of the second plate is insulated while the other three sides are held at 0°. Which plate cools down the fastest? How much faster? Assuming the thermal diffusivity $\gamma = 1$, how long do you have to wait until every point on each plate is within 1° of its equilibrium temperature? *Hint*: Once t is no longer small, the series solution is well approximated by its first term.

♡ 11.2.10. *Multiple choice*: On a unit square that is subject to Dirichlet boundary conditions, the eigenvalues of the Laplace operator are
(*a*) all simple, (*b*) at most double, or (*c*) can have arbitrarily large multiplicity.

♡ 11.2.11. The thermodynamics of a thin circular cylindrical shell of radius a and height h, e.g., the side of a tin can after its top and bottom are removed, is modeled by the heat equation $\dfrac{\partial u}{\partial t} = \gamma \left(\dfrac{1}{a^2} \dfrac{\partial^2 u}{\partial \theta^2} + \dfrac{\partial^2 u}{\partial z^2} \right)$, in which $u(t, \theta, z)$ measures the temperature of the point on the cylinder at time $t > 0$, angle $-\pi < \theta \leq \pi$, and height $0 < z < h$. Keep in mind that $u(t, \theta, z)$ must be a 2π–periodic function of the angular coordinate θ. Assume that the cylinder is everywhere insulated, while its two circular ends at held at 0°. Given an initial temperature distribution at time $t = 0$, write down a series formula for the cylinder's temperature at subsequent times. What is the eventual equilibrium temperature? How fast does the cylinder return to equilibrium?

♡ 11.2.12. Consider the initial-boundary value problem

$$u_t = u_{xx} + u_{yy}, \qquad u(0, x, y) = 0, \qquad 0 < x, y < \pi, \qquad t > 0,$$

for the heat equation in a square subject to the Dirichlet conditions

$$u(0, y) = u(\pi, y) = 0 = u(x, 0), \qquad u(x, \pi) = f(x), \qquad 0 < x, y < \pi.$$

Write out an eigenfunction series formulas for
(*a*) the equilibrium solution $u_\star(x, y) = \lim\limits_{t \to \infty} u(t, x, y)$; (*b*) the solution $u(t, x, y)$.

11.2.13. Solve Exercise 11.2.1 when one long side of the plate is held at 100°.
Hint: See Exercise 11.2.12.

Heating of a Disk — Preliminaries

Let us perform a similar analysis of the thermodynamics of a circular disk. For simplicity (or by choice of suitable physical units), we will assume that the disk

$$D = \{ x^2 + y^2 \leq 1 \} \subset \mathbb{R}^2$$

has unit radius and unit diffusivity $\gamma = 1$. We shall solve the heat equation on D subject to homogeneous Dirichlet boundary values of zero temperature at the circular edge

$$\partial D = C = \{ x^2 + y^2 = 1 \}.$$

Thus, the full initial-boundary value problem is

$$\begin{aligned} &\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\,, && x^2 + y^2 < 1, \\ &u(t, x, y) = 0, && x^2 + y^2 = 1, \\ &u(0, x, y) = f(x, y), && x^2 + y^2 \leq 1. \end{aligned} \qquad t > 0, \tag{11.48}$$

We remark that a simple rescaling of space and time, as outlined in Exercise 11.4.7, can be used to recover the solution for an arbitrary diffusion coefficient and a disk of arbitrary radius from this particular case.

Since we are working in a circular domain, we instinctively pass to polar coordinates (r, θ). In view of the polar coordinate formula (4.105) for the Laplace operator, the heat equation and boundary and initial conditions assume the form

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\,\frac{\partial u}{\partial r} + \frac{1}{r^2}\,\frac{\partial^2 u}{\partial \theta^2}\,, \qquad u(t, 1, \theta) = 0, \qquad u(0, r, \theta) = f(r, \theta), \tag{11.49}$$

where the solution $u(t, r, \theta)$ is defined for all $0 \le r \le 1$ and $t \ge 0$. To ensure that the solution represents a single-valued function on the entire disk, it is required to be a 2π–periodic function of the angular variable:

$$u(t, r, \theta + 2\pi) = u(t, r, \theta).$$

To obtain the separable solutions

$$u(t, r, \theta) = e^{-\lambda t}\, v(r, \theta), \tag{11.50}$$

we need to solve the polar coordinate form of the Helmholtz equation

$$\frac{\partial^2 v}{\partial r^2} + \frac{1}{r}\,\frac{\partial v}{\partial r} + \frac{1}{r^2}\,\frac{\partial^2 v}{\partial \theta^2} + \lambda\, v = 0, \qquad \begin{array}{r} 0 \le r < 1, \\ -\pi < \theta \le \pi, \end{array} \tag{11.51}$$

subject to the boundary conditions

$$v(1, \theta) = 0, \qquad v(r, \theta + 2\pi) = v(r, \theta). \tag{11.52}$$

To solve the polar Helmholtz boundary value problem (11.51–52), we invoke a further separation of variables by writing

$$v(r, \theta) = p(r)\, q(\theta). \tag{11.53}$$

Substituting this ansatz into (11.51), collecting all terms involving r and all terms involving θ, and then equating both to a common separation constant, we are led to the pair of ordinary differential equations

$$r^2\,\frac{d^2 p}{dr^2} + r\,\frac{dp}{dr} + (\lambda r^2 - \mu)\, p = 0, \qquad \frac{d^2 q}{d\theta^2} + \mu\, q = 0, \tag{11.54}$$

where λ is the Helmholtz eigenvalue, and μ the separation constant.

Let us start with the equation for $q(\theta)$. The second boundary condition in (11.52) requires that $q(\theta)$ be 2π–periodic. Therefore, the required solutions are the elementary trigonometric functions

$$q(\theta) = \cos m\theta \qquad \text{or} \qquad \sin m\theta, \qquad \text{where} \qquad \mu = m^2, \tag{11.55}$$

with $m = 0, 1, 2, \dots$ a nonnegative integer.

Substituting the formula for the separation constant, $\mu = m^2$, the differential equation for $p(r)$ takes the form

$$r^2\,\frac{d^2 p}{dr^2} + r\,\frac{dp}{dr} + (\lambda r^2 - m^2)\, p = 0, \qquad 0 \le r \le 1. \tag{11.56}$$

Ordinarily, one imposes two boundary conditions in order to pin down a solution to such a second-order ordinary differential equation. But our Dirichlet condition, namely $p(1) = 0$, specifies its value at only one of the endpoints. The other endpoint is a *singular point* for the ordinary differential equation, because the coefficient of the highest-order derivative, namely r^2, vanishes at $r = 0$. This situation might remind you of our solution to the Euler differential equation (4.111) in the context of separable solutions to the Laplace equation on the disk. As there, we require that the solution be bounded at $r = 0$, and so seek eigensolutions that satisfy the boundary conditions

$$| p(0) | < \infty, \qquad p(1) = 0. \tag{11.57}$$

While (11.56) appears in a variety of applications, it is more challenging than any ordinary differential equation we have encountered so far. Indeed, most solutions cannot be written in terms of the elementary functions (rational functions, trigonometric functions, exponentials, logarithms, etc.) you see in first-year calculus. Nevertheless, owing to their ubiquity in physical applications, its solutions have been extensively studied and tabulated, and so are, in a sense, well known, [**86**, **85**, **119**].

To simplify the subsequent analysis, we make a preliminary rescaling of the independent variable, replacing r by

$$z = \sqrt{\lambda}\, r.$$

(We know the eigenvalue $\lambda > 0$, since we are dealing with a positive definite boundary value problem.) Note that, by the chain rule,

$$\frac{dp}{dr} = \sqrt{\lambda}\, \frac{dp}{dz}, \qquad \frac{d^2p}{dr^2} = \lambda\, \frac{d^2p}{dz^2},$$

and hence

$$r\, \frac{dp}{dr} = z\, \frac{dp}{dz}, \qquad r^2\, \frac{d^2p}{dr^2} = z^2\, \frac{d^2p}{dz^2}.$$

The net effect is to eliminate the eigenvalue parameter λ (or, rather, hide it in the change of variables), so that (11.56) assumes the slightly simpler form

$$z^2\, \frac{d^2p}{dz^2} + z\, \frac{dp}{dz} + (z^2 - m^2)\, p = 0. \tag{11.58}$$

The resulting ordinary differential equation (11.58) is known as *Bessel's equation*, named after the early-nineteenth-century German astronomer Wilhelm Bessel, who first encountered its solutions, now known as *Bessel functions*, in his study of planetary orbits. Special cases had already appeared in the investigations of Daniel Bernoulli on vibrations of a hanging chain, and in those of Fourier on the thermodynamics of a cylindrical body. To make further progress, we need to take time out to study their basic properties, and this will require us to develop the method of power series solutions of ordinary differential equations. With this in hand, we can then return to complete our solution to the heat equation on a disk.

11.3 Series Solutions of Ordinary Differential Equations

When confronted with a novel ordinary differential equation, we have several available options for deriving and understanding its solutions. For instance, the "look-up" method

relies on published handbooks. One of the most useful references that collects many solved differential equations is the classic German compendium by Kamke, [**62**]. Two more recent English-language handbooks are [**93**, **127**]. In addition, many symbolic computer algebra programs, including MATHEMATICA and MAPLE, will produce solutions, when expressible in terms of both elementary and special functions, to a wide range of differential equations.

Of course, use of numerical integration to approximate solutions, [**24**, **60**, **80**], is always an option. Numerical methods do, however, have their limitations, and are best accompanied by some understanding of the underlying theory, coupled with qualitative or quantitative expectations of how the solutions should behave. Furthermore, numerical methods provide less than adequate insight into the nature of the special functions that regularly appear as solutions of the particular differential equations arising in separation of variables. A numerical approximation cannot, in itself, establish rigorous mathematical properties of the solutions of the differential equation.

A more classical means of constructing and approximating the solutions of differential equations is based on their power series expansions, a.k.a. Taylor series. The Taylor expansion of a solution at a point x_0 is found by substituting a general power series into the differential equation and equating coefficients of the various powers of $x - x_0$. The initial conditions at x_0 serve to uniquely determine the coefficients and hence all the derivatives of the solution at the initial point. The Taylor expansion of a special function is an effective tool for deducing some of its key properties, as well as providing a means of computing reasonable numerical approximations to its values within the radius of convergence of the series. (However, serious numerical computations more often rely on nonconvergent asymptotic expansions, [**85**].)

In this section, we provide a brief introduction to the basic series solution techniques for ordinary differential equations, concentrating on second-order linear differential equations, since these form by far the most important class of examples arising in applications. At a regular point, the method will produce a standard Taylor expansion for the solution, while so-called regular singular points require a slightly more general type of series expansion. Generalizations to irregular singular points, higher-order equations, nonlinear equations, and even linear and nonlinear systems are deferred to more advanced texts, including [**54**, **59**].

The Gamma Function

Before delving into the machinery of series solutions and special functions, we need to introduce the gamma function, which effectively generalizes the factorial operation to noninteger. Recall that the *factorial* of a nonnegative integer $n \geq 0$ is defined inductively by the iterative formula

$$n\,! = n \cdot (n-1)\,!, \qquad \text{starting with} \qquad 0\,! = 1. \tag{11.59}$$

When n is a positive integer, the iteration terminates, yielding the familiar expression

$$n\,! = n(n-1)(n-2)\,\cdots\,3 \cdot 2 \cdot 1. \tag{11.60}$$

However, for more general values of n, the iteration never stops, and it cannot be used to compute its factorial. Our goal is to circumvent this difficulty, and introduce a function $f(x)$ that is defined for *all* values of x and will play the role of such a factorial. First,

mimicking (11.59), the function should satisfy the functional equation

$$f(x) = x\, f(x-1) \tag{11.61}$$

where defined. If, in addition, $f(0) = 1$, then we know that $f(n) = n!$ whenever n is a nonnegative integer, and hence such a function will extend the definition of the factorial to more general real and complex numbers.

A moment's thought should convince the reader that there are many possible ways to construct such a function; see Exercise 11.3.6 for a nonstandard example. The most important version is due to Euler. The modern definition of Euler's gamma function relies on an integral formula discovered by the eighteenth-century French mathematician Adrien–Marie Legendre, who will play a starring role in Chapter 12.

Definition 11.2. The *gamma function* is defined by

$$\Gamma(x) = \int_0^\infty e^{-t}\, t^{x-1}\, dt. \tag{11.62}$$

The first fact is that, for real x, the gamma function integral converges only when $x > 0$; otherwise the singularity of t^{x-1} at $t = 0$ is too severe. The key property that turns the gamma function into a substitute for the factorial function relies on an elementary integration by parts:

$$\Gamma(x+1) = \int_0^\infty e^{-t}\, t^x\, dt = -e^{-t}\, t^x \,\Big|_{t=0}^{\infty} + x \int_0^\infty e^{-t}\, t^{x-1}\, dt.$$

The boundary terms vanish whenever $x > 0$, while the final integral is merely $\Gamma(x)$. Therefore, the gamma function satisfies the *recurrence relation*

$$\Gamma(x+1) = x\,\Gamma(x). \tag{11.63}$$

If we set $f(x) = \Gamma(x+1)$, then (11.63) becomes (11.61). Moreover, by direct integration,

$$\Gamma(1) = \int_0^\infty e^{-t}\, dt = 1.$$

Combining this with the recurrence relation (11.63), we deduce that

$$\Gamma(n+1) = n! \tag{11.64}$$

whenever $n \geq 0$ is a nonnegative integer. Therefore, we can identify $x!$ with the value $\Gamma(x+1)$ whenever $x > -1$ is *any* real number.

Remark: The reader may legitimately ask why not replace t^{x-1} by t^x in the definition of $\Gamma(z)$, which would avoid the $n+1$ in (11.64). There is no good answer; we are merely following a well-established precedent set by Legendre and enshrined in all subsequent works.

Thus, at integer values of x, the gamma function agrees with the elementary factorial. A few other values can be computed exactly. One important case is at $x = \frac{1}{2}$. Using the substitution $t = s^2$, with $dt = 2\,s\,ds$, we obtain

$$\Gamma\left(\tfrac{1}{2}\right) = \int_0^\infty e^{-t}\, t^{-1/2}\, dt = \int_0^\infty 2\, e^{-s^2}\, ds = \sqrt{\pi}, \tag{11.65}$$

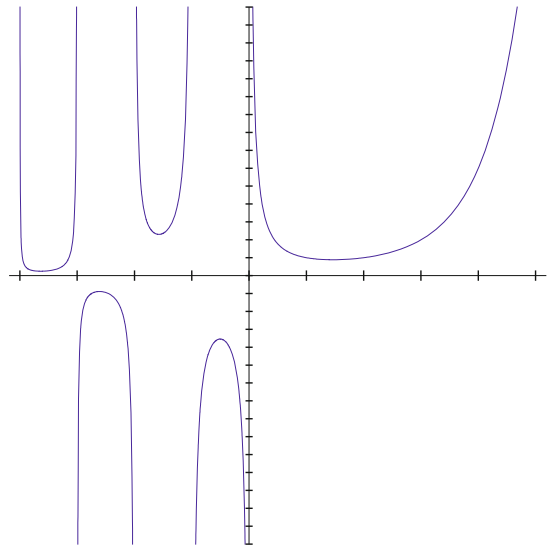


Figure 11.3. The gamma function.

where the final integral was evaluated in (2.100). Thus, using the identification with the factorial function, we identify this value with $\left(-\frac{1}{2}\right)! = \sqrt{\pi}$. The recurrence relation (11.63) will then produce the value of the gamma function at all half-integers $\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \dots$. For example,

$$\Gamma\left(\tfrac{3}{2}\right) = \tfrac{1}{2}\,\Gamma\left(\tfrac{1}{2}\right) = \tfrac{1}{2}\,\sqrt{\pi}, \tag{11.66}$$

and hence $\frac{1}{2}! = \frac{1}{2}\sqrt{\pi}$. The recurrence relation can also be employed to extend the definition of $\Gamma(x)$ to (most) negative values of x. For example, setting $x = -\frac{1}{2}$ in (11.63), we have

$$\Gamma\left(\tfrac{1}{2}\right) = -\tfrac{1}{2}\,\Gamma\left(-\tfrac{1}{2}\right), \qquad \text{so} \qquad \Gamma\left(-\tfrac{1}{2}\right) = -2\,\Gamma\left(\tfrac{1}{2}\right) = -2\sqrt{\pi}\,.$$

The only points at which this device fails are the negative integers, and indeed, $\Gamma(x)$ has a singularity when $x = -1, -2, -3, \dots$. A graph[†] of the gamma function is displayed in Figure 11.3.

Remark: Most special functions of importance for applications arise as solutions to fairly simple ordinary differential equations. The gamma function is a significant exception. Indeed, it can be proved, [**11**], that the gamma function *does not* satisfy *any* algebraic differential equation!

Regular Points

We are now ready to develop the method of series solutions to ordinary differential equations. Before we proceed to develop the general computational machinery, a naïve calculation in an elementary example will be enlightening.

[†] The axes are at different scales; the tick marks are at integer values.

Example 11.3. Consider the initial value problem

$$\frac{d^2u}{dx^2} + u = 0, \qquad u(0) = 1, \qquad u'(0) = 0. \tag{11.67}$$

Let us investigate whether we can construct an analytic solution in the form of a convergent power series

$$u(x) = u_0 + u_1 x + u_2 x^2 + u_3 x^3 + \cdots = \sum_{n=0}^{\infty} u_n x^n \tag{11.68}$$

that is based at the initial point $x_0 = 0$. Term-by-term differentiation yields the following series expansions† for its derivatives:

$$\begin{aligned} \frac{du}{dx} &= u_1 + 2u_2 x + 3u_3 x^2 + 4u_4 x^3 + \cdots &&= \sum_{n=0}^{\infty} (n+1)\, u_{n+1} x^n, \\ \frac{d^2u}{dx^2} &= 2u_2 + 6u_3 x + 12u_4 x^2 + 20u_5 x^3 + \cdots &&= \sum_{n=0}^{\infty} (n+1)(n+2)\, u_{n+2} x^n. \end{aligned} \tag{11.69}$$

The next step is to substitute the series (11.68–69) into the differential equation and collect common powers of x:

$$\frac{d^2u}{dx^2} + u = (2u_2 + u_0) + (6u_3 + u_1)x + (12u_4 + u_2)x^2 + (20u_5 + u_3)x^3 + \cdots = 0.$$

At this point, one focuses attention on the individual coefficients, appealing to the following basic observation:

Two convergent power series are equal if and only if all their coefficients are equal.

In particular, a power series represents the zero function‡ if and only if all its coefficients are 0. In this manner we obtain the following infinite sequence of algebraic *recurrence relations* among the coefficients:

$$\begin{aligned} 1 \qquad & 2u_2 + u_0 = 0, \\ x \qquad & 6u_3 + u_1 = 0, \\ x^2 \qquad & 12u_4 + u_2 = 0, \\ x^3 \qquad & 20u_5 + u_3 = 0, \\ x^4 \qquad & 30u_6 + u_4 = 0, \\ \vdots \qquad & \vdots \\ x^n \qquad & (n+1)(n+2)\, u_{n+2} + u_n = 0. \end{aligned} \tag{11.70}$$

Now, the initial conditions serve to prescribe the first two coefficients:

$$u(0) = u_0 = 1, \qquad u'(0) = u_1 = 0.$$

---

† When working with the series in summation form, it helps to re-index in order to display the term of degree n.

‡ Here it is essential that we work with analytic functions, since this result is *not* true for C^∞ functions! For example, the function e^{-1/x^2} has identically zero power series at $x_0 = 0$; see Exercise 11.3.21.

We then solve the recurrence relations in order: The first determines $u_2 = -\frac{1}{2}\,u_0 = -\frac{1}{2}$; the second, $u_3 = -\frac{1}{6}\,u_1 = 0$; next, $u_4 = -\frac{1}{12}\,u_2 = \frac{1}{24}$; then $u_5 = -\frac{1}{20}\,u_3 = 0$; then $u_6 = -\frac{1}{30}\,u_4 = -\frac{1}{720}$; and so on. In general, it is not hard to see that

$$u_{2k} = \frac{(-1)^k}{(2\,k)\,!}\,, \qquad u_{2k+1} = 0, \qquad k = 0, 1, 2, \ \dots \ .$$

Hence, the required series solution is

$$u(x) = 1 - \tfrac{1}{2}\,x^2 + \tfrac{1}{24}\,x^3 - \tfrac{1}{720}\,x^6 + \ \cdots \ = \sum_{k=0}^{\infty} \frac{(-1)^k}{k\,!}\,x^{2k},$$

which, by the ratio test, converges for all x. We have thus recovered the well-known Taylor series for $\cos x$, which is indeed the solution to the initial value problem. Changing the initial conditions to $u(0) = u_0 = 0$, $u'(0) = u_1 = 1$, will similarly produce the usual Taylor expansion of $\sin x$. Note that the generation of the Taylor series does not rely on any a priori knowledge of trigonometric functions or the direct solution method for linear constant-coefficient ordinary differential equations.

Building on this experience, let us describe the general method. We shall concentrate on solving a second-order homogeneous linear differential equation

$$p(x)\,\frac{d^2u}{dx^2} + q(x)\,\frac{du}{dx} + r(x)\,u = 0. \tag{11.71}$$

The coefficients $p(x), q(x), r(x)$ are assumed to be analytic functions on some common domain. This means that, at a point x_0 within the domain, they admit convergent power series expansions

$$\begin{aligned} p(x) &= p_0 + p_1\,(x - x_0) + p_2\,(x - x_0)^2 + \ \cdots \ , \\ q(x) &= q_0 + q_1\,(x - x_0) + q_2\,(x - x_0)^2 + \ \cdots \ , \\ r(x) &= r_0 + r_1\,(x - x_0) + r_2\,(x - x_0)^2 + \ \cdots \ . \end{aligned} \tag{11.72}$$

We expect that solutions to the differential equation are also analytic. This expectation is justified, provided that the equation is *regular* at the point x_0, in the following sense.

Definition 11.4. A point $x = x_0$ is a *regular point* of a second-order linear ordinary differential equation (11.71) if the leading coefficient does not vanish there:

$$p_0 = p(x_0) \neq 0.$$

A point where $p(x_0) = 0$ is known as a *singular point*.

In short, at a regular point, the second-order derivative term does not disappear, and so the equation is "genuinely" of second order.

Remark: The definition of a singular point assumes that the other two coefficients do not also vanish there, so that either $q(x_0) \neq 0$ or $r(x_0) \neq 0$. If all three functions happen to vanish at x_0, we can cancel any common factor $(x - x_0)^k$, and hence, without loss of generality, assume that at least one of the coefficient functions is nonzero at x_0.

Proofs of the basic existence theorem for differential equations at regular points can be found in [**18**, **54**, **59**].

Theorem 11.5. *Let x_0 be a regular point for the second-order homogeneous linear ordinary differential equation* (11.71). *Then there exists a unique, analytic solution $u(x)$ to the initial value problem*

$$u(x_0) = a, \qquad u'(x_0) = b. \tag{11.73}$$

The radius of convergence of the power series for $u(x)$ is at least as large as the distance from the regular point x_0 to the nearest singular point of the differential equation in the complex plane.

Thus, every solution to an analytic differential equation at a regular point x_0 can be expanded in a convergent power series

$$u(x) = u_0 + u_1(x - x_0) + u_2(x - x_0)^2 + \cdots = \sum_{n=0}^{\infty} u_n (x - x_0)^n. \tag{11.74}$$

Since the power series necessarily coincides with the Taylor series for $u(x)$, its coefficients[†]

$$u_n = \frac{u^{(n)}(x_0)}{n!}$$

are multiples of the derivatives of the function at the point x_0. In particular, the first two coefficients,

$$u_0 = u(x_0) = a, \qquad u_1 = u'(x_0) = b, \tag{11.75}$$

are fixed by the initial conditions. The remaining coefficients will then be uniquely prescribed thanks to the uniqueness of solutions to initial value problems.

Near a regular point, the second-order differential equation (11.71) admits two linearly independent analytic solutions, which we denote by $\widehat{u}(x)$ and $\widetilde{u}(x)$. The general solution can be written as a linear combination of the two basis solutions:

$$u(x) = a\,\widehat{u}(x) + b\,\widetilde{u}(x). \tag{11.76}$$

A convenient choice is to have the first satisfy the initial conditions

$$\widehat{u}(x_0) = 1, \qquad \widehat{u}'(x_0) = 0, \tag{11.77}$$

and the second satisfy

$$\widetilde{u}(x_0) = 0, \qquad \widetilde{u}'(x_0) = 1, \tag{11.78}$$

although other conventions may be used depending on the circumstances. Given (11.77–78), the linear combination (11.76) automatically satisfies the initial conditions (11.73).

The basic computational strategy to construct the power series solution to the initial value problem is a straightforward adaptation of the method used in Example 11.3. One substitutes the known power series (11.72) for the coefficient functions and the unknown power series (11.74) for the solution into the differential equation (11.71). Multiplying out the formulas and collecting the common powers of $x - x_0$ will result in a (complicated) power series whose individual coefficients must be equated to zero. The lowest-order terms are multiples of $(x - x_0)^0 = 1$, i.e., the constant terms. They produce a linear relation

$$u_2 = R_2(u_0, u_1) = R_2(a, b)$$

[†] Some authors prefer to include the $n!$'s in the original power series; this is purely a matter of personal taste.

that prescribes the coefficient u_2 in terms of the initial data (11.75). The coefficient of $(x - x_0)$ leads to a relation

$$u_3 = R_3(u_0, u_1, u_2) = R_3(a, b, R_2(a, b))$$

that prescribes u_3 in terms of the initial data and the previously computed coefficient u_2. And so on. At the n^{th} stage of the procedure, the coefficient of $(x - x_0)^n$ produces the linear *recurrence relation*

$$u_{n+2} = R_n(u_0, u_1, \dots, u_{n+1}), \qquad n = 0, 1, 2, \dots , \tag{11.79}$$

that will prescribe the $(n + 2)^{\text{nd}}$ order coefficient in terms of the previously computed coefficients. In this fashion, we will have constructed a formal power series solution to the differential equation at a regular point. The one remaining issue is whether the resulting power series actually converges. The full analysis can be found in [**54**, **59**], and will serve to complete the proof of the general Existence Theorem 11.5.

Rather than continue on in general, the best way to learn the method is to work through another, less trivial, example.

The Airy Equation

We will illustrate the procedure by constructing power series solutions to the *Airy equation*

$$\frac{d^2u}{dx^2} = x\,u. \tag{11.80}$$

This second-order linear ordinary differential equation, which arises in applications to optics, rainbows, and dispersive waves, has solutions that cannot be expressed in terms of elementary functions.

For the Airy equation (11.80), the leading coefficient is constant, and so every point is a regular point. For simplicity, we will look only for power series based at the origin $x_0 = 0$, and therefore of the form (11.68). Equating the two series

$$\begin{aligned} u''(x) = 2\,u_2 + 6\,u_3\,x + 12\,u_4\,x^2 + 20\,u_5 x^3 + \cdots &= \sum_{n=0}^{\infty} (n+1)(n+2)\,u_{n+2}\,x^n, \\ x\,u(x) = u_0\,x + u_1\,x^2 + u_2\,x^3 + \cdots &= \sum_{n=1}^{\infty} u_{n-1}\,x^n, \end{aligned}$$

leads to the following recurrence relations relating the coefficients:

$$\begin{aligned} 1 \qquad & 2\,u_2 = 0, \\ x \qquad & 6\,u_3 = u_0, \\ x^2 \qquad & 12\,u_4 = u_1, \\ x^3 \qquad & 20\,u_5 = u_2, \\ x^4 \qquad & 30\,u_6 = u_3, \\ \vdots \qquad & \vdots \\ x^n \qquad & (n+1)(n+2)\,u_{n+2} = u_{n-1}. \end{aligned}$$

As before, we solve them in order: The first equation determines u_2. The second prescribes $u_3 = \frac{1}{6} u_0$ in terms of u_0. Next, we find $u_4 = \frac{1}{12} u_1$ in terms of u_1, followed by $u_5 = \frac{1}{20} u_2 = 0$; then $u_6 = \frac{1}{30} u_3 = \frac{1}{180} u_0$ is first given in terms of u_3, but we already know the latter in terms of u_0. And so on.

Let us now construct two basis solutions. The first has the initial conditions

$$u_0 = \widehat{u}(0) = 1, \qquad u_1 = \widehat{u}'(0) = 0.$$

The recurrence relations imply that the only nonzero coefficients c_n occur when $n = 3k$ is a multiple of 3. Moreover,

$$u_{3k} = \frac{u_{3k-3}}{3k(3k-1)}.$$

A straightforward induction proves that

$$u_{3k} = \frac{1}{3k(3k-1)(3k-3)(3k-4)\cdots 6\cdot 5\cdot 3\cdot 2}.$$

The resulting solution is

$$\widehat{u}(x) = 1 + \tfrac{1}{6}x^3 + \tfrac{1}{180}x^6 + \cdots = 1 + \sum_{k=1}^{\infty} \frac{x^{3k}}{3k(3k-1)(3k-3)(3k-4)\cdots 6\cdot 5\cdot 3\cdot 2}. \tag{11.81}$$

Note that the denominator is similar to a factorial, except every third term is omitted. A straightforward application of the ratio test confirms that the series converges for all (complex) x, in conformity with the general Theorem 11.5, which guarantees an infinite radius of convergence because the Airy equation has no singular points.

Similarly, starting with the initial conditions

$$u_0 = \widetilde{u}(0) = 0, \qquad u_1 = \widetilde{u}'(0) = 1,$$

we find that the only nonzero coefficients u_n occur when $n = 3k+1$. The recurrence relation

$$u_{3k+1} = \frac{u_{3k-2}}{(3k+1)(3k)} \quad \text{yields} \quad u_{3k+1} = \frac{1}{(3k+1)(3k)(3k-2)(3k-3)\cdots 7\cdot 6\cdot 4\cdot 3}.$$

The resulting solution is

$$\widetilde{u}(x) = x + \tfrac{1}{12}x^4 + \tfrac{1}{504}x^7 + \cdots = x + \sum_{k=1}^{\infty} \frac{x^{3k+1}}{(3k+1)(3k)(3k-2)(3k-3)\cdots 7\cdot 6\cdot 4\cdot 3}. \tag{11.82}$$

Again, the denominator skips every third term in the product. Every solution to the Airy equation can be written as a linear combination of these two basis power series solutions:

$$u(x) = a\,\widehat{u}(x) + b\,\widetilde{u}(x), \qquad \text{where} \qquad a = u(0), \qquad b = u'(0).$$

Both power series (11.81, 82), converge quite rapidly, and so the first few terms will provide a reasonable approximation to the solutions for moderate values of x.

We have, in fact, already encountered another solution to the Airy equation. According to formula (8.97), the integral

$$\mathrm{Ai}(x) = \frac{1}{\pi}\int_0^{\infty} \cos\bigl(s\,x + \tfrac{1}{3}s^3\bigr)\,ds \tag{11.83}$$

defines the *Airy function of the first kind*. Let us prove that it satisfies the Airy differential equation (11.80):

$$\frac{d^2}{dx^2}\,\mathrm{Ai}(x) = x\,\mathrm{Ai}(x).$$

Before differentiating, we recall the integration by parts argument in (8.96) to re-express the Airy integral in absolutely convergent form:

$$\mathrm{Ai}(x) = \frac{2}{\pi}\int_0^\infty \frac{s\,\sin\bigl(s\,x + \frac{1}{3}\,s^3\bigr)}{(x+s^2)^2}\,ds.$$

We are now permitted to differentiate under the integral sign, producing (after some algebra)

$$\frac{d^2}{dx^2}\,\mathrm{Ai}(x) - x\,\mathrm{Ai}(x) = \frac{2}{\pi}\int_0^\infty \frac{d}{ds}\left[\frac{s\,(x+s^2)\cos\bigl(s\,x+\frac{1}{3}\,s^3\bigr) - \sin\bigl(s\,x+\frac{1}{3}\,s^3\bigr)}{(x+s^2)^3}\right] ds = 0.$$

Thus, the Airy function must be a certain linear combination of the two basic series solutions:

$$\mathrm{Ai}(x) = \mathrm{Ai}(0)\,\widehat{u}(x) + \mathrm{Ai}'(0)\,\widetilde{u}(x).$$

Its values at $x = 0$ are, in fact, given by

$$\begin{aligned}
\mathrm{Ai}(0) &= \frac{1}{\pi}\int_0^\infty \cos\bigl(\tfrac{1}{3}\,s^3\bigr)\,ds = \frac{\Gamma\bigl(\frac{1}{3}\bigr)}{2\,\pi\,3^{1/6}} = \frac{1}{3^{2/3}\,\Gamma\bigl(\frac{2}{3}\bigr)} \approx .355028\,,\\
\mathrm{Ai}'(0) &= -\,\frac{1}{\pi}\int_0^\infty s\,\sin\bigl(\tfrac{1}{3}\,s^3\bigr)\,ds = -\,\frac{3^{1/6}\,\Gamma\bigl(\frac{2}{3}\bigr)}{2\,\pi} = -\,\frac{1}{3^{1/3}\,\Gamma\bigl(\frac{1}{3}\bigr)} \approx -.258819.
\end{aligned} \tag{11.84}$$

The second and third expressions involve the gamma function (11.62); a proof, based on complex integration, can be found in [**85**; p. 54].

Exercises

11.3.1. Find (*a*) $\Gamma\bigl(\frac{5}{2}\bigr)$, (*b*) $\Gamma\bigl(\frac{7}{2}\bigr)$, (*c*) $\Gamma\bigl(-\frac{3}{2}\bigr)$, (*d*) $\Gamma\bigl(-\frac{5}{2}\bigr)$.

11.3.2. Prove that $\Gamma\bigl(n+\frac{1}{2}\bigr) = \dfrac{\sqrt{\pi}\,(2n)!}{2^{2n}\,n!}$ for every positive integer n.

11.3.3. Let $x \in \mathbb{C}$ be complex. (*a*) Prove that the gamma function integral (11.62) converges, provided $\mathrm{Re}\,x > 0$. (*b*) Is formula (11.63) valid when x is complex?

♢ 11.3.4. Prove that $\Gamma(x) = \displaystyle\int_0^1 (-\log s)^{x-1}\,ds$, and hence, for $0 \le n \in \mathbb{Z}$, we have $n\,! = \displaystyle\int_0^1 (-\log s)^n\,ds$. *Remark*: Euler first established the latter identity directly, and used it to define the gamma function.

11.3.5. Evaluate $\displaystyle\int_0^\infty \sqrt{x}\,e^{-x^3}\,dx$.

♢ 11.3.6. Can you construct a function $f(x)$ that satisfies the factorial functional equation (11.61) and has the values $f(x) = 1$ for $0 \le x \le 1$? If so, is $f(x) = \Gamma(x+1)$?

11.3.7. Explain how to construct the power series for $\sin x$ by solving the differential equation (11.67).

11.3.8. Construct two independent power series solutions to the Euler equation $x^2 u'' - 2u = 0$ based at the point $x_0 = 1$.

11.3.9. Construct two independent power series solutions to the equation $u'' + x^2 u = 0$ based at the point $x_0 = 0$.

11.3.10. Consider the ordinary differential equation $u'' + 2x\,u' + 2u = 0$. (*a*) Find two linearly independent power series solutions in powers of x. (*b*) What is the radius of convergence of your power series? (*c*) By inspection of your series, find one solution to the equation expressible in terms of elementary functions. (*d*) Find an explicit (non-series) formula for the second independent power series solution.

11.3.11. Answer Exercise 11.3.10 for the equation $u'' + \frac{1}{2}x\,u' - \frac{1}{2}u = 0$, which is a special case of equation (8.63).

11.3.12. Consider the ordinary differential equation $u'' + x\,u' + 2u = 0$. (a) Find two linearly independent power series solutions based at $x_0 = 0$. (*b*) Write down the power series for the solution to the initial value problem $u(0) = 1$, $u'(0) = -1$. (*c*) What is the radius of convergence of your power series solution in part (a)? Can you justify this by direct inspection of your power series?

♢ 11.3.13. The *Hermite equation* of order n is

$$\frac{d^2u}{dx^2} - 2x\,\frac{du}{dx} + 2n\,u = 0. \tag{11.85}$$

Assuming $n \in \mathbb{N}$ is a nonnegative integer: (a) Find two linearly independent power series solutions based at $x_0 = 0$, and then show that one of your solutions is a polynomial of degree n. (*b*) Prove that the Hermite polynomial $H_n(x)$ defined in (8.64) solves the Hermite equation (11.85) and hence is a multiple of the polynomial solution you found in part (a). What is the multiple? (*c*) Prove that the Hermite polynomials are orthogonal with respect to the inner product $\langle\, u\,,v\,\rangle = \displaystyle\int_{-\infty}^{\infty} u(x)\,v(x)\,e^{-x^2}\,dx$.

11.3.14. Use the ratio test to directly determine the radius of convergence of the series solutions (11.81, 82) to the Airy equation.

11.3.15. Write down the general solution to the following ordinary differential equations:
(*a*) $u'' + (x - c)\,u = 0$, where c is a fixed constant;
(*b*) $u'' = \lambda x\,u$, where $\lambda \neq 0$ is a fixed nonzero constant.

♢ 11.3.16. The *Airy function of the second kind* is defined by

$$\mathrm{Bi}(x) = \frac{1}{\pi}\int_0^\infty \Big[\exp\big(s\,x - \tfrac{1}{3}\,s^3\big) + \sin\big(s\,x + \tfrac{1}{3}\,s^3\big)\Big]\,ds. \tag{11.86}$$

(*a*) Prove that $\mathrm{Bi}(x)$ is well defined and a solution to the Airy equation. (*b*) Given that[†]

$$\mathrm{Bi}(0) = \frac{1}{3^{1/6}\Gamma\big(\frac{2}{3}\big)}\,, \qquad \mathrm{Bi}'(0) = \frac{3^{1/6}}{\Gamma\big(\frac{1}{3}\big)}\,, \tag{11.87}$$

explain why every solution to the Airy equation can be written as a linear combination of $\mathrm{Ai}(x)$ and $\mathrm{Bi}(x)$. (*c*) Write the two series solutions (11.81, 82) in terms of $\mathrm{Ai}(x)$ and $\mathrm{Bi}(x)$.

11.3.17. Use the Fourier transform to construct an L^2 solution to the Airy equation. Can you identify your solution?

♢ 11.3.18. Apply separation of variables to the Tricomi equation (4.137), and write down all separable solutions. *Hint*: See Exercise 11.3.15(*b*) and Exercise 11.3.16.

[†] See [**85**; p. 54] for a proof.

♡ 11.3.19. (*a*) Show that $u(x) = \sum_{n=1}^{\infty} (n-1)!\, x^n$ is a power series solution to the first-order linear ordinary differential equation $x^2 u' - u + x = 0$. (*b*) For which x does the series converge? (*c*) Find an analytic formula for the general solution to the equation. (*d*) Find a second-order homogeneous linear ordinary differential equation that has this power series as a (formal) solution. *Remark*: The lesson of this exercise is that not all power series solutions to ordinary differential equations converge. Theorem 11.5 guarantees convergence at a regular point, but in this example the power series is based at the singular point $x_0 = 0$.

11.3.20. *True or false*: The only function $f(x)$ that has identically zero Taylor series is the zero function.

♢ 11.3.21. Define $f(x) = \begin{cases} e^{-1/x^2}, & x \neq 0, \\ 0, & x = 0. \end{cases}$ (*a*) Prove that f is a C^∞ function for all $x \in \mathbb{R}$. (*b*) Prove that $f(x)$ is not analytic by showing that its Taylor series at $x_0 = 0$ does not converge to $f(x)$ when $x \neq 0$.

Regular Singular Points

As we have just seen, constructing power series solutions at regular points is a reasonably straightforward computational exercise: one writes down a power series with arbitrary coefficients, substitutes into the differential equation along with a pair of initial conditions, and recursively solves for the coefficients. Finding a general formula for the coefficients might be challenging, but producing their successive numerical values, degree by degree, is a mechanical exercise.

However, at a singular point, the solutions cannot be typically written as an ordinary power series, and one needs to be cleverer. Of course, you may object — why not just solve the equation away from the singular point and be done with it. But there are multiple reasons not to do this. First, one may be unable to discover a general formula for the power series coefficients at regular points. Second, the most informative and interesting behavior of solutions is typically found at the singular points, and so series solutions based at singular points are particularly enlightening. And finally, one of the boundary conditions required for us to complete our construction of separable solutions to partial differential equations often occurs at a singular point.

Singular points appear in two guises. The easier to handle, and, fortunately, the ones that arise in almost all applications, are known as "regular singular points". Irregular singular points are nastier, and we will not make any attempt to understand them in this text; the curious reader is referred to [**54**, **59**].

Definition 11.6. A second-order linear homogeneous ordinary differential equation that can be written the form

$$(x - x_0)^2\, a(x)\, \frac{d^2u}{dx^2} + (x - x_0)\, b(x)\, \frac{du}{dx} + c(x)\, u = 0, \tag{11.88}$$

where $a(x), b(x)$, and $c(x)$ are analytic at $x = x_0$ and, moreover, $a(x_0) \neq 0$, is said to have a *regular singular point* at x_0.

The simplest example of a second-order equation with a regular singular point at $x_0 = 0$ is the *Euler equation*

$$a\, x^2 u'' + b\, x\, u' + c\, u = 0, \tag{11.89}$$

with a, b, c all constant and $a \neq 0$. Note that all other points are regular points. Euler equations can be readily solved by substituting the power ansatz $u(x) = x^r$. We find

$$a x^2 u'' + b x u' + c u = a r(r-1)x^r + b r x^r + c x^r = 0,$$

provided the exponent r satisfies the *indicial equation*

$$a r (r-1) + b r + c = 0.$$

If this quadratic equation has two distinct roots $r_1 \neq r_2$, we obtain two linearly independent (possibly complex) solutions $\widehat{u}(x) = x^{r_1}$ and $\widetilde{u}(x) = x^{r_2}$. The general solution $u(x) = c_1 x^{r_1} + c_2 x^{r_2}$ is a linear combination thereof. Note that unless r_1 or r_2 is a nonnegative integer, all nonzero solutions have a singularity at the singular point $x = 0$. A repeated root, $r_1 = r_2$, has only one power solution, $\widehat{u}(x) = x^{r_1}$, and requires an additional logarithmic term, $\widetilde{u}(x) = x^{r_1} \log x$, for the second independent solution. In this case, the general solution has the form $u(x) = c_1 x^{r_1} + c_2 x^{r_1} \log x$.

The series solution method at more general regular singular points is modeled on the simple example of the Euler equation. One now seeks a solution that has a series expansion of the form

$$u(x) = (x-x_0)^r \sum_{n=0}^{\infty} u_n (x-x_0)^n = u_0 (x-x_0)^r + u_1 (x-x_0)^{r+1} + u_2 (x-x_0)^{r+2} + \cdots. \quad (11.90)$$

The exponent r is known as the *index*. If $r = 0$, or, more generally, if r is a positive integer, then (11.90) is an ordinary power series, but we allow the possibility of a non-integral, or even complex, index r. We can assume, without any loss of generality, that the leading coefficient $u_0 \neq 0$. Indeed, if $u_k \neq 0$ is the first nonzero coefficient, then the series begins with the term $u_k (x - x_0)^{r+k}$, and we merely replace r by $r + k$ to write it in the form (11.90). Since any scalar multiple of a solution is a solution, we can further assume that $u_0 = 1$, in which case we call (11.90) a *normalized Frobenius series* in honor of the German mathematician Georg Frobenius, who systematically established the calculus of series solutions at regular singular points in the late 1800s. The index r, and the higher-order coefficients $u_1, u_2, \ldots$, are then found by substituting the normalized Frobenius series into the differential equation (11.88) and equating the coefficients of the powers of $x - x_0$ to zero.

Warning: Unlike those in ordinary power series expansions, the coefficients $u_0 = 1$ and u_1 are *not* prescribed by the initial conditions at the point x_0.

Since

$$\begin{aligned} u(x) &= (x - x_0)^r + u_1 (x - x_0)^{r+1} + \cdots, \\ (x - x_0)\, u'(x) &= r\, (x - x_0)^r + (r+1) u_1 (x - x_0)^{r+1} + \cdots, \\ (x - x_0)^2\, u''(x) &= r\, (r-1)\, (x - x_0)^r + (r+1)\, r\, u_1 (x - x_0)^{r+1} + \cdots, \end{aligned}$$

the terms of lowest order in the equation are multiples of $(x - x_0)^r$. Equating their coefficients to zero produces a quadratic equation of the form

$$s_0\, r\, (r-1) + t_0\, r + r_0 = 0, \quad (11.91)$$

where

$$s_0 = s(x_0) = \tfrac{1}{2} p''(x_0), \qquad t_0 = t(x_0) = q'(x_0), \qquad r_0 = r(x_0),$$

are the leading coefficients in the power series expansions of the individual coefficient functions. The quadratic equation (11.91) is known as the *indicial equation*, since it determines the possible indices r in the Frobenius expansion (11.90) of a solution.

As with the Euler equation, the quadratic indicial equation usually has two roots, say r_1 and r_2, which provide two allowable indices, and one thus expects to find two independent Frobenius expansions. Usually, this expectation is realized, but there is an important exception. The general result is summarized in the following list:

(*i*) If $r_2 - r_1$ is not an integer, then there are two linearly independent solutions $\widehat{u}(x)$ and $\widetilde{u}(x)$, each having convergent normalized Frobenius expansions of the form (11.90).

(*ii*) If $r_1 = r_2$, then there is only one solution $\widehat{u}(x)$ with a normalized Frobenius expansion (11.90). One can construct a second independent solution of the form

$$\widetilde{u}(x) = \log(x - x_0)\, \widehat{u}(x) + v(x), \qquad \text{where} \qquad v(x) = \sum_{n=1}^{\infty} v_n (x - x_0)^{n+r_1} \tag{11.92}$$

is a convergent Frobenius series.

(*iii*) Finally, if $r_2 = r_1 + k$, where $k > 0$ is a positive integer, then there is a nonzero solution $\widehat{u}(x)$ with a convergent Frobenius expansion corresponding to the smaller index r_1. One can construct a second independent solution of the form

$$\widetilde{u}(x) = c \log(x - x_0)\, \widehat{u}(x) + v(x), \qquad \text{where} \qquad v(x) = x^{r_2} + \sum_{n=1}^{\infty} v_n (x - x_0)^{n+r_2} \tag{11.93}$$

is a convergent Frobenius series, and c is a constant, which may be 0, in which case the second solution $\widetilde{u}(x)$ is also of Frobenius form.

Thus, in every case, the differential equation has at least one nonzero solution with a convergent Frobenius expansion. If the second independent solution does not have a Frobenius expansion, then it requires an additional logarithmic term of a well-prescribed form. Rather than try to develop the general theory in any more detail here, we will content ourselves to work through a couple of particular examples.

Example 11.7. Consider the second-order ordinary differential equation

$$\frac{d^2u}{dx^2} + \left(\frac{1}{x} + \frac{x}{2} \right) \frac{du}{dx} + u = 0. \tag{11.94}$$

We look for series solutions based at $x = 0$. Note that, upon multiplying by x^2, the equation takes the form

$$x^2 u'' + x\big(1 + \tfrac{1}{2}x^2\big)u' + x^2 u = 0,$$

and hence $x_0 = 0$ is a regular singular point, with $a(x) = 1$, $b(x) = 1 + \frac{1}{2}x^2$, $c(x) = x^2$. We thus look for a solution that can be represented by a Frobenius expansion:

$$\begin{aligned}
u(x) &= x^r + u_1 x^{r+1} + \cdots + u_n x^{n+r} + \cdots, \\
x\, u'(x) &= r\, x^r + (r+1) u_1 x^{r+1} + \cdots + (n+r) u_n x^{n+r} + \cdots, \\
\tfrac{1}{2} x^3 u'(x) &= \tfrac{1}{2} r\, x^{r+2} + \tfrac{1}{2}(r+1) u_1 x^{r+3} + \cdots + \tfrac{1}{2}(n+r-2)\, u_{n-2} x^{n+r} + \cdots, \\
x^2 u''(x) &= r(r-1)\, x^r + (r+1)\, r\, u_1 x^{r+1} + \cdots + (n+r)(n+r-1) u_n x^{n+r} + \cdots.
\end{aligned} \tag{11.95}$$

Substituting into the differential equation, we find that the coefficient of x^r leads to the indicial equation

$$r^2 = 0.$$

There is only one root, $r = 0$, and hence, even though we are at a singular point, the Frobenius expansion reduces to an ordinary power series. The coefficient of $x^{r+1} = x$ tells us that $u_1 = 0$. The general recurrence relation, for $n \geq 2$, is

$$n^2 u_n + \tfrac{1}{2} n\, u_{n-2} = 0,$$

and hence

$$u_n = -\frac{u_{n-2}}{2\,n}\,.$$

Therefore, the odd coefficients $u_{2k+1} = 0$ are all zero, while the even ones are

$$u_{2k} = -\frac{u_{2k-2}}{4\,k} = \frac{u_{2k-4}}{4\,k\,(4\,k-4)} = -\frac{u_{2k-6}}{4\,k(4\,k-4)(4\,k-8)} = \cdots = \frac{(-1)^k}{4^k\,k!}\,, \quad \text{since} \quad u_0 = 1.$$

The resulting power series assumes a recognizable form:

$$\widehat{u}(x) = \sum_{k=1}^{\infty} u_{2k} x^{2\,k} = \sum_{k=1}^{\infty} \frac{1}{k!} \left(-\frac{x^2}{4}\right)^k = e^{-x^2/4},$$

which is an explicit elementary solution to the ordinary differential equation (11.94).

Since there is only one root to the indicial equation, the second solution $\widetilde{u}(x)$ will require a logarithmic term. It can be constructed by a second application of the Frobenius method using the more complicated form (11.92). Alternatively, since the first solution is known, we can use a well-known reduction trick, [**23**]. Given one solution $\widehat{u}(x)$ to a second-order linear ordinary differential equation, the general solution can be found by substituting the ansatz

$$u(x) = v(x)\,\widehat{u}(x) = v(x)\,e^{-x^2/4} \tag{11.96}$$

into the equation. In this case,

$$\begin{aligned} u'' + \left(\frac{1}{x} + \frac{x}{2}\right) u' + u &= v\left[\widehat{u}'' + \left(\frac{1}{x} + \frac{x}{2}\right)\widehat{u}' + \widehat{u}\right] + v'\left[2\,\widehat{u}' + \left(\frac{1}{x} + \frac{x}{2}\right)\widehat{u}\right] + v''\,\widehat{u} \\ &= e^{-x^2/4}\left(v'' + \frac{v'}{x}\right). \end{aligned}$$

If u is to be a solution, v' must satisfy a linear first-order ordinary differential equation:

$$v'' + \frac{v'}{x} = 0, \qquad \text{and hence} \qquad v' = \frac{c}{x}\,, \qquad v = c\,\log x + d,$$

where c, d are arbitrary constants. We conclude that the general solution to the original differential equation is

$$\widetilde{u}(x) = v(x)\,\widehat{u}(x) = (c\,\log x + d)\,e^{-x^2/4}. \tag{11.97}$$

Bessel's Equation

Perhaps the most important "non-elementary" ordinary differential equation is

$$x^2\,u'' + x\,u' + (x^2 - m^2)\,u = 0, \tag{11.98}$$

known as *Bessel's equation* of *order* m. We assume here that the order $m \geq 0$ is a nonnegative real number. (Exercise 11.3.30 investigates Bessel equations of imaginary order.) The Bessel equation arises from separation of variables in a variety of partial differential equations, including the Laplace, heat, and wave equations on a disk, a cylinder, and a spherical ball.

The Bessel equation cannot (except in a few particular instances) be solved in terms of elementary functions, and so the use of power series is essential. The leading coefficient, $p(x) = x^2$, is nonzero *except* when $x = 0$, and so all points except the origin are regular. Therefore, at any $x_0 \neq 0$, the standard power series construction can be used to produce the solutions of the Bessel equation. However, the recurrence relations for the coefficients are not particularly easy to solve in closed form. Moreover, applications tend to demand understanding the behavior of solutions at the singular point $x_0 = 0$.

Comparison with (11.88) immediately shows that $x_0 = 0$ is a regular singular point, and so we seek solutions in Frobenius form. We substitute the first, second, and fourth expressions in (11.95) into the Bessel equation and then equate the coefficients of the various powers of x to zero. The lowest power, x^r, provides the indicial equation

$$r(r-1) + r - m^2 = r^2 - m^2 = 0.$$

It has two solutions, $r = \pm m$, except when $m = 0$, for which $r = 0$ is the only index.

The higher powers of x lead to recurrence relations for the coefficients u_n in the Frobenius series. Replacing m^2 by r^2 produces

$$\begin{aligned}
x^{r+1}: &\qquad \big[\,(r+1)^2 - r^2\,\big]u_1 = (2r+1)u_1 = 0, && u_1 = 0,\\
x^{r+2}: &\qquad \big[\,(r+2)^2 - r^2\,\big]u_2 + 1 = (4r+4)u_2 + 1 = 0, && u_2 = -\frac{1}{4r+4},\\
x^{r+3}: &\qquad \big[\,(r+3)^2 - r^2\,\big]u_3 + u_1 = (6r+9)u_3 + u_1 = 0, && u_3 = -\frac{u_1}{6r+9} = 0,
\end{aligned}$$

and, in general,

$$x^{r+n}: \qquad \big[\,(r+n)^2 - r^2\,\big]u_n + u_{n-2} = n(2r+n)u_n + u_{n-2} = 0.$$

Thus, the general recurrence relation is

$$u_n = -\frac{1}{n(2r+n)}\,u_{n-2}, \qquad n = 2, 3, 4, \ldots. \tag{11.99}$$

Starting with $u_0 = 1$, $u_1 = 0$, it is easy to deduce that all $u_n = 0$ for all odd $n = 2k+1$, while for even $n = 2k$,

$$\begin{aligned}
u_{2k} &= -\frac{u_{2k-2}}{4k(k+r)} = \frac{u_{2k-4}}{16k(k-1)(r+k)(r+k-1)} = \cdots\\
&= \frac{(-1)^k}{2^{2k}\,k(k-1)\cdots 3\cdot 2\,(r+k)(r+k-1)\cdots(r+2)(r+1)}\,.
\end{aligned}$$

We have thus found the series solution

$$\widehat{u}(x) = \sum_{k=0}^{\infty} u_{2k}\,x^{r+2k} = \sum_{k=0}^{\infty} \frac{(-1)^k x^{r+2k}}{2^{2k}\,k!\,(r+k)(r+k-1)\cdots(r+2)(r+1)}\,. \tag{11.100}$$

So far, we have not paid attention to the precise values of the indices $r = \pm m$. In order to continue the recurrence, we need to ensure that the denominators in (11.99) are

never 0. Since $n > 0$, a vanishing denominator will appear whenever $2r + n = 0$, and so $r = -\frac{1}{2}n$ is either a negative integer $-1, -2, -3, \dots$ or half-integer $-\frac{1}{2}, -\frac{3}{2}, -\frac{5}{2}, \dots$. This will occur when the order $m = -r = \frac{1}{2}n$ is either an integer or a half-integer. Indeed, these are precisely the situations in which the two indices, namely $r_1 = -m$ and $r_2 = m$, differ by an integer, $r_2 - r_1 = n$, and so we are in the tricky case (*iii*) of the Frobenius method.

There is, in fact, a major difference between the integral and the half-integral cases. Recall that the odd coefficients $u_{2k+1} = 0$ in the Frobenius series automatically vanish, and so we only have to worry about the recurrence relation (11.99) for *even* values of n. When $n = 2k$, the factor $2r + n = 2(r + k) = 0$ vanishes only when $r = -k$ is a negative integer; the half-integral values do not, in fact cause problems. Therefore, if the order $m \geq 0$ is *not* an integer, then the Bessel equation of order m admits two linearly independent Frobenius solutions, given by the expansions (11.100) with exponents $r = +m$ and $r = -m$. On the other hand, if m is an integer, there is only one Frobenius solution, namely the expansion (11.100) for the positive index $r = +m$. The Frobenius recurrence with index $r = -m$ breaks down, and the second independent solution must include a logarithmic term; details appear below.

By convention, the standard *Bessel function* of order m is obtained by multiplying the Frobenius solution (11.100) with $r = m$ by

$$\frac{1}{2^m\, m!}, \qquad \text{or, more generally,} \qquad \frac{1}{2^m\, \Gamma(m+1)}, \tag{11.101}$$

where the first factorial form can be used if m is a nonnegative integer, while the more general gamma function expression must be employed for non-integral values of m. The result is

$$\begin{aligned} J_m(x) &= \sum_{k=0}^{\infty} \frac{(-1)^k x^{m+2k}}{2^{2k+m}\, k!\, (m+k)!} \\ &= \frac{1}{2^m\, m!} \left[x^m - \frac{x^{m+2}}{4\,(m+1)} + \frac{x^{m+4}}{32\,(m+1)(m+2)} - \frac{x^{m+6}}{384\,(m+1)(m+2)(m+3)} + \cdots \right]. \end{aligned} \tag{11.102}$$

When m is non-integral, the $(m+k)!$ should be replaced by $\Gamma(m+k+1)$, and $m!$ by $\Gamma(m+1)$. With this convention, the series is well defined for all real m except when $m = -1, -2, -3, \dots$ is a negative integer. Actually, if m is a negative integer, the first m terms in the series vanish, because, at negative integer values, $\Gamma(-n) = \infty$. With this convention, one can prove that

$$J_{-m}(x) = (-1)^m J_m(x), \qquad m = 1, 2, 3, \dots . \tag{11.103}$$

A simple application of the ratio test tells us that the power series converges for all (complex) values of x, and hence $J_m(x)$ is everywhere analytic. Indeed, the convergence is quite rapid when x is of moderate size, and so summing the series is a reasonably effective method for computing the Bessel function $J_m(x)$ — although in serious applications one adopts more sophisticated numerical techniques based on asymptotic expansions and integral formulas, [**85**, **86**]. In particular, we note that

$$J_0(0) = 1, \qquad J_m(0) = 0, \qquad m > 0. \tag{11.104}$$

Figure 11.4 displays graphs of the first four Bessel functions for $0 \leq x \leq 20$; the vertical axes range from $-.5$ to 1.0. Most software packages, both symbolic and numeric, include

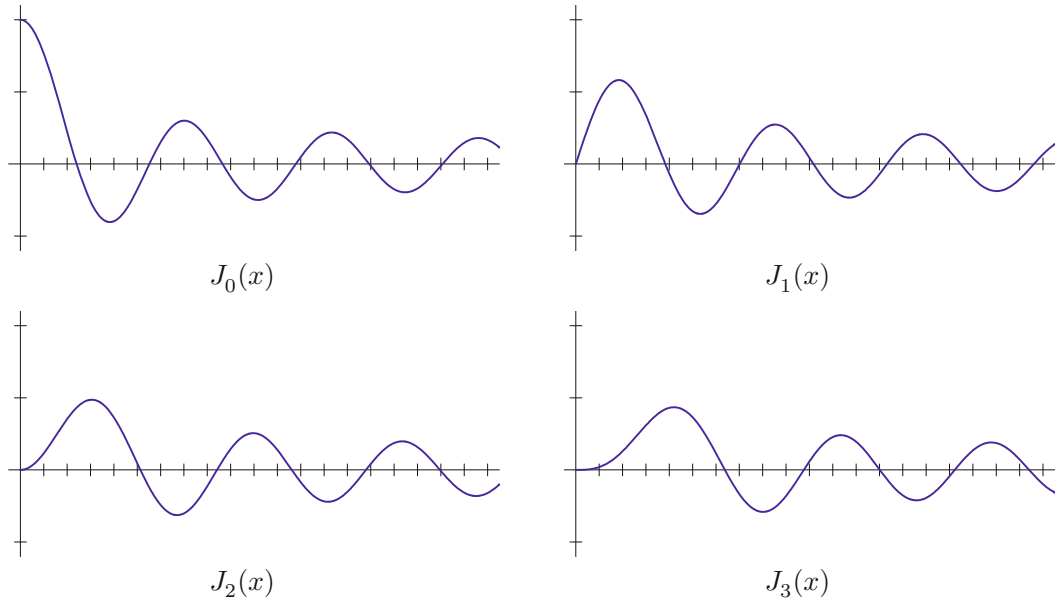


Figure 11.4. Bessel functions.

routines for accurately evaluating and graphing Bessel functions, and their properties can be regarded as well known.

Example 11.8. Consider the Bessel equation of order $m = \frac{1}{2}$. There are two indices, $r = \pm\frac{1}{2}$, and the Frobenius method yields two independent solutions: $J_{1/2}(x)$ and $J_{-1/2}(x)$. For the first, with $r = \frac{1}{2}$, the recurrence relation (11.99) takes the form

$$u_n = -\frac{u_{n-2}}{(n+1)\,n}\,.$$

Starting with $u_0 = 1$ and $u_1 = 0$, the general formula is easily found to be

$$u_n = \begin{cases} \dfrac{(-1)^k}{(n+1)\,!}\,, & n = 2\,k \text{ even}, \\ 0 & n = 2\,k+1 \text{ odd}. \end{cases}$$

Therefore, the resulting solution is

$$\widehat{u}(x) = \sqrt{x}\,\sum_{k=0}^{\infty} \frac{(-1)^k}{(2\,k+1)\,!}\,x^{2k} = \frac{1}{\sqrt{x}}\,\sum_{k=0}^{\infty} \frac{(-1)^k}{(2\,k+1)\,!}\,x^{2k+1} = \frac{\sin x}{\sqrt{x}}\,.$$

According to (11.101), the Bessel function of order $\frac{1}{2}$ is obtained by dividing this function by

$$\sqrt{2}\,\Gamma\left(\tfrac{3}{2}\right) = \sqrt{\frac{\pi}{2}}\,,$$

where we used (11.66) to evaluate the gamma function at $\frac{3}{2}$. Therefore,

$$J_{1/2}(x) = \sqrt{\frac{2}{\pi x}}\,\sin x\,. \tag{11.105}$$

Similarly, for the other index $r = -\frac{1}{2}$, the recurrence relation

$$u_n = -\frac{u_{n-2}}{n\,(n-1)}$$

leads to the formula

$$u_n = \begin{cases} \dfrac{(-1)^k}{n\,!}\,, & n = 2\,k \text{ even}, \\ 0 & n = 2\,k+1 \text{ odd}, \end{cases}$$

for its coefficients, corresponding to the solution

$$\widetilde{u}(x) = x^{-1/2} \sum_{k=0}^{\infty} \frac{(-1)^k}{(2\,k)\,!}\, x^{2k} = \frac{\cos x}{\sqrt{x}}\,.$$

Therefore, in view of (11.101) and (11.65), the Bessel function of order $-\frac{1}{2}$ is

$$J_{-1/2}(x) = \frac{\sqrt{2}}{\Gamma\left(\frac{1}{2}\right)}\,\frac{\cos x}{\sqrt{x}} = \sqrt{\frac{2}{\pi x}}\,\cos x\,. \tag{11.106}$$

As we noted above, if m is not an integer, the two independent solutions to the Bessel equation of order m are $J_m(x)$ and $J_{-\,m}(x)$. However, when m is an integer, (11.103) implies that these two solutions are constant multiples of each other, and so one must look elsewhere for a second independent solution. One method is to use a generalized Frobenius expansion involving a logarithmic term, i.e., (11.92) when $m = 0$ (see Exercise 11.3.33) or (11.93) when $m > 0$. A second approach is to employ the reduction procedure used in Example 11.7. Yet another option relies on the following limiting procedure; see [**85**, **119**] for full details.

Theorem 11.9. *If $m > 0$ is* not *an integer, then the Bessel functions $J_m(x)$ and $J_{-\,m}(x)$ provide two linearly independent solutions to the Bessel equation of order m. On the other hand, if $m = 0, 1, 2, 3, \dots$ is an integer, then a second independent solution, traditionally denoted by $Y_m(x)$ and called the* Bessel function of the second kind *of order m, can be found as a limiting case*

$$Y_m(x) = \lim_{\nu \to m} \frac{J_\nu(x)\cos\nu\,\pi - J_{-\,\nu}(x)}{\sin\nu\,\pi} \tag{11.107}$$

of a certain linear combination of Bessel functions of non-integral order ν.

With some further analysis, it can be shown that the Bessel function of the second kind of order m has the logarithmic Frobenius expansion

$$Y_m(x) = \frac{2}{\pi}\left(\gamma + \log\frac{x}{2}\right) J_m(x) + \sum_{k=0}^{\infty} b_k x^{2k-m}, \qquad m = 0, 1, 2, \dots\,, \tag{11.108}$$

with coefficients

$$b_k = \begin{cases} -\dfrac{(m-k-1)\,!}{\pi\,2^{2k-m}\,k\,!}\,, & 0 \le k \le m-1, \\ \dfrac{(-1)^{k-m-1}(h_{k-m} + h_k)}{\pi\,2^{2k-m}\,k\,!\,(k-m)\,!}\,, & k \ge m, \end{cases}$$

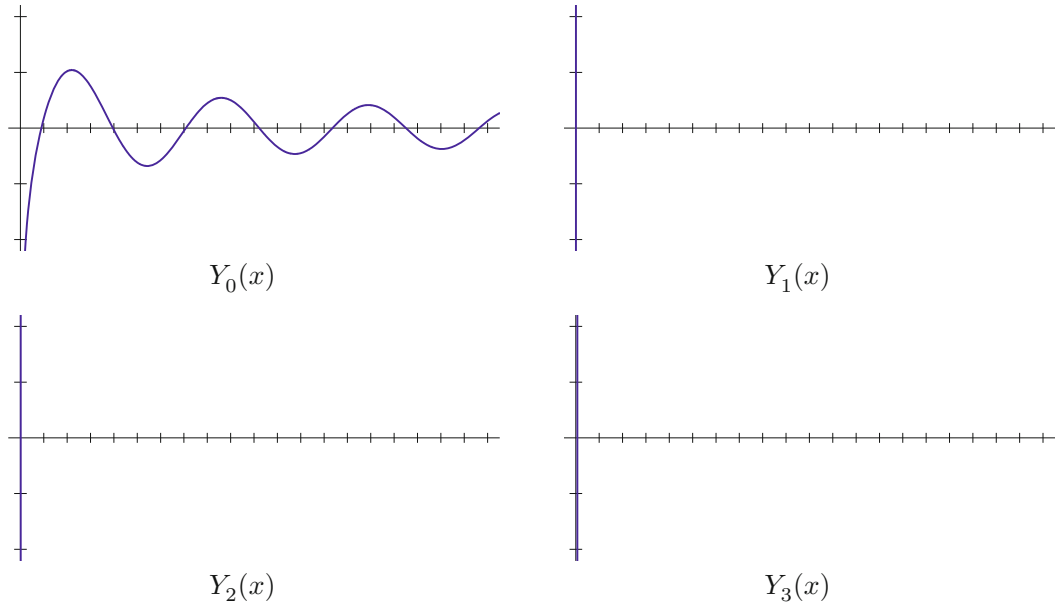


Figure 11.5. Bessel functions of the second kind.

where

$$h_0 = 0, \qquad h_k = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{k}, \qquad k > 0,$$

while

$$\gamma = \lim_{k \to \infty} \left(h_k - \log k \right) \approx .5772156649 \ldots \tag{11.109}$$

is known as the *Euler* or *Euler–Mascheroni constant*. All Bessel functions of the second kind have a singularity at the origin $x = 0$; indeed, by inspection of (11.108), we find that the leading asymptotics as $x \to 0$ are

$$Y_0(x) \sim \frac{2}{\pi} \log x, \qquad Y_m(x) \sim -\frac{2^m (m-1)!}{\pi\, x^m}, \qquad m > 0. \tag{11.110}$$

Figure 11.5 contains graphs of the first four Bessel function of the second kind on the interval $0 < x \leq 20$; the vertical axis ranges from -1 to 1.

Finally, we show how Bessel functions of different orders are interconnected by two important recurrence relations.

Proposition 11.10. *The Bessel functions are related by the following formulae*:

$$\frac{dJ_m}{dx} + \frac{m}{x} J_m(x) = J_{m-1}(x), \qquad -\frac{dJ_m}{dx} + \frac{m}{x} J_m(x) = J_{m+1}(x). \tag{11.111}$$

Proof: Differentiating the power series

$$x^m J_m(x) = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2m+2k}}{2^{2k+m}\, k!\, (m+k)!}$$

produces

$$\begin{aligned}\frac{d}{dx}\,[\,x^m J_m(x)\,] &= \sum_{k=0}^{\infty} \frac{(-1)^k\, 2\,(m+k) x^{2m+2k-1}}{2^{2k+m}\, k!\,(m+k)!} \\ &= x^m \sum_{k=0}^{\infty} \frac{(-1)^k x^{m-1+2k}}{2^{2k+m-1}\, k!\,(m-1+k)!} = x^m\, J_{m-1}(x).\end{aligned} \tag{11.112}$$

Expansion of the left-hand side of this formula leads to

$$x^m\, \frac{dJ_m}{dx} + m\, x^{m-1} J_m(x) = \frac{d}{dx}\,[\,x^m J_m(x)\,] = x^m J_{m-1}(x),$$

which establishes the first recurrence formula (11.111). The second is proved by a similar manipulation involving differentiation of $x^{-m}\, J_m(x)$. *Q.E.D.*

For example, using the second recurrence formula (11.111) along with (11.105), we can write the Bessel function of order $\frac{3}{2}$ in elementary terms:

$$\begin{aligned}J_{3/2}(x) &= -\,\frac{dJ_{1/2}(x)}{dx} + \frac{1}{2\,x}\, J_{1/2}(x) \\ &= -\sqrt{\frac{2}{\pi}}\left(\frac{\cos x}{x^{1/2}} - \frac{\sin x}{2\,x^{3/2}}\right) + \sqrt{\frac{2}{\pi}}\,\frac{\sin x}{2\,x^{3/2}} = \sqrt{\frac{2}{\pi}}\,\frac{\sin x - x\cos x}{x^{3/2}}\,.\end{aligned} \tag{11.113}$$

Iterating, one concludes that Bessel functions of half-integral order, $m = \pm\frac{1}{2}\,, \pm\frac{3}{2}\,, \pm\frac{5}{2}\,, \ldots$, are all elementary functions, in that they can be written in terms of trigonometric functions and powers of $\sqrt{x}\,$. We will make use of these functions in our treatment of the three-dimensional heat and wave equations in spherical geometry. On the other hand, all of the other Bessel functions are non-elementary special functions.

With this, we conclude our brief introduction to the method of Frobenius and the basics of Bessel functions. The reader interested in delving further into either the general method or the host of additional properties of Bessel functions is encouraged to consult a more specialized text, e.g., [**59**, **85**, **119**].

Exercises

11.3.22. Consider the ordinary differential equation $2\,x\,u'' + u' + x\,u = 0$. (a) Prove that $x = 0$ is a regular singular point. (*b*) Find two independent series solutions in powers of x.

♡ 11.3.23. Consider the differential equation $\dfrac{u''}{2-x} = \dfrac{u}{x^2}\,$. (a) Classify all $x_0 \in \mathbb{R}$ as either a (*i*) regular point; (*ii*) regular singular point; and/or (*iii*) irregular singular point. Explain your answers. (*b*) Find a series solution to the equation based at the point $x_0 = 0$, or explain why none exists. What is the radius of convergence of your series?

11.3.24. Consider the differential equation $u'' + \left(1 - \dfrac{1}{x}\right) u' + u = 0$.

(a) Classify all $x_0 \in \mathbb{R}$ as either (*i*) a regular point; (*ii*) a regular singular point; (*iii*) an irregular singular point; (*iv*) none of the above. Explain your answers.

(*b*) Write out the first five nonzero terms in a series solution.

11.3.25. Consider the differential equation $4x\,u'' + 2u' + u = 0$. (*a*) Classify the values of x for which the equation has regular points, regular singular points, and irregular singular points. (*b*) Find two independent series solutions, in powers of x. For what values of x do your series converge? (*c*) By inspection of your series, write the general solution to the equation in terms of elementary functions.

♡ 11.3.26. The *Chebyshev differential equation* is $(1 - x^2)u'' - x\,u' + m^2 u = 0$. (*a*) Find all (*i*) regular points; (*ii*) regular singular points; (*iii*) irregular singular points. (*b*) Show that if m is an integer, the equation has a polynomial solution of degree m, known as a *Chebyshev polynomial*. Write down the Chebyshev polynomials of degrees 1, 2, and 3. (*c*) For $m = 1$, find two linearly independent series solutions based at the point $x_0 = 1$.

11.3.27. Write the following Bessel functions in terms of elementary functions:

$$(a)\ J_{5/2}(x), \quad (b)\ J_{7/2}(x), \quad (c)\ J_{-3/2}(x).$$

♢ 11.3.28. Prove the identity (11.103).

11.3.29. Suppose that $u(x)$ solves Bessel's equation. (*a*) Find a second order ordinary differential equation satisfied by the function $w(x) = \sqrt{x}\,u(x)$. (*b*) Use this result to rederive the formulas for $J_{1/2}(x)$ and $J_{-1/2}(x)$.

♢ 11.3.30. Let $m \geq 0$ be real, and consider the *modified Bessel equation* of order m:

$$x^2\,u'' + x\,u' - (x^2 + m^2)\,u = 0. \tag{11.114}$$

(*a*) Explain why $x_0 = 0$ is a regular singular point.
(*b*) Use the method of Frobenius to construct a series solution based at $x_0 = 0$. Can you relate your solutions to the Bessel function $J_m(x)$?

♢ 11.3.31. (*a*) Let a, b, c be constants with $b, c \neq 0$. Show that the function $u(x) = x^a J_0(b\,x^c)$ solves the ordinary differential equation

$$x^2\,\frac{d^2u}{dx^2} + (1 - 2a)\,x\,\frac{du}{dx} + (b^2c^2x^{2c} + a^2)\,u = 0.$$

What is the general solution to this equation?
(*b*) Find the general solution to the ordinary differential equation

$$x^2\,\frac{d^2u}{dx^2} + \alpha\,x\,\frac{du}{dx} + (\beta\,x^{2c} + \gamma)\,u = 0,$$

for constants α, β, γ, c with $\beta, c \neq 0$.

♡ 11.3.32. Let $k > 0$ be a constant. The ordinary differential equation $\dfrac{d^2u}{dt^2} + e^{-2t}\,u = 0$ describes the vibrations of a weakening spring whose stiffness $k(t) = e^{-2t}$ is exponentially decaying in time. (*a*) Show that this equation can be solved in terms of Bessel functions of order 0. *Hint*: Perform a change of variables. (*b*) Does the solution tend to 0 as $t \to \infty$?

♡ 11.3.33. We know that $\widehat{u}(x) = J_0(x)$ is a solution to the Bessel equation of order 0, namely

$$x\,u'' + u' + x\,u = 0. \tag{11.115}$$

In accordance with the general Frobenius method, construct a second solution of the form

$$\widetilde{u}(x) = J_0(x)\log x + \sum_{n=1}^{\infty} v_n x^n.$$

11.3.34. Is it possible to have all solutions to an ordinary differential equation bounded at a regular singular point? If not, explain why not. If true, give an example where this happens.

11.4 The Heat Equation in a Disk, Continued

Now that we have acquired some familiarity with the solutions to Bessel's ordinary differential equation, we are ready to analyze the separable solutions to the heat equation in a polar geometry. At the end of Section 11.2, we were left with the task of solving the Bessel equation (11.58) of integer order m. As we now know, there are two independent solutions, namely the Bessel function of the first kind J_m, (11.102), and the more complicated Bessel function of the second kind Y_m, (11.107), and hence the general solution has the form

$$p(z) = c_1 J_m(z) + c_2 Y_m(z),$$

for constants c_1, c_2. Reverting to our original radial coordinate $r = z/\sqrt{\lambda}$, we conclude that every solution to the radial equation (11.56) has the form

$$p(r) = c_1 J_m\big(\sqrt{\lambda}\, r\big) + c_2 Y_m\big(\sqrt{\lambda}\, r\big).$$

Now, the singular point $r = 0$ represents the center of the disk, and the solutions must remain bounded there. While this is true for $J_m(z)$, the second Bessel function $Y_m(z)$ has, according to (11.110), a singularity at $z = 0$ and so is unsuitable for the present purposes. (On the other hand, it plays a role in other situations, e.g., the heat equation on an annular ring.) Thus, every separable solution that is bounded at $r = 0$ comes from the rescaled Bessel function of the first kind of order m:

$$p(r) = J_m\big(\sqrt{\lambda}\, r\big). \tag{11.116}$$

The Dirichlet boundary condition at the disk's rim $r = 1$ requires

$$p(1) = J_m\big(\sqrt{\lambda}\,\big) = 0.$$

Therefore, in order that λ be a bona fide eigenvalue, $\sqrt{\lambda}$ must be a *root* of the m^{th} order Bessel function J_m.

Remark: We already know, thanks to the positive definiteness of the Dirichlet boundary value problem, that the Helmholtz eigenvalues must all be positive, $\lambda > 0$, and so there will be no difficulty in taking its square root.

The graphs of $J_m(z)$ strongly indicate, and, indeed, it can be rigorously proved, [**85**, **119**], that as z increases above 0, each Bessel function oscillates, with slowly decreasing amplitude, between positive and negative values. In fact, asymptotically,

$$J_m(z) \;\sim\; \sqrt{\frac{2}{\pi z}}\, \cos\big[\, z - \big(\tfrac{1}{2}m + \tfrac{1}{4}\big)\pi \,\big] \qquad \text{as} \qquad z \longrightarrow \infty, \tag{11.117}$$

and so the oscillations become essentially the same as a (phase-shifted) cosine whose amplitude decreases like $z^{-1/2}$. As a consequence, there exists an infinite sequence of *Bessel roots*, which we number in increasing order:

$$\begin{aligned} &J_m(\zeta_{m,n}) = 0, \qquad \text{where} \\ &\qquad 0 < \zeta_{m,1} < \zeta_{m,2} < \zeta_{m,3} < \cdots \quad \text{with} \quad \zeta_{m,n} \longrightarrow \infty \quad \text{as} \quad n \longrightarrow \infty. \end{aligned} \tag{11.118}$$

It is worth emphasizing that the Bessel functions are *not* periodic, and so their roots are not evenly spaced. However, as a consequence of (11.117), the large Bessel roots are asymptotically close to the evenly spaced roots of the shifted cosine:

$$\zeta_{m,n} \;\sim\; \big(n + \tfrac{1}{2}m - \tfrac{1}{4}\big)\pi \qquad \text{as} \qquad n \longrightarrow \infty. \tag{11.119}$$

Owing to their physical importance in a wide range of problems, the Bessel roots have been extensively tabulated. The accompanying table displays all Bessel roots that are < 12 in magnitude. The columns of the table are indexed by m, the order of the Bessel function, and the rows by n, the root number.

Table of Bessel Roots $\zeta_{m,n}$

n \ m	0	1	2	3	4	5	6	7	...
1	2.4048	3.8317	5.1356	6.3802	7.5883	8.7715	9.9361	11.0864	...
2	5.5201	7.0156	8.4172	9.7610	11.0647	⋮	⋮	⋮	
3	8.6537	10.1735	11.6198	⋮	⋮				
4	11.7915	⋮	⋮						
⋮	⋮								

Remark: According to (11.102),

$$J_m(0) = 0 \qquad \text{for} \qquad m > 0, \qquad \text{while} \qquad J_0(0) = 1.$$

However, we do not count 0 as a bona fide Bessel root, since it does not lead to a valid eigenfunction for the Helmholtz boundary value problem.

Summarizing our progress so far, the eigenvalues

$$\lambda_{m,n} = \zeta_{m,n}^2, \qquad n = 1, 2, 3, \ldots, \qquad m = 0, 1, 2, \ldots, \tag{11.120}$$

of the Bessel boundary value problem (11.56–57) are the squares of the roots of the Bessel function of order m. The corresponding eigenfunctions are

$$w_{m,n}(r) = J_m(\zeta_{m,n}\, r), \qquad n = 1, 2, 3, \ldots, \qquad m = 0, 1, 2, \ldots, \tag{11.121}$$

defined for $0 \le r \le 1$. Combining (11.121) with the formula (11.55) for the angular components, we conclude that the separable solutions (11.53) to the polar Helmholtz boundary value problem (11.51) are

$$\begin{aligned} v_{0,n}(r) &= J_0(\zeta_{0,n}\, r), \\ v_{m,n}(r,\theta) &= J_m(\zeta_{m,n}\, r)\, \cos m\theta, \\ \widehat{v}_{m,n}(r,\theta) &= J_m(\zeta_{m,n}\, r)\, \sin m\theta, \end{aligned} \qquad \text{where} \qquad m, n = 1, 2, 3, \ldots . \tag{11.122}$$

These solutions define the *normal modes* for the unit disk; Figure 11.6 plots the first few of them. The eigenvalues $\lambda_{0,n}$ are simple, and contribute radially symmetric eigenfunctions, whereas the eigenvalues $\lambda_{m,n}$ for $m > 0$ are double, and produce two linearly independent separable eigenfunctions, with trigonometric dependence on the angular variable.

Recalling the original ansatz (11.50), we have at last produced the basic separable eigensolutions

$$\begin{aligned} u_{0,n}(t,r) &= e^{-\zeta_{0,n}^2 t}\, v_{0,n}(r) = e^{-\zeta_{0,n}^2 t}\, J_0(\zeta_{0,n}\, r), \\ u_{m,n}(t,r,\theta) &= e^{-\zeta_{m,n}^2 t}\, v_{m,n}(r,\theta) = e^{-\zeta_{m,n}^2 t}\, J_m(\zeta_{m,n} r)\, \cos m\theta, \\ \widehat{u}_{m,n}(t,r,\theta) &= e^{-\zeta_{m,n}^2 t}\, \widehat{v}_{m,n}(r,\theta) = e^{-\zeta_{m,n}^2 t}\, J_m(\zeta_{m,n} r)\, \sin m\theta, \end{aligned} \qquad m, n = 1, 2, 3, \ldots , \tag{11.123}$$

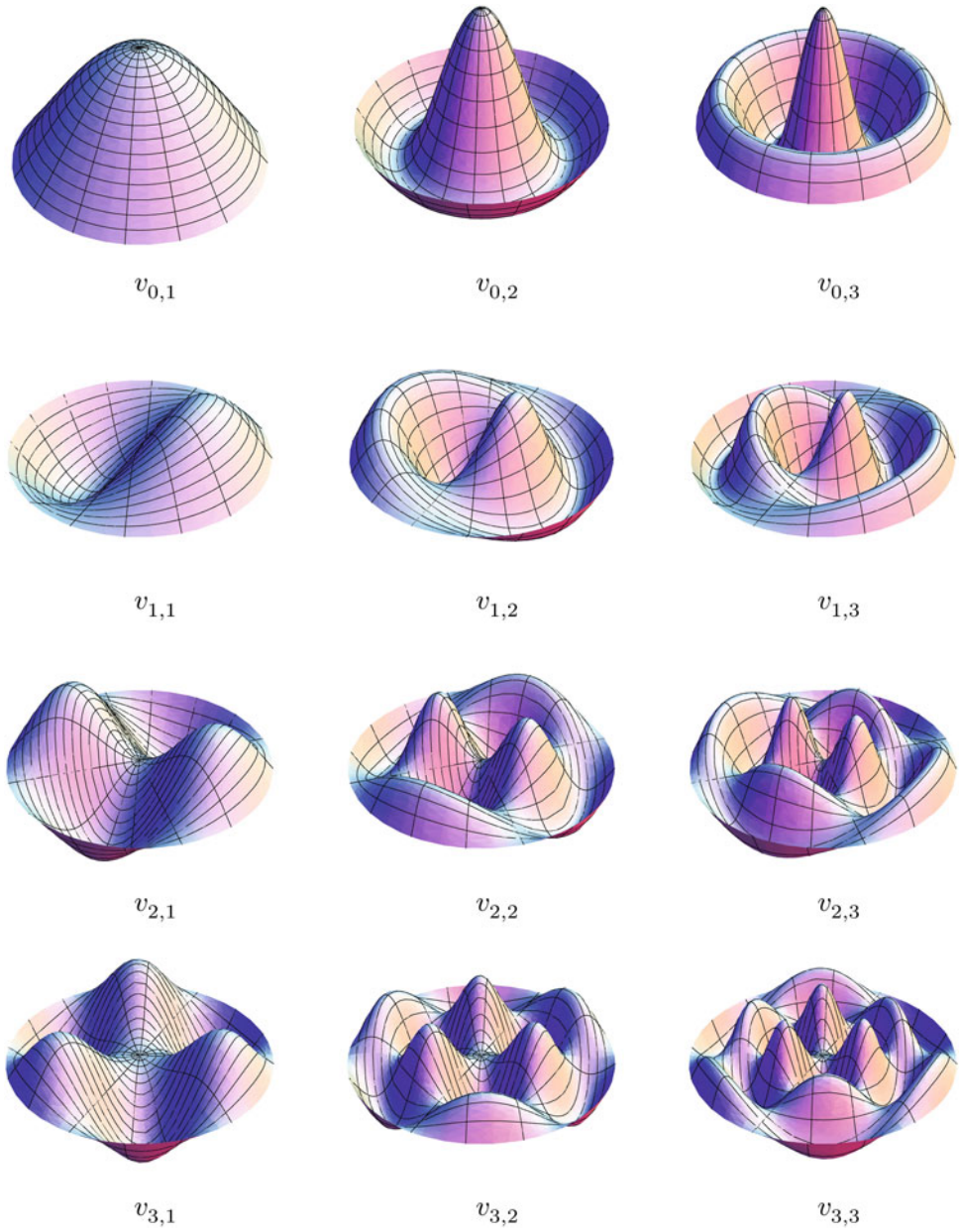


Figure 11.6. Normal modes for a disk.

to the homogeneous Dirichlet boundary value problem for the heat equation on the unit disk. The general solution is obtained by linear superposition, in the form of an infinite series

$$u(t,r,\theta) = \frac{1}{2}\sum_{n=1}^{\infty} a_{0,n}\, u_{0,n}(t,r) + \sum_{m,n=1}^{\infty} \big[\, a_{m,n}\, u_{m,n}(t,r,\theta) + b_{m,n}\, \widehat{u}_{m,n}(t,r,\theta)\,\big], \quad (11.124)$$

where the initial factor of $\frac{1}{2}$ is included, as with ordinary Fourier series, for later conve-

nience. As usual, the coefficients $a_{m,n}, b_{m,n}$ are determined by the initial condition

$$u(0,r,\theta) = \frac{1}{2}\sum_{n=1}^{\infty} a_{0,n}\, v_{0,n}(r) + \sum_{m,n=1}^{\infty} \big[\, a_{m,n}\, v_{m,n}(r,\theta) + b_{m,n}\, \widehat{v}_{m,n}(r,\theta)\,\big] = f(r,\theta). \tag{11.125}$$

This requires that we expand the initial data into a *Fourier–Bessel series* in the eigenfunctions. As before, it is possible to prove, [**34**], that the separable eigenfunctions are *complete* — there are no other eigenfunctions — and hence every (reasonable) function defined on the unit disk can be written as a convergent series in the Bessel eigenfunctions.

Theorem 9.33 gurantees that the eigenfunctions are orthogonal† with respect to the standard L^2 inner product

$$\langle\, u\,, v\,\rangle = \iint_D u(x,y)\, v(x,y)\, dx\, dy = \int_0^1 \int_{-\pi}^{\pi} u(r,\theta)\, v(r,\theta)\; r\, d\theta\, dr$$

on the unit disk. (Note the extra factor of r coming from the polar coordinate form of the area element $dx\, dy = r\, dr\, d\theta$.) The L^2 norms of the Fourier–Bessel eigenfunctions are given by the interesting formulae

$$\| v_{0,n} \| = \sqrt{\pi}\; \big|\, J_1(\zeta_{0,n}) \,\big|\,, \qquad \| v_{m,n} \| = \| \widehat{v}_{m,n} \| = \sqrt{\frac{\pi}{2}}\; \big|\, J_{m+1}(\zeta_{m,n}) \,\big|\,, \tag{11.126}$$

which involve the value of the Bessel function of the next-higher order at the appropriate Bessel root. A proof of (11.126) can be found in Exercise 11.4.22, while numerical values are provided in the accompanying table.

Norms of the Fourier–Bessel Eigenfunctions $\| v_{m,n} \| = \| \widehat{v}_{m,n} \|$

n \ m	0	1	2	3	4	5	6	7
1	.9202	.5048	.4257	.3738	.3363	.3076	.2847	.2658
2	.6031	.3761	.3401	.3126	.2906	.2725	.2572	.2441
3	.4811	.3130	.2913	.2736	.2586	.2458	.2347	.2249
4	.4120	.2737	.2589	.2462	.2352	.2255	.2169	.2092
5	.3661	.2462	.2353	.2257	.2171	.2095	.2025	.1962

Orthogonality of the eigenfunctions implies that the coefficients in the Fourier–Bessel

† For the two independent eigenfunctions corresponding to one of the double eigenvalues, orthogonality must be verified by hand, but, in this case, it follows easily from the orthogonality of their trigonometric components.

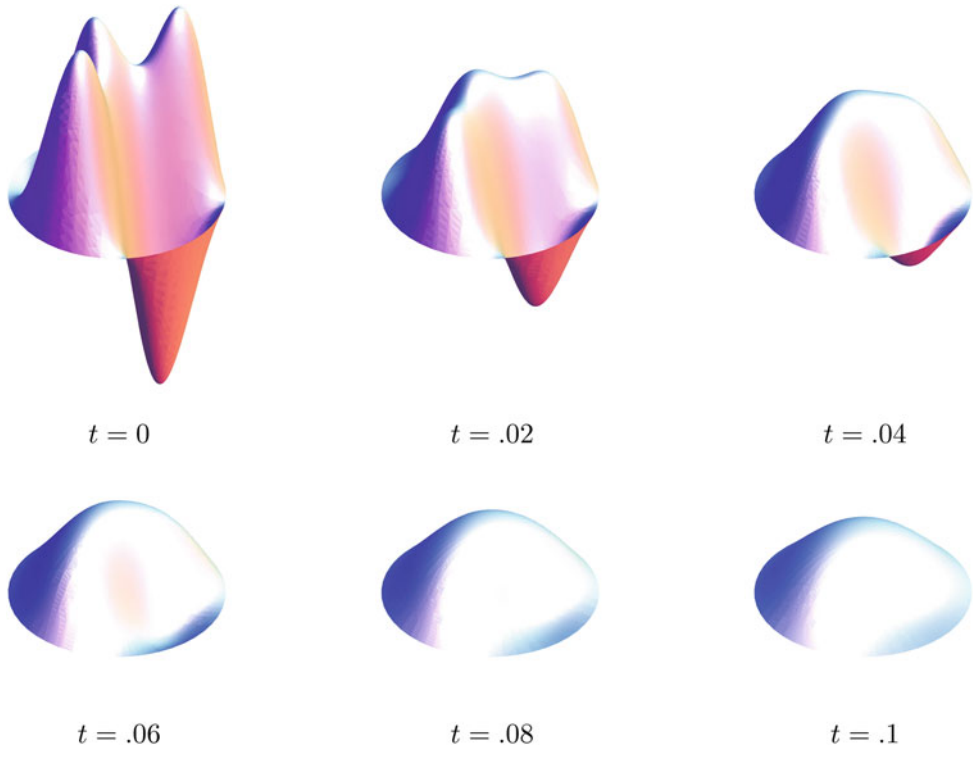


Figure 11.7. Heat diffusion in a disk.

series (11.125) are given by the inner product formulae

$$a_{0,n} = 2\,\frac{\langle\, f\,,v_{0,n}\,\rangle}{\|\,v_{0,n}\,\|^2} = \frac{2}{\pi\, J_1(\zeta_{0,n})^2}\int_0^1\int_{-\pi}^{\pi} f(r,\theta)\, J_0(\zeta_{0,n}\,r)\; r\,d\theta\,dr,$$
$$a_{m,n} = \frac{\langle\, f\,,v_{m,n}\,\rangle}{\|\,v_{m,n}\,\|^2} = \frac{2}{\pi\, J_{m+1}(\zeta_{m,n})^2}\int_0^1\int_{-\pi}^{\pi} f(r,\theta)\, J_m(\zeta_{m,n}\,r)\; r\cos m\,\theta\,d\theta\,dr, \tag{11.127}$$
$$b_{m,n} = \frac{\langle\, f\,,\widehat{v}_{m,n}\,\rangle}{\|\,\widehat{v}_{m,n}\,\|^2} = \frac{2}{\pi\, J_{m+1}(\zeta_{m,n})^2}\int_0^1\int_{-\pi}^{\pi} f(r,\theta)\, J_m(\zeta_{m,n}\,r)\; r\sin m\,\theta\,d\theta\,dr.$$

In accordance with the general theory, each individual separable solution (11.123) to the heat equation decays exponentially fast, at a rate $\lambda_{m,n} = \zeta_{m,n}^2$ prescribed by the square of the corresponding Bessel root. In particular, the dominant mode, meaning the one that persists the longest, is

$$u_{0,1}(t,r,\theta) = e^{-\,\zeta_{0,1}^2\,t}\; J_0(\zeta_{0,1}\,r). \tag{11.128}$$

Its decay rate is prescribed by the smallest positive eigenvalue:

$$\zeta_{0,1}^2 \;\approx\; 5.783, \tag{11.129}$$

which is the square of the smallest root of the Bessel function $J_0(z)$. Since $J_0(z) > 0$ for $0 \le z < \zeta_{0,1}$, the dominant eigenfunction $v_{0,1}(r,\theta) = J_0(\zeta_{0,1}\,r) > 0$ is radially symmetric and strictly positive within the entire disk. Consequently, for most initial conditions

(specifically those for which $a_{0,1} \neq 0$), the disk's temperature distribution eventually becomes entirely of one sign and radially symmetric, while decaying exponentially fast to zero at the rate given by (11.129). See Figure 11.7 for a plot of a typical solution. Note how, in accordance with the theory, the solution soon acquires a radial symmetry as it decays to thermal equilibrium.

Exercises

11.4.1. At the initial time $t_0 = 0$, a concentrated unit heat source is instantaneously applied at position $x = \frac{1}{2}$, $y = 0$, to a circular metal disk of unit radius and unit thermal diffusivity whose outside edge is held at 0°. Write down an eigenfunction series for the resulting temperature distribution at time $t > 0$. *Hint*: Be careful working with the delta function in polar coordinates; see Exercise 6.3.6.

11.4.2. Solve Exercise 11.4.1 when the concentrated unit heat source is instantaneously applied at the center of the disk.

♡ 11.4.3. (*a*) Write down the Fourier–Bessel series for the solution to the heat equation on a unit disk with $\gamma = 1$, whose circular edge is held at 0° and subject to the initial conditions $u(0, x, y) \equiv 1$ for $x^2 + y^2 \leq 1$. *Hint*: Use (11.112) to evaluate the integrals for the coefficients. (*b*) Approximate the time $t_\star \geq 0$ after which the temperature of the disk is everywhere $\leq .5^\circ$.

♣ 11.4.4. (*a*) Write down the first three nonzero terms in the Fourier–Bessel series for the solution to the heat equation on a unit disk with $\gamma = 1$ whose circular edge is held at 0° subject to the initial conditions $u(0, r, \theta) = 1 - r$ for $r \leq 1$. Use numerical integration to evaluate the coefficients. (*b*) Use your approximation to determine at which times $t \geq 0$ the temperature of the disk is everywhere $\leq .5^\circ$.

11.4.5. Prove that every separable eigenfunction of the Dirichlet boundary value problem for the Helmholtz equation in the unit disk can be written in the form

$$c\, J_m(\zeta_{m,n}\, r)\, \cos(m\theta - \alpha) \quad \text{for fixed } c \neq 0 \text{ and } -\pi < \alpha \leq \pi.$$

11.4.6. Suppose the initial data $f(r, \theta)$ in (11.49) satisfies $\displaystyle\int_0^1 \int_{-\pi}^{\pi} f(r, \theta)\, J_0(\zeta_{0,1}\, r)\, r\, d\theta\, dr = 0$. (*a*) What is the decay rate to equilibrium of the resulting heat equation solution $u(t, r, \theta)$? (*b*) Prove that, generically, the asymptotic temperature distribution has half the disk above the equilibrium temperature and the other half below. Can you predict the diameter that separates the two halves? (*c*) If you know that $a_{0,1} = 0$, and also that the long-time temperature distribution is radially symmetric, what is the (generic) decay rate? What is the asymptotic temperature distribution?

♢ 11.4.7. Show how to use a scaling symmetry to solve the heat equation in a disk of radius R knowing the solution in a disk of radius 1.

11.4.8. Use rescaling, as in Exercise 11.4.7, to produce the solution to the Dirichlet initial-boundary value problem for a disk of radius 2 with diffusion coefficient $\gamma = 5$.

11.4.9. If it takes a disk of unit radius 3 minutes to reach (approximate) thermal equilibrium, how long will it take a disk of radius 2 made out of the same material and subject to the same homogeneous boundary conditions to reach equilibrium?

11.4.10. Assuming Dirichlet boundary conditions, does a square or a circular disk of the same area reach thermal equilibrium faster? Use your intuition first, and then check using the explicit formulas.

11.4.11. Answer Exercise 11.4.10 when the square and circle have the same perimeter.

11.4.12. Which reaches thermal equilibrium faster: a disk whose edge is held at 0° or a disk of the same radius that is fully insulated?

11.4.13. A circular metal disk is removed from an oven and then fully insulated.
True or false: (*a*) The eventual equilibrium temperature is constant.
(*b*) For large $t \gg 0$, the temperature $u(t, x, y)$ becomes more and more radially symmetric. If false, what can you say about the temperature profile at large times?

♡ 11.4.14. (*a*) Write down an eigenfunction series formula for the temperature dynamics of a disk of radius 1 that has an insulated boundary. (*b*) What is the eventual equilibrium temperature? (*c*) Is the rate of decay to thermal equilibrium (*i*) faster, (*ii*) slower, or (*iii*) the same as a disk with Dirichlet boundary conditions?

♡ 11.4.15. Write out a series solution for the temperature in a half-disk of radius 1, subject to (*a*) homogeneous Dirichlet boundary conditions on its entire boundary; (*b*) homogeneous Dirichlet conditions on the circular part of its boundary and homogeneous Neumann conditions on the straight part. (*c*) Which of the two boundary conditions results in a faster return to equilibrium temperature? How much faster?

11.4.16. A large sheet of metal is heated to 100°. A circular disk and a semi-circular half-disk of the same radius are cut out of it. Their edges are then held at 0°, while being fully insulated from above and below.
(*a*) *True or false*: The half-disk goes to thermal equilibrium twice as fast as the disk.
(*b*) If you need to wait 20 minutes for the circular disk to cool down enough to be picked up in your bare hands, how long do you need to wait to pick up the semi-circular disk?

♣ 11.4.17. Two identical plates have the shape of an annular ring $\{\, 1 < r < 2 \,\}$ with inner radius 1 and outer radius 2. The first has an insulated inner boundary and outer boundary held at 0°, while the second has an insulated outer boundary and inner boundary held at 0°. If both start out at the same temperature, which reaches thermal equilibrium faster? Quantify the rates of decay.

♡ 11.4.18. Let $m \geq 0$ be a nonnegative integer. In this exercise, we investigate the completeness of the eigenfunctions of the Bessel boundary value problem (11.56–57). To this end, define the Sturm–Liouville linear differential operator

$$S[u] = -\frac{1}{x}\,\frac{d}{dx}\left(x\,\frac{du}{dx}\right) + \frac{m^2}{x^2}\,u,$$

subject to the boundary conditions $|\,u'(0)\,| < \infty$, $u(1) = 0$, and either $|\,u(0)\,| < \infty$ when $m = 0$, or $u(0) = 0$ when $m > 0$.

(*a*) Show that S is self-adjoint relative to the inner product $\langle\, f\,, g\,\rangle = \displaystyle\int_0^1 f(x)\,g(x)\,x\,dx$.
(*b*) Prove that the eigenfunctions of S are the rescaled Bessel functions $J_m(\zeta_{m,n}\,x)$ for $n = 1, 2, 3, \ldots$. What are the orthogonality relations?
(*c*) Find the Green's function $G(x;\xi)$ and modified Green's function $\widehat{G}(x;\xi)$, cf. (9.59), associated with the boundary value problem $S[u] = 0$.
(*d*) Use the criterion of Theorem 9.47 to prove that the eigenfunctions are complete.

11.4.19. Determine the Bessel roots $\zeta_{1/2,n}$. Do they satisfy the asymptotic formula (11.119)?

♣ 11.4.20. Use a numerical root finder to compute the first 10 Bessel roots $\zeta_{3/2,n}$, $n = 1, \ldots, 10$. Compare your values with the asymptotic formula (11.119).

♢ 11.4.21. Prove that $J_{m-1}(\zeta_{m,n}) = -\,J_{m+1}(\zeta_{m,n})$.

♢ 11.4.22. In this exercise, we prove formula (11.126).
(*a*) First, use the recurrence formulae (11.111) to prove

$$\frac{d}{dx}\left[\,x^2\Big(\,J_m(x)^2 - J_{m-1}(x)\,J_{m+1}(x)\,\Big)\,\right] = 2\,x\,J_m(x)^2.$$

(*b*) Integrate both sides of the previous formula from 0 to the Bessel zero $\zeta_{m,n}$ and then

use Exercise 11.4.21 to show that

$$\int_0^{\zeta_{m,n}} x\, J_m(x)^2\, dx = -\frac{\zeta_{m,n}^2}{2}\, J_{m-1}(\zeta_{m,n})\, J_{m+1}(\zeta_{m,n}) = \frac{\zeta_{m,n}^2}{2}\, J_{m+1}(\zeta_{m,n})^2.$$

(*c*) Next, use a change of variables to establish the identity

$$\int_0^1 z\, J_m(\zeta_{m,n}\, z)^2\, dz = \tfrac{1}{2}\, J_{m+1}(\zeta_{m,n})^2.$$

(*d*) Finally, use the formulae for $v_{m,n}$ and $\widehat{v}_{m,n}$ to complete the proof of (11.126).

♢ 11.4.23. Prove directly that the eigenfunctions $v_{m,n}(r,\theta)$ and $\widehat{v}_{m,n}(r,\theta)$ in (11.122) are orthogonal with respect to the L^2 inner product on the unit disk.

11.4.24. Establish the following alternative formulae for the eigenfunction norms:

$$\| v_{0,n} \| = \sqrt{\pi}\, \big|\, J_0'(\zeta_{0,n}) \,\big|, \qquad \| v_{m,n} \| = \| \widehat{v}_{m,n} \| = \sqrt{\frac{\pi}{2}}\, \big|\, J_m'(\zeta_{m,n}) \,\big|.$$

11.5 The Fundamental Solution to the Planar Heat Equation

As we learned in Section 4.1, the fundamental solution to the heat equation measures the temperature distribution resulting from a concentrated initial heat source, e.g., a hot soldering iron applied instantaneously at a single point on a metal plate. The physical problem is modeled mathematically by taking a delta function as the initial data along with the relevant homogeneous boundary conditions. Once the fundamental solution is known, one is able to use linear superposition to recover the solution generated by any other initial data.

As in our one-dimensional analysis, we shall concentrate on the most tractable case, in which the domain is the entire plane: $\Omega = \mathbb{R}^2$. Thus, our first goal is to solve the initial value problem

$$u_t = \gamma\, \Delta u, \qquad u(0,x,y) = \delta(x-\xi)\, \delta(y-\eta), \tag{11.130}$$

for $t > 0$ and $(x,y) \in \mathbb{R}^2$. The solution $u = F(t, \mathbf{x}; \boldsymbol{\xi}) = F(t,x,y;\xi,\eta)$ to this initial value problem is known as the *fundamental solution* for the heat equation on $\mathbb{R}^2$.

The quickest route to the desired formula relies on the following means of combining solutions of the one-dimensional heat equation to produce solutions of the two-dimensional version.

Lemma 11.11. *Let $v(t,x)$ and $w(t,x)$ be any two solutions to the one-dimensional heat equation $u_t = \gamma\, u_{xx}$. Then their product*

$$u(t,x,y) = v(t,x)\, w(t,y) \tag{11.131}$$

is a solution to the two-dimensional heat equation $u_t = \gamma\, (u_{xx} + u_{yy})$.

Proof: Our assumptions imply that $v_t = \gamma\, v_{xx}$, while $w_t = \gamma\, w_{yy}$ when we write $w(t,y)$ as a function of t and y. Therefore, differentiating (11.131), we find

$$\frac{\partial u}{\partial t} = \frac{\partial v}{\partial t}\, w + v\, \frac{\partial w}{\partial t} = \gamma\, \frac{\partial^2 v}{\partial x^2}\, w + \gamma\, v\, \frac{\partial^2 w}{\partial y^2} = \gamma \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right),$$

and hence $u(t,x,y)$ solves the two-dimensional heat equation. *Q.E.D.*

For example, if

$$v(t,x) = e^{-\gamma\alpha^2 t}\,\sin\alpha x, \qquad w(t,y) = e^{-\gamma\beta^2 t}\,\sin\beta y,$$

are separable solutions of the one-dimensional heat equation, then

$$u(t,x,y) = e^{-\gamma(\alpha^2+\beta^2)t}\,\sin\alpha x\,\sin\beta y$$

are the separable solutions we used to solve the heat equation on a rectangle. A more interesting case is to choose

$$v(t,x) = \frac{1}{2\sqrt{\pi\gamma t}}\,e^{-(x-\xi)^2/(4\gamma t)}, \qquad w(t,y) = \frac{1}{2\sqrt{\pi\gamma t}}\,e^{-(y-\eta)^2/(4\gamma t)}, \tag{11.132}$$

to be the fundamental solutions (8.14) to the one-dimensional heat equation at respective locations $x = \xi$ and $y = \eta$. Multiplying these two solutions together produces the fundamental solution for the two-dimensional problem.

Theorem 11.12. *The* fundamental solution *to the heat equation $u_t = \gamma\Delta u$ corresponding to a unit delta function placed at position $(\xi,\eta) \in \mathbb{R}^2$ at the initial time $t_0 = 0$ is*

$$F(t,x,y;\xi,\eta) = \frac{1}{4\pi\gamma t}\,e^{-[(x-\xi)^2+(y-\eta)^2]/(4\gamma t)}. \tag{11.133}$$

Proof: Since we already know that both function (11.132) are solutions to the one-dimensional heat equation, Lemma 11.11 guarantees that their product, which equals (11.133), solves the two-dimensional heat equation for $t > 0$. Moreover, at the initial time,

$$u(0,x,y) = v(0,x)\,w(0,y) = \delta(x-\xi)\,\delta(y-\eta)$$

is a product of delta functions, and hence the result follows. Indeed, the total heat

$$\iint u(t,x,y)\,dx\,dy = \int_{-\infty}^{\infty} v(t,x)\,dx\,\int_{-\infty}^{\infty} w(t,y)\,dy = 1, \qquad t \ge 0,$$

remains constant, while

$$\lim_{t\to 0^+} u(t,x,y) = \begin{cases} \infty, & (x,y) = (\xi,\eta), \\ 0, & \text{otherwise}, \end{cases}$$

has the standard delta function limit at the initial time instant. *Q.E.D.*

Figure 11.8 depicts the evolution of the fundamental solution when $\gamma = 1$ at the indicated times. Observe that the initially concentrated temperature spreads out in a radially symmetric manner, while the total amount of heat remains constant. At any individual point $(x,y) \neq (0,0)$, the initially zero temperature rises slightly at first, but then decays monotonically back to zero at a rate proportional to $1/t$. As in the one-dimensional case, since the fundamental solution is > 0 for all $t > 0$, the heat energy has an infinite speed of propagation.

Both the one- and two-dimensional fundamental solutions have bell-shaped profiles known as *Gaussian filters*. The most important difference is the initial factor. In a one-dimensional medium, the fundamental solution decays in proportion to $1/\sqrt{t}$, whereas in the plane the decay is more rapid, being proportional to $1/t$. The physical explanation is that the heat energy is able to spread out in two independent directions, and hence diffuses

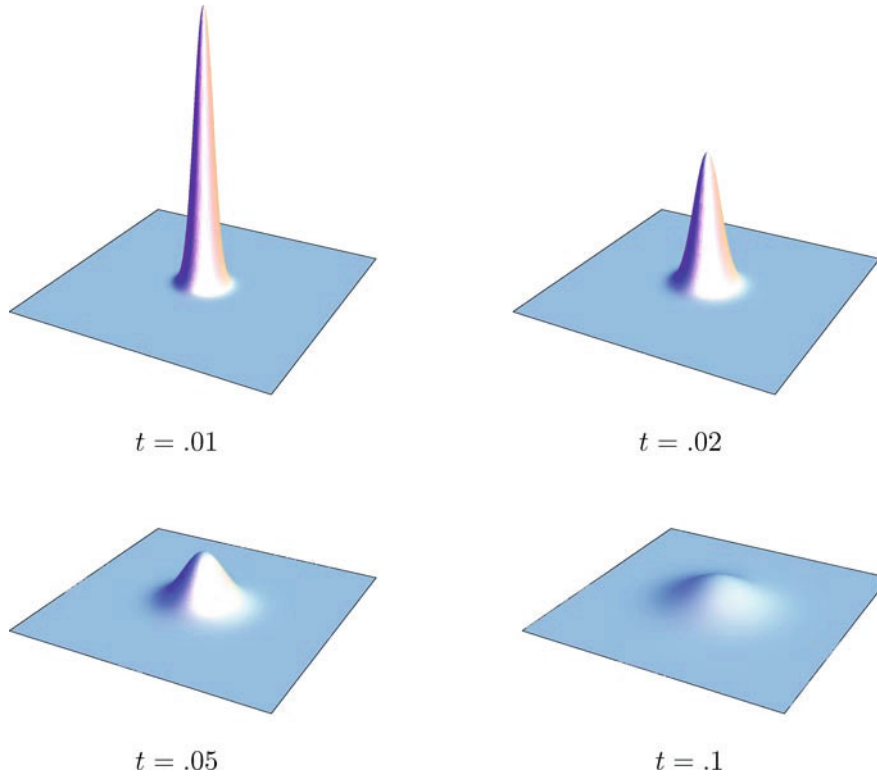


Figure 11.8. Fundamental solution of the planar heat equation.

away from its initial source more rapidly. As we shall see, the decay in three-dimensional space is more rapid still, being proportional to $t^{-3/2}$ for similar reasons; see (12.120).

The principal use of the fundamental solution is for solving the general initial value problem. We express the initial temperature distribution as a superposition of delta function impulses,

$$u(0, x, y) = f(x, y) = \iint f(\xi, \eta)\, \delta(x - \xi, y - \eta)\, d\xi\, d\eta,$$

where, at the point $(\xi, \eta) \in \mathbb{R}^2$, the impulse has magnitude $f(\xi, \eta)$. Linearity implies that the solution is then given by the same superposition of fundamental solutions.

Theorem 11.13. *The solution to the initial value problem*

$$u_t = \gamma\, \Delta u, \qquad u(t, x, y) = f(x, y), \qquad (x, y) \in \mathbb{R}^2,$$

for the planar heat equation is given by the linear superposition formula

$$u(t, x, y) = \frac{1}{4\pi\gamma t} \iint f(\xi, \eta)\, e^{-[(x-\xi)^2 + (y-\eta)^2]/(4\gamma t)}\, d\xi\, d\eta. \qquad (11.134)$$

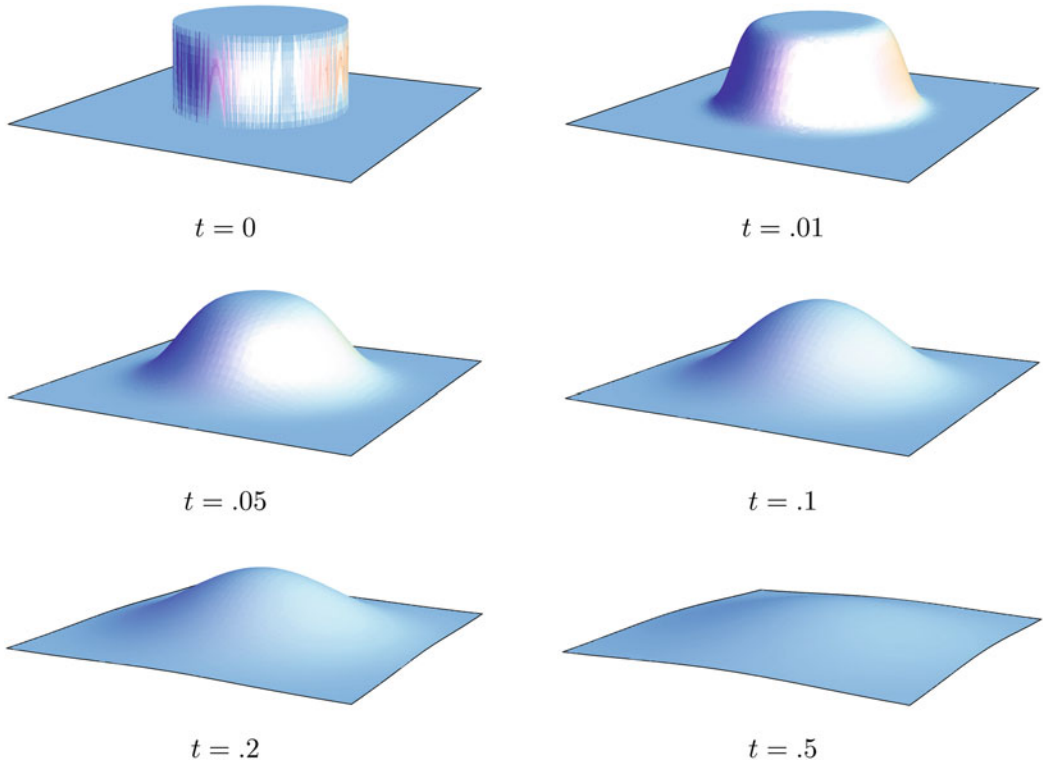


Figure 11.9. Diffusion of a disk.

We can interpret the solution formula (11.134) as a two-dimensional *convolution*

$$u(t, x, y) = F(t, x, y) * f(x, y) \tag{11.135}$$

of the initial data with a one-parameter family of progressively wider and shorter Gaussian filters

$$F(t, x, y) = F(t, x, y; 0, 0) = \frac{1}{4\pi\gamma t}\, e^{-(x^2+y^2)/(4\gamma t)}. \tag{11.136}$$

As in (7.54), such a convolution can be interpreted as a Gaussian weighted averaging of the function $f(x, y)$, which has the effect of smoothing out the initial data.

Example 11.14. If our initial temperature distribution is constant on a circular region, say

$$u(0, x, y) = \begin{cases} 1 & x^2 + y^2 < 1, \\ 0, & \text{otherwise}, \end{cases}$$

then the solution can be evaluated using (11.134), as follows:

$$u(t, x, y) = \frac{1}{4\pi t} \iint_D e^{-[(x-\xi)^2+(y-\eta)^2]/(4t)}\, d\xi\, d\eta,$$

where the integral is over the unit disk $D = \{\, \xi^2 + \eta^2 \le 1 \,\}$. Unfortunately, the integral cannot be expressed in terms of elementary functions. On the other hand, numerical

evaluation of the integral is straightforward. A plot of the resulting radially symmetric solution appears in Figure 11.9. One could also interpret this solution as the diffusion of an animal population in a uniform isotropic environment or bacteria in a similarly uniform large petri dish that are initially confined to a small circular region.

Exercises

11.5.1. Solve the following initial value problem: $u_t = 5\,(u_{xx} + u_{yy}), \quad u(0, x, y) = e^{-(x^2+y^2)}$.

11.5.2. Write down an integral formula for the solution to the following initial value problem:
$$u_t = 3(u_{xx} + u_{yy}), \qquad u(0, x, y) = (1 + x^2 + y^2)^{-2}.$$

11.5.3. At the initial time $t = 0$, a unit heat source is instantaneously applied at the origin of the (x, y)–plane. For $t > 0$, what is the maximum temperature experienced at a point $(x, y) \neq \mathbf{0}$? At what time is the maximum temperature achieved? Does the temperature approach an equilibrium value as $t \to \infty$? If so, how fast?

11.5.4. (*a*) Find an eigenfunction series representation of the fundamental solution for the heat equation $u_t = \Delta u$ on the unit square $\{\, 0 \leq x, y \leq 1 \,\}$ when subject to homogeneous Dirichlet boundary conditions. (*b*) Write the solution to the initial value problem $u(0, x, y) = f(x, y)$ in terms of the fundamental solution. (*c*) Discuss how your formula is related to the Fourier series solution (11.43).

11.5.5. Let $u(t, x, y)$ be a solution to the heat equation on all of $\mathbb{R}^2$ such that u and $\| \nabla u \| \to 0$ rapidly as $\| \mathbf{x} \| \to \infty$. (a) Prove that the total heat $H(t) = \iint u(t, x, y)\, dx\, dy$ is constant. (*b*) Explain how this can be reconciled with the statement that $u(t, x, y) \to 0$ as $t \to \infty$ at all points $(x, y) \in \mathbb{R}^2$.

♢ 11.5.6. Consider the initial value problem $u_t = \gamma\, \Delta u + H(t, x, y)$, $u(0, x, y) = 0$, for the inhomogeneous heat equation on the entire (x, y)–plane, where $H(t, x, y)$ represents a time-varying external heat source. Derive an integral formula for its solution. *Hint*: Mimic the solution method in Section 8.1.

11.5.7. A flat plate of infinite extent with unit thermal diffusivity starts off at 0°. From then on, a unit heat source is continually applied at the origin. Find the resulting temperature distribution. Does the temperature eventually reach a steady state? *Hint*: Use Exercise 11.5.6.

♡ 11.5.8. Building on Example 11.14, we model the "diffusion" of a set $D \subset \mathbb{R}^2$ as the solution $u(t, x, y)$ to the heat equation $u_t = \Delta u$ subject to the initial condition $u(0, x, y) = \chi_D(x, y)$, where $\chi_D(x, y) = \begin{cases} 1, & (x, y) \in D, \\ 0, & (x, y) \notin D, \end{cases}$ is the *characteristic function* of the set D.
(*a*) Write down a formula for the diffusion of the set D.
(*b*) *True or false*: At each t, the diffusion $u(t, x, y)$ is the characteristic function of a set D_t.
(*c*) Prove that $0 < u(t, x, y) < 1$ for all (x, y) and $t > 0$. (*d*) What is $\lim_{t \to \infty} u(t, x, y)$?
(*e*) Write down a formula for the diffusion of a unit square $D = \{\, 0 \leq x, y \leq 1 \,\}$, and then plot the result at several times. Discuss what you observe.

11.5.9. (*a*) Explain why the delta function on $\mathbb{R}^2$ satisfies the scaling law $\delta(x, y) = \beta^2\, \delta(\beta x, \beta\, y)$, for $\beta > 0$. (*b*) Verify that the fundamental solution to the heat equation on $\mathbb{R}^2$ obeys the same scaling law: $F(t, x, y) = \beta^2\, F(\beta^2\, t, \beta\, x, \beta\, y)$. (*c*) Is the fundamental solution a similarity solution?

11.5.10. (a) Find the fundamental solution on $\mathbb{R}^2$ to the cable equation $u_t = \gamma \Delta u - \alpha u$, where $\alpha > 0$ is constant. (b) Use your solution to write down a formula for the solution to the general initial value problem $u(0, x, y) = f(x, y)$ for $(x, y) \in \mathbb{R}^2$.

11.5.11. (a) Prove that if $v(t, x)$ and $w(t, x)$ solve the dispersive wave equation (8.90), then their product $u(t, x, y) = v(t, x)\, w(t, y)$ solves the two-dimensional dispersive equation

$$u_t + u_{xxx} + u_{yyy} = 0.$$

(b) What is the fundamental solution on $\mathbb{R}^2$ of the latter equation? (c) Write down an integral formula for the solution to the initial value problem $u(0, x, y) = f(x, y)$ for $(x, y) \in \mathbb{R}^2$.

11.5.12. Define the two-dimensional convolution $f * g$ of functions $f(x, y)$ and $g(x, y)$ so that equation (11.135) is valid.

11.6 The Planar Wave Equation

Let us next consider the two-dimensional wave equation

$$\frac{\partial^2 u}{\partial t^2} = c^2 \Delta u = c^2 \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right), \tag{11.137}$$

which models the unforced transverse vibrations of a homogeneous membrane, e.g., a drum. Here, $u(t, x, y)$ represents the vertical displacement of the membrane at time t and position $(x, y) \in \Omega$, where the domain $\Omega \subset \mathbb{R}^2$, assumed bounded, represents the undeformed shape. The constant $c^2 > 0$ encapsulates the membrane's physical properties — density, tension, stiffness, etc.; its square root, c, is called, as in the one-dimensional case, the *wave speed*, since it represents the speed of propagation of localized signals.

Remark: In this simplified model, we are only allowing small, transverse (vertical) displacements of the membrane. Large elastic vibrations lead to the nonlinear partial differential equations of elastodynamics, [**7**]. In particular, the bending vibrations of a flexible elastic plate are governed by a more complicated fourth-order partial differential equation.

The solution $u(t, x, y)$ to the wave equation will be uniquely specified once we impose suitable boundary and initial conditions. The Dirichlet conditions

$$u(t, x, y) = h(x, y), \qquad (x, y) \in \partial\Omega, \tag{11.138}$$

correspond to gluing our membrane to a fixed boundary — a rim; more generally, we can also allow h to depend on t, modeling a membrane attached to a moving boundary. On the other hand, the homogeneous Neumann conditions

$$\frac{\partial u}{\partial \mathbf{n}}\,(t, x, y) = 0, \qquad (x, y) \in \partial\Omega, \tag{11.139}$$

represent a free boundary where the membrane is not attached to any support — although in this model, its edge is allowed to move only in a vertical direction. Mixed boundary conditions attach part of the boundary and leave the remaining portion free to vibrate:

$$u = h \quad \text{on} \quad D \subsetneq \partial\Omega, \qquad \frac{\partial u}{\partial \mathbf{n}} = 0 \quad \text{on} \quad N = \partial\Omega \setminus D. \tag{11.140}$$

Since the wave equation is of second order in time, to uniquely specify the solution we need to impose two initial conditions,

$$u(0,x,y) = f(x,y), \qquad \frac{\partial u}{\partial t}\,(0,x,y) = g(x,y), \qquad (x,y) \in \Omega. \tag{11.141}$$

The first specifies the membrane's initial displacement, while the second prescribes its initial velocity.

Separation of Variables

Unfortunately, the d'Alembert solution method does not apply to the two-dimensional wave equation in any obvious manner. The reason is that, unlike the one-dimensional version (2.69), one cannot factorize the planar wave operator $\Box = \partial_t^2 - c^2\,\partial_x^2 - c^2\,\partial_y^2$, thus precluding any sort of reduction to a first-order partial differential equation. However, this is not the end of the story, and we will return to this issue at the end of Section 12.6.

We thus fall back on our universal solution tool for linear partial differential equations — separation of variables. According to the general framework established in Section 9.5, the separable solutions to the wave equation have the trigonometric form

$$u_k(t,x,y) = \cos(\omega_k\, t)\, v_k(x,y) \qquad \text{and} \qquad \widetilde{u}_k(t,x,y) = \sin(\omega_k\, t)\, v_k(x,y). \tag{11.142}$$

Substituting back into the wave equation, we find that $v_k(x,y)$ must be an eigenfunction of the associated Helmholtz boundary value problem

$$c^2 \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) + \lambda_k\, v = 0, \tag{11.143}$$

whose eigenvalue $\lambda_k = \omega_k^2$ equals the square of the vibrational frequency. According to Theorem 9.47, on a bounded domain, there is an infinite number of such *normal modes* with progressively faster vibrational frequencies: $\omega_k \to \infty$ as $k \to \infty$. In addition, in the positive semi-definite case — which occurs under homogeneous Neumann boundary conditions — there is a single constant null eigenfunction, leading to the additional separable solutions

$$u_0(t,x,y) = 1 \qquad \text{and} \qquad \widetilde{u}_0(t,x,y) = t. \tag{11.144}$$

The first represents a stationary membrane that has been displaced to a fixed height, while the second represents a membrane that is moving off in the vertical direction with constant unit speed. (Think of the membrane moving in outer space unaffected by any external gravitational force.) As in Section 9.5, the general solution can be written as an infinite series in the eigensolutions (11.142). Unfortunately, as we know, the Helmholtz boundary value problem can be explicitly solved only on a rather restricted class of domains. Here we will content ourselves with investigating the two most important cases: rectangular and circular membranes.

Remark: The vibrational frequencies represent the tones and overtones one hears when the drum membrane vibrates. An interesting question is whether two drums of different shapes can have identical sounds — the exact same vibrational frequencies. Or, more descriptively, can one "hear" the shape of a drum? It was not until 1992 that the answer was shown to be no, but for quite subtle reasons. See [**47**] for a discussion and some examples of differently shaped drums that have the same vibrational frequencies.

Vibration of a Rectangular Drum

Let us first consider the vibrations of a membrane in the shape of a rectangle

$$R = \{\, 0 < x < a,\ 0 < y < b \,\},$$

with side lengths a and b, whose edges are fixed to the (x, y)–plane. Thus, we seek to solve the wave equation

$$u_{tt} = c^2 \Delta u = c^2 (u_{xx} + u_{yy}), \qquad 0 < x < a, \qquad 0 < y < b, \tag{11.145}$$

subject to the initial and boundary conditions

$$\begin{aligned} u(t, 0, y) = u(t, a, y) = 0 = u(t, x, 0) = u(t, x, b), \qquad & 0 < x < a, \\ u(0, x, y) = f(x, y), \qquad u_t(0, x, y) = g(x, y), \qquad & 0 < y < b. \end{aligned} \tag{11.146}$$

As we saw in Section 11.2, the eigenfunctions and eigenvalues for the associated Helmholtz equation on a rectangle,

$$c^2 (v_{xx} + v_{yy}) + \lambda\, v = 0, \qquad (x, y) \in R, \tag{11.147}$$

when subject to the homogeneous Dirichlet boundary conditions

$$v(0, y) = v(a, y) = 0 = v(x, 0) = v(x, b), \qquad 0 < x < a, \qquad 0 < y < b, \tag{11.148}$$

are

$$v_{m,n}(x, y) = \sin \frac{m\,\pi\, x}{a}\ \sin \frac{n\,\pi\, y}{b}\,, \qquad \text{where} \qquad \lambda_{m,n} = \pi^2\, c^2 \left(\frac{m^2}{a^2} + \frac{n^2}{b^2} \right), \tag{11.149}$$

with $m, n = 1, 2, \ldots$. The fundamental frequencies of vibration are the square roots of the eigenvalues, so

$$\omega_{m,n} = \sqrt{\lambda_{m,n}} = \pi\, c \sqrt{\frac{m^2}{a^2} + \frac{n^2}{b^2}}\,, \qquad m, n = 1, 2, \ldots\ . \tag{11.150}$$

The frequencies will depend upon the underlying geometry — meaning the side lengths — of the rectangle, as well as the wave speed c, which, in turn, is a function of the membrane's density and stiffness. The higher the wave speed, or the smaller the rectangle, the faster the vibrations. In layman's terms, (11.150) quantifies the observation that smaller, stiffer drums made of less-dense material vibrate faster.

According to (11.142), the normal modes of vibration of our rectangle are

$$\begin{aligned} u_{m,n}(t, x, y) &= \cos \left(\pi\, c \sqrt{\frac{m^2}{a^2} + \frac{n^2}{b^2}}\ t \right)\ \sin \frac{m\,\pi\, x}{a}\ \sin \frac{n\,\pi\, y}{b}\,, \\ \widetilde{u}_{m,n}(t, x, y) &= \sin \left(\pi\, c \sqrt{\frac{m^2}{a^2} + \frac{n^2}{b^2}}\ t \right)\ \sin \frac{m\,\pi\, x}{a}\ \sin \frac{n\,\pi\, y}{b}\,. \end{aligned} \tag{11.151}$$

The general solution can then be written as a double Fourier series

$$u(t, x, y) = \sum_{m,n=1}^{\infty} \big[\, a_{m,n}\, u_{m,n}(t, x, y) + b_{m,n}\, \widetilde{u}_{m,n}(t, x, y) \,\big]$$

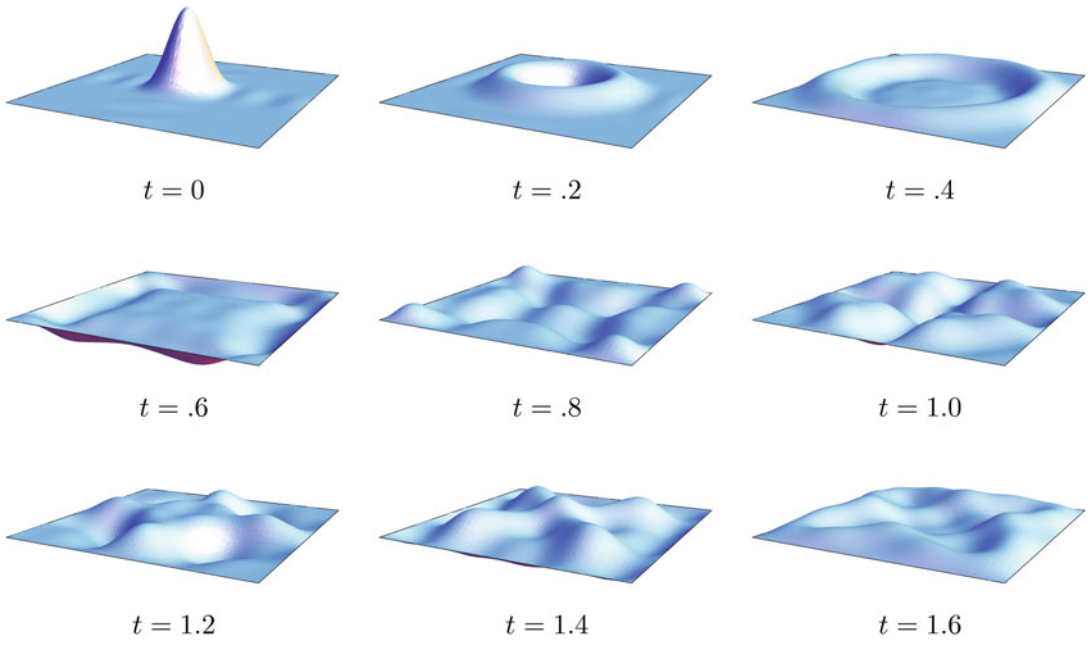


Figure 11.10. Vibrations of a square.

in the normal modes. The coefficients $a_{m,n}, b_{m,n}$ are fixed by the initial displacement $u(0, x, y) = f(x, y)$ and the initial velocity $u_t(0, x, y) = g(x, y)$. Indeed, the usual orthogonality relations among the eigenfunctions imply

$$a_{m,n} = \frac{\langle\, v_{m,n}\,, f\,\rangle}{\|\, v_{m,n}\,\|^2} = \frac{4}{a\,b}\int_0^b\int_0^a f(x,y)\,\sin\frac{m\,\pi\,x}{a}\,\sin\frac{n\,\pi\,y}{b}\,dx\,dy, \tag{11.152}$$

$$b_{m,n} = \frac{\langle\, v_{m,n}\,, g\,\rangle}{\omega_{m,n}\,\|\, v_{m,n}\,\|^2} = \frac{4}{\pi\,c\,\sqrt{m^2\,b^2 + n^2\,a^2}}\int_0^b\int_0^a g(x,y)\,\sin\frac{m\,\pi\,x}{a}\,\sin\frac{n\,\pi\,y}{b}\,dx\,dy.$$

Since the fundamental frequencies are not rational multiples of each other, the general solution is a genuinely quasiperiodic superposition of the various normal modes.

In Figure 11.10, we plot the solution resulting from the initially concentrated displacement[†]

$$u(0, x, y) = f(x, y) = e^{-100\,[\,(x-.5)^2+(y-.5)^2\,]}$$

at the center of a unit square, so $a = b = 1$, with unit wave speed, $c = 1$. Note that, unlike a concentrated displacement of a one-dimensional string, which remains concentrated at all subsequent times and periodically repeats, the initial displacement here spreads out in a radially symmetric manner and propagates to the edges of the rectangle, where it reflects

[†] The alert reader may object that the initial displacement $f(x, y)$ does not exactly satisfy the Dirichlet boundary conditions on the edges of the rectangle. But this does not prevent the existence of a well-defined (weak) solution to the initial value problem, whose initial boundary discontinuities will subsequently propagate into the square. However, here these are so tiny as to be unnoticeable in the solution graphs.

and then interacts with itself. Moreover, due to the quasiperiodicity of the solution, the drum's motion never exactly repeats, and the initially concentrated displacement never quite reforms.

Vibration of a Circular Drum

Let us next analyze the vibrations of a circular membrane of unit radius. In polar coordinates, the planar wave equation (11.137) takes the form

$$\frac{\partial^2 u}{\partial t^2} = c^2 \left(\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\,\frac{\partial u}{\partial r} + \frac{1}{r^2}\,\frac{\partial^2 u}{\partial \theta^2} \right). \tag{11.153}$$

We will again consider the homogeneous Dirichlet boundary value problem

$$u(t, 1, \theta) = 0, \qquad t \geq 0, \qquad -\pi \leq \theta \leq \pi, \tag{11.154}$$

along with initial conditions

$$u(0, r, \theta) = f(r, \theta), \qquad \frac{\partial u}{\partial t}\,(0, r, \theta) = g(r, \theta), \tag{11.155}$$

representing the initial displacement and velocity of the membrane. As always, we build up the general solution as a quasiperiodic linear combination of the normal modes as specified by the eigenfunctions for the associated Helmholtz boundary value problem.

As we saw in Section 11.2, the eigenfunctions of the Helmholtz equation on a disk of radius 1, say, subject to homogeneous Dirichlet boundary conditions, are products of trigonometric and Bessel functions:

$$\begin{aligned} v_{0,n}(r, \theta) &= J_0(\zeta_{0,n}\, r), \\ v_{m,n}(r, \theta) &= J_m(\zeta_{m,n}\, r)\, \cos m\theta, \qquad\qquad m, n = 1, 2, 3, \ldots . \\ \widetilde{v}_{m,n}(r, \theta) &= J_m(\zeta_{m,n}\, r)\, \sin m\theta, \end{aligned} \tag{11.156}$$

Here r, θ are the usual polar coordinates, while $\zeta_{m,n} > 0$ denotes the n^{th} (positive) root of the m^{th} order Bessel function $J_m(z)$, cf. (11.118). The corresponding eigenvalue is its square, $\lambda_{m,n} = \zeta_{m,n}^2$, and hence the natural frequencies of vibration are equal to the Bessel roots scaled by the wave speed:

$$\omega_{m,n} = c\,\sqrt{\lambda_{m,n}}\; = c\,\zeta_{m,n}. \tag{11.157}$$

A table of their values (for the case $c = 1$) can be found in the preceding section. The Bessel roots do not follow any easily discernible pattern, and are not rational multiples of each other. This result, known as *Bourget's hypothesis*, [**119**; p. 484], was rigorously proved by the German pure mathematician Carl Ludwig Siegel in 1929, [**106**]. Thus, the vibrations of a circular drum are also truly quasiperiodic, thereby providing a mathematical explanation of why drums sound dissonant.

The frequencies $\omega_{0,n} = c\,\zeta_{0,n}$ correspond to simple eigenvalues, with a single radially symmetric eigenfunction $J_0(\zeta_{0,n}\, r)$, while the "angular modes" $\omega_{m,n}$, for $m > 0$, are double, each possessing two linearly independent eigenfunctions (11.156). According to the general

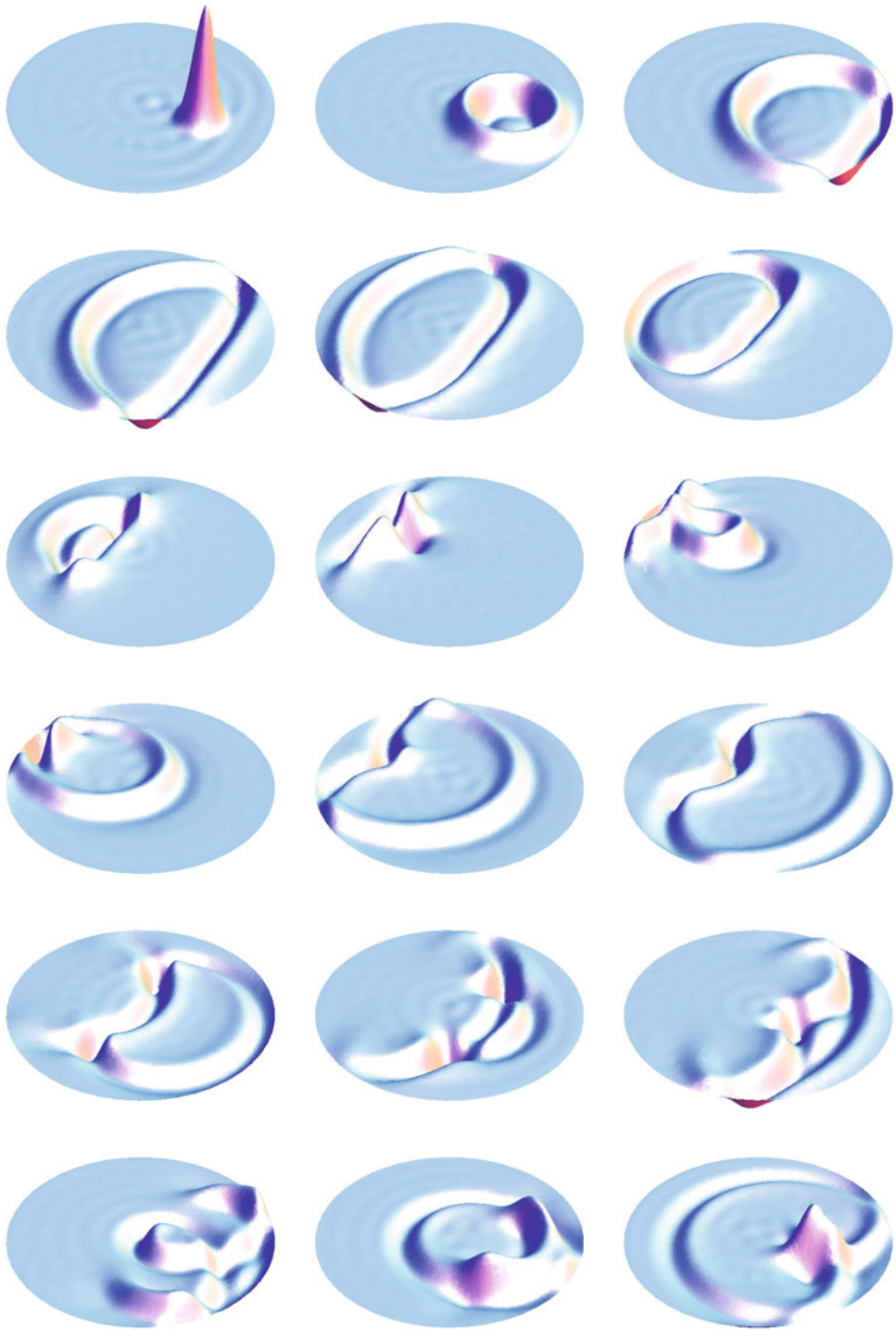


Figure 11.11. Vibrations of a disk.

formula (11.142), each eigenfunction engenders two independent normal modes of vibration, having the explicit forms

$$\begin{aligned} &\cos(c\,\zeta_{0,n}\,t)\; J_0(\zeta_{0,n}\,r), && \sin(c\,\zeta_{0,n}\,t)\; J_0(\zeta_{0,n}\,r),\\ &\cos(c\,\zeta_{m,n}\,t)\; J_m(\zeta_{m,n}\,r)\; \cos m\theta, && \sin(c\,\zeta_{m,n}\,t)\; J_m(\zeta_{m,n}\,r)\; \cos m\theta,\\ &\cos(c\,\zeta_{m,n}\,t)\; J_m(\zeta_{m,n}\,r)\; \sin m\theta, && \sin(c\,\zeta_{m,n}\,t)\; J_m(\zeta_{m,n}\,r)\; \sin m\theta. \end{aligned} \tag{11.158}$$

The general solution to (11.153–154) is then expressed as a Fourier–Bessel series:

$$\begin{aligned} u(t,r,\theta) = \frac{1}{2}\sum_{n=1}^{\infty} \big[\, a_{0,n}\cos(c\,\zeta_{0,n}\,t) + c_{0,n}\sin(c\,\zeta_{0,n}\,t)\,\big]\, J_0(\zeta_{0,n}\,r)\\ + \sum_{m,n=1}^{\infty} \big[\, \big(\, a_{m,n}\cos(c\,\zeta_{m,n}\,t) + c_{m,n}\sin(c\,\zeta_{m,n}\,t)\,\big)\cos m\theta\\ + \big(\, b_{m,n}\cos(c\,\zeta_{m,n}\,t) + d_{m,n}\sin(c\,\zeta_{m,n}\,t)\,\big)\sin m\theta\,\big]\, J_m(\zeta_{m,n}\,r), \end{aligned} \tag{11.159}$$

whose coefficients $a_{m,n}, b_{m,n}, c_{m,n}, d_{m,n}$ are determined, as usual, by the initial displacement and velocity of the membrane (11.155). In Figure 11.11, the vibrations due to an initially off-center concentrated displacement are displayed; the wave speed is $c = 1$, and the time interval between successive plots is $\Delta t = .3$. Again, the motion is only quasiperiodic and, no matter how long you wait, never quite returns to its original configuration.

Exercises

11.6.1. Use your physical intuition to decide whether the following statements are *true or false*. Then justify your answer.
(*a*) Increasing the stiffness of a membrane increases the wave speed.
(*b*) Increasing the density of a membrane increases the wave speed.
(*c*) Increasing the size of a membrane increases the wave speed

11.6.2. Two uniform membranes have the same shape, but are made out of different materials. Assuming that they are both subject to the same homogeneous boundary conditions, how are their vibrational frequencies related?

11.6.3. List the numerical values of the six lowest vibrational frequencies of a unit square with wave speed $c = 1$ when subject to homogeneous Dirichlet boundary conditions. How many linearly independent normal modes are associated with each of these frequencies?

♡ 11.6.4. The rectangular membrane $R = \{\, -1 < x < 1,\ 0 < y < 1\,\}$ has its two short sides attached to the (x, y)–plane, while its long sides are left free. The membrane is initially displaced so that its right half is one unit above, while its left half is one unit below the plane, and then released with zero initial velocity. (This discontinuous initial data serves to model a very sharp transition region.) Assume that the physical units are chosen so the wave speed $c = 1$. (*a*) Write down an initial-boundary value problem that governs the vibrations of the membrane. (*b*) What are the fundamental frequencies of vibration of the membrane? (*c*) Find the eigenfunction series solution that describes the subsequent motion of the membrane. (*d*) Is the motion (*i*) periodic? (*ii*) quasiperiodic? (*iii*) unstable? (*iv*) chaotic? Explain your answer.

11.6.5. Determine the solution to the following initial-boundary value problems for the wave equation on the rectangle $R = \{\, 0 < x < 2,\ 0 < y < 1\,\}$:

(*a*) $\begin{cases} u_{tt} = u_{xx} + u_{yy}, & u(t,x,0) = u(t,x,1) = u(t,0,y) = u(t,2,y) = 0,\\ u(0,x,y) = \sin \pi y, & u_t(0,x,y) = \sin \pi y; \end{cases}$

(*b*) $\begin{cases} u_{tt} = u_{xx} + u_{yy}, & u(t,x,0) = u(t,x,1) = \dfrac{\partial u}{\partial x}(t,0,y) = \dfrac{\partial u}{\partial x}(t,2,y) = 0, \\ u(0,x,y) = \sin \pi y, & u_t(0,x,y) = \sin \pi y; \end{cases}$

(*c*) $\begin{cases} u_{tt} = u_{xx} + u_{yy}, & u(t,x,0) = u(t,x,1) = u(t,0,y) = u(t,2,y) = 0, \\ u(0,x,y) = \begin{cases} 1, & 0 < x < 1, \\ 0, & 1 < x < 2, \end{cases} & u_t(0,x,y) = 0; \end{cases}$

(*d*) $\begin{cases} u_{tt} = 2\,u_{xx} + 2\,u_{yy}, & u(t,x,0) = u(t,x,1) = u(t,0,y) = u(t,2,y) = 0, \\ u(0,x,y) = 0, & u_t(0,x,y) = \begin{cases} 1, & 0 < x < 1, \\ 0, & 1 < x < 2. \end{cases} \end{cases}$

11.6.6. *True or false*: The more sides of a rectangle that are tied down, the faster it vibrates.

11.6.7. Answer Exercise 11.6.3 when (*a*) two adjacent sides of the square are tied down and the other two are left free; (*b*) two opposite sides of the square are tied down and the other two are left free; (*c*) the membrane is freely floating in outer space.

11.6.8. A square drum has two sides fixed to a support and two sides left free. Does the drum vibrate faster if the fixed and free sides are adjacent to each other or on opposite sides?

11.6.9. Write down a periodic solution to the wave equation on a unit square, subject to homogeneous Dirichlet boundary conditions, that is *not* a normal mode. Does it vibrate at a fundamental frequency?

11.6.10. A rectangular drum with side lengths 1 cm by 2 cm and unit wave speed $c = 1$ has its boundary fixed to the (x,y)–plane while subject to a periodic external forcing of the form $F(t,x,y) = \cos(\omega\, t)\, h(x,y)$. (*a*) At which frequencies ω will the forcing incite resonance in the drum? (*b*) If ω is a resonant frequency, write down the condition(s) on $h(x,y)$ that ensure excitation of a resonant mode.

11.6.11. The right half of a rectangle of side lengths 1 by 2 is initially displaced, while the left half is quiescent. *True or false*: The ensuing vibrations are restricted to the right half of the membrane.

♡ 11.6.12. A torus (inner tube) can be obtained by gluing together each of the two pairs of opposite sides of a rubber rectangle. The (small) vibrations of the torus are described by the following periodic initial-boundary value problem for the wave equation, in which x, y represent angular variables:

$$\begin{array}{lll} u_{tt} = c^2 \Delta u = c^2(u_{xx} + u_{yy}), & u(0,x,y) = f(x,y), & u_t(0,x,y) = g(x,y), \\ u(t,-\pi,y) = u(t,\pi,y), & u_x(t,-\pi,y) = u_x(t,\pi,y), & -\pi < x < \pi, \\ u(t,x,-\pi) = u(t,x,\pi), & u_x(t,x,-\pi) = u_x(t,x,\pi), & -\pi < y < \pi. \end{array}$$

(*a*) Find the fundamental frequencies and normal modes of vibration. (*b*) Write down a series for the solution. (*c*) Discuss the stability of a vibrating torus. Is the motion (*i*) periodic; (*ii*) quasiperiodic; (*iii*) chaotic; (*iv*) none of these?

11.6.13. The *forced wave equation* $u_{tt} = c^2 \Delta u + F(x,y)$ on a bounded domain $\Omega \subset \mathbb{R}^2$ models a membrane subject to a constant external forcing function $F(x,y)$. Write down an eigenfunction series solution to the forced wave equation when the membrane is subject to homogeneous Dirichlet boundary conditions and initial conditions $u(0,x,y) = f(x,y)$, $u_t(0,x,y) = g(x,y)$. *Hint*: Expand the forcing function in an eigenfunction series.

11.6.14. A circular drum of radius $\zeta_{0,1} \approx 2.4048$ has initial displacement and velocity

$$u(0,x,y) = 0, \qquad \frac{\partial u}{\partial t}(0,x,y) = 2\, J_0\big(\sqrt{x^2 + y^2}\,\big).$$

Assuming that the circular edge of the drum is fixed to the (x,y)–plane, describe, both qualitatively and quantitatively, its subsequent motion.

11.6.15. Write out the integral formulae for the coefficients in the Fourier–Bessel series solution (11.159) to the wave equation in a circular disk in terms of the initial data $u(0,r,\theta) = f(r,\theta)$, $u_t(0,r,\theta) = g(r,\theta)$.

11.6.16. A circular drum at rest is struck with a concentrated blow at its center. Write down

an eigenfunction series describing the resulting vibration.

♡ 11.6.17. (*a*) Set up and solve the initial-boundary value problem for the vibrations of a uniform circular drum of unit radius that is freely floating in space. (*b*) Discuss the stability of the drum's motion. (*c*) Are the vibrations slower or faster than when its edges are fixed to a plane?

11.6.18. A flat quarter-disk of radius 1 has its circular edge and one of its straight edges attached to the (x, y)–plane, while the other straight edge is left free. At time $t = 0$ the disk is struck with a hammer (unit delta function) at its midpoint, i.e., at radius $\frac{1}{2}$ and halfway between the straight edges. (*a*) Write down an initial-boundary value problem for the subsequent vibrations of the quarter-disk. *Hint*: Be careful with the form of the delta function in polar coordinates; see Exercise 6.3.6. (*b*) Assuming that the physical units are chosen so that the wave speed $c = 1$, determine the quarter-disk's vibrational frequencies. (*c*) Write down an eigenfunction series solution for the subsequent motion. (*d*) Is the motion unstable? periodic? If so, what is the period?

11.6.19. *True or false*: Assuming homogeneous Dirichlet boundary conditions, the fundamental frequencies of a vibrating half-disk are exactly twice those of the full disk of the same radius.

♡ 11.6.20. The edge of a circular drum is moved periodically up and down, so $u(t, 1, \theta) = \cos \omega t$. Assuming that the drum is initially at rest, discuss its response.

♣ 11.6.21. A drum is in the shape of a circular annulus with outer radius 1 meter and inner radius .5 meter. Find numerical values for its first three fundamental vibrational frequencies.

♡ 11.6.22. A homogeneous rope of length 1 and weight 1 is suspended from the ceiling. Taking x as the vertical coordinate, with $x = 1$ representing the fixed end and $x = 0$ the free end, the planar displacement $u(t, x)$ of the rope satisfies the initial-boundary value problem

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial}{\partial x}\left(x\,\frac{\partial u}{\partial x}\right), \qquad \begin{array}{ll} |\,u(t,0)\,| < \infty, & u(t,1) = 0, \\ u(0,x) = f(x), & \dfrac{\partial u}{\partial t}\,(0,x) = g(x), \end{array} \qquad t > 0, \quad 0 < x < 1.$$

(*a*) Find the solution. *Hint*: Let $y = \sqrt{x}$. (*b*) Are the vibrations periodic or quasiperiodic? (*c*) Describe the behavior of the rope when subject to uniform periodic external forcing

$$F(t, x) = a \cos \omega t.$$

Scaling and Symmetry

Symmetry methods can also be effectively employed in the analysis of the wave equation. Let us consider the simultaneous rescaling

$$t \longmapsto \alpha\, t, \qquad x \longmapsto \beta\, x, \qquad y \longmapsto \beta\, y, \tag{11.160}$$

of time and space, whose effect is to change the function $u(t, x, y)$ into a rescaled version

$$U(t, x, y) = u(\alpha\, t, \beta\, x, \beta\, y). \tag{11.161}$$

The chain rule is employed to relate their derivatives:

$$\frac{\partial^2 U}{\partial t^2} = \alpha^2\, \frac{\partial^2 u}{\partial t^2}\,, \qquad \frac{\partial^2 U}{\partial x^2} = \beta^2\, \frac{\partial^2 u}{\partial x^2}\,, \qquad \frac{\partial^2 U}{\partial y^2} = \beta^2\, \frac{\partial^2 u}{\partial y^2}\,.$$

Therefore, if u satisfies the wave equation

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

then U satisfies the rescaled wave equation

$$U_{tt} = \frac{\alpha^2 c^2}{\beta^2}\,\Delta U = C^2\,\Delta U, \quad \text{where the rescaled wave speed is} \quad C = \frac{\alpha\, c}{\beta}\,. \tag{11.162}$$

In particular, rescaling only time by setting $\alpha = 1/c$, $\beta = 1$, results in a unit wave speed $C = 1$. In other words, we are free to choose our unit of time measurement so as to fix the wave speed equal to 1.

If we set $\alpha = \beta$, scaling time and space in the same proportion, then the wave speed does not change, $C = c$, and so

$$t \;\longmapsto\; \beta\, t, \qquad\qquad x \;\longmapsto\; \beta\, x, \qquad\qquad y \;\longmapsto\; \beta\, y, \tag{11.163}$$

defines a *symmetry transformation* for the wave equation: If $u(t, x, y)$ is any solution to the wave equation, then so is its rescaled version

$$U(t, x, y) = u(\beta t, \beta x, \beta y) \tag{11.164}$$

for any choice of scale parameter $\beta \neq 0$. Observe that if $u(t, x, y)$ is defined on a domain Ω, then the rescaled solution $U(t, x, y)$ will be defined on the rescaled domain

$$\widetilde{\Omega} = \frac{1}{\beta}\,\Omega = \left\{ \left(\frac{x}{\beta}\,, \frac{y}{\beta} \right) \;\middle|\; (x, y) \in \Omega \right\} = \{\, (x, y) \mid\; (\beta x,\; \beta y) \in \Omega \,\}\,. \tag{11.165}$$

For instance, setting the scaling parameter $\beta = 2$ halves the size of the domain. The normal modes for the rescaled domain have the form

$$\begin{aligned} U_n(t, x, y) &= u_n(\beta t, \beta x, \beta y) = \cos(\beta\, \omega_n\, t)\; v_n(\beta x, \beta y), \\ \widetilde{U}_n(t, x, y) &= \widetilde{u}_n(\beta t, \beta x, \beta y) = \sin(\beta\, \omega_n\, t)\; v_n(\beta x, \beta y), \end{aligned}$$

and hence the rescaled vibrational frequencies are $\Omega_n = \beta\, \omega_n$. Thus, when $\beta < 1$, the rescaled membrane is larger by a factor $1/\beta$, and its vibrations are slowed down by the reciprocal factor β. For instance, a drum that is twice as large will vibrate twice as slowly, and hence have an octave lower overall tone. Musically, this means that all drums of a similar shape have the same pattern of overtones, differing only in their overall pitch, which is a function of their size, tautness, and density.

In particular, choosing $\beta = 1/R$ will rescale the unit disk into a disk of radius R. The fundamental frequencies of the rescaled disk are

$$\Omega_{m,n} = \beta\, \omega_{m,n} = \frac{c}{R}\, \zeta_{m,n}, \tag{11.166}$$

where c is the wave speed and $\zeta_{m,n}$ are the Bessel roots, defined in (11.118). Observe that the ratios $\omega_{m,n}/\omega_{m',n'}$ between vibrational frequencies remain the same, independent of the size of the disk R and the wave speed c. We define the *relative vibrational frequencies*

$$\rho_{m,n} = \frac{\omega_{m,n}}{\omega_{0,1}} = \frac{\zeta_{m,n}}{\zeta_{0,1}}\,, \qquad \text{in proportion to} \qquad \omega_{0,1} = \frac{c\, \zeta_{0,1}}{R} \approx 2.4\, \frac{c}{R}\,, \tag{11.167}$$

which is the drum's dominant, or lowest, vibrational frequency. The relative frequencies $\rho_{m,n}$ are independent of the size, stiffness or composition of the drum membrane. In the following table, we display a list of all relative vibrational frequencies (11.167) that are < 6. Once the lowest frequency $\omega_{0,1}$ has been determined — either theoretically, numerically,

or experimentally — all the higher overtones $\omega_{m,n} = \rho_{m,n}\,\omega_{0,1}$ are simply obtained by rescaling.

Relative Vibrational Frequencies of a Circular Disk

n \ m	0	1	2	3	4	5	6	7	8	9	...
1	1.000	1.593	2.136	2.653	3.155	3.647	4.132	4.610	5.084	5.553	...
2	2.295	2.917	3.500	4.059	4.601	5.131	5.651	⋮	⋮	⋮	
3	3.598	4.230	4.832	5.412	5.977	⋮	⋮				
4	4.903	5.540	⋮	⋮	⋮						
⋮	⋮	⋮									

Exercises

11.6.23. *True or false*: Two rectangular membranes, made out of the same material and both subject to Dirichlet boundary conditions, have the same relative vibrational frequencies if and only if they are have similar shapes.

11.6.24. *True or false*: (*a*) The vibrational frequencies of a square with side lengths $a = b = 2$ are four times as slow as those of a square with side lengths $a = b = 1$.
(*b*) The vibrational frequencies of a rectangle with side lengths $a = 2$, $b = 1$, are twice as slow as those of a square with side lengths $a = b = 1$.

11.6.25. A vibrating rectangle of unknown size has wave speed $c = 1$ and is subject to homogeneous Dirichlet boundary conditions. How many of its lowest vibrational frequencies do you need to know in order to determine the size of the rectangle?

11.6.26. Answer Exercise 11.6.25 when the rectangle is subject to homogeneous Neumann boundary conditions.

♣ 11.6.27. A circular drum has the A above middle C, which has a frequency of 440 Hertz, as its lowest tone. What notes are the first five overtones nearest? Try playing these on a piano or guitar. Or, if you have a synthesizer, try assembling notes of these frequencies to see how closely it reproduces the dissonant sound of a drum.

11.6.28. In an orchestra, a circular snare drum of radius 1 foot sits near a second circular drum made out of the same material. Vibrations of the first drum are observed to excite an undesired resonant vibration in its partner. What are the possible radii of the second drum?

11.6.29. *True or false*: The relative vibrational frequencies of a half-disk, subject to Dirichlet boundary conditions, are a subset of the relative vibrational frequencies of a full disk.

11.6.30. *True or false*: If $u(t,x,y) = \cos(\omega t)\,v(x,y)$ is a normal mode of vibration for a unit square subject to homogeneous Dirichlet boundary conditions, then the function $\widehat{u}(t,x,y) = \cos(\omega t)\,v\big(\frac{1}{2}\,x, \frac{1}{3}\,y\big)$ is a normal mode of vibration for a 2×3 rectangle that is subject to the same boundary conditions, but with a possibly different wave speed. If true, how are the wave speeds of the two rectangles related?

11.6.31. Prove that if $u(t,x,y)$ is a solution to the two-dimensional wave equation, so is the translated function $U(t,x,y) = u(t - t_0, x - x_0, y - y_0)$, for any constants t_0, x_0, y_0.

♢ 11.6.32. (*a*) Prove that if $u(t, x, y)$ solves the wave equation, so does $U(t, x, y) = u(-t, x, y)$. Thus, unlike the heat equation, the wave equation is time-reversible, and its solutions can be unambiguously followed backwards in time. (*b*) Suppose $u(t, x, y)$ solves the initial value problem (11.141). Write down the initial value problem satisfied by $U(t, x, y)$.

11.6.33. (*a*) Prove that, on $\mathbb{R}^2$, the solution to the pure displacement initial value problem $u_{tt} = c^2 \Delta u$, $u(0, x, y) = f(x, y)$, $u_t(0, x, y) = 0$, is an even function of t.

(*b*) Prove that the solution to the pure velocity initial value problem $u_{tt} = c^2 \Delta u$, $u(0, x, y) = 0$, $u_t(0, x, y) = g(x, y)$, is an odd function of t. *Hint*: Use Exercise 11.6.32 and uniqueness of solutions to the initial value problem.

11.6.34. Suppose $v(t, x)$ is any solution to the one-dimensional wave equation $v_{tt} = v_{xx}$. Prove that $u(t, x, y) = v(t, a\,x + b\,y)$, for any constants $(a, b) \neq (0, 0)$, solves the two-dimensional wave equation $u_{tt} = c^2(u_{xx} + u_{yy})$ for some choice of wave speed. Describe the behavior of such solutions.

11.6.35. A *traveling-wave solution* to the two-dimensional wave equation has the form $u(t, x, y) = v(x - a\,t, y - a\,t)$, where a is a constant. Find the partial differential equation satisfied by the function $v(\xi, \eta)$. Is the equation hyperbolic?

11.6.36. Is the counterpart of Lemma 11.11 valid for the wave equation? In other words, if $v(t, x)$ and $w(t, x)$ are any two solutions to the one-dimensional wave equation, is their product $u(t, x, y) = v(t, x)\, w(t, y)$ a solution to the two-dimensional wave equation?

11.6.37. (*a*) How would you solve an initial-boundary value problem for the wave equation on a rectangle that is not aligned with the coordinate axes? (*b*) Apply your method to set up and solve an initial-boundary value problem on the square $R = \{\, |\, x + y\, | < 1,\ |\, x - y\, | < 1\,\}$.

Chladni Figures and Nodal Curves

When a membrane vibrates, its individual atoms typically move up and down in a quasiperiodic manner. As such, there is little correlation between their motions at different locations. However, if the membrane is set to vibrate in a pure eigenmode, say

$$u_n(t, x, y) = \cos(\omega_n\, t)\, v_n(x, y), \tag{11.168}$$

then all points move up and down at a common frequency $\omega_n = \sqrt{\lambda_n}$, which is the square root of the eigenvalue corresponding to the eigenfunction $v_n(x, y)$. The exceptions are the points where the eigenfunction vanishes:

$$v_n(x, y) = 0, \tag{11.169}$$

which remain stationary. The set of all points $(x, y) \in \Omega$ that satisfy (11.169) is known as the n^{th} *Chladni figure* of the domain Ω, named in honor of the eighteenth-century German physicist and musician Ernst Chladni who first observed them experimentally by exciting a metal plate with his violin bow, [**43**]. The mathematical models governing such vibrating plates were formulated by the French mathematician Sophie Germain in the early 1800s. It can be shown that, in general, each Chladni figure consists of a finite system of *nodal curves*, [**34**, **43**], that partition the membrane into disjoint *nodal regions*. As the membrane vibrates, the nodal curves remain stationary, while each nodal region is entirely either above or below the equilibrium plane, except, momentarily, when the *entire* membrane has zero displacement. As Chladni discovered in his original experiments, scattering small

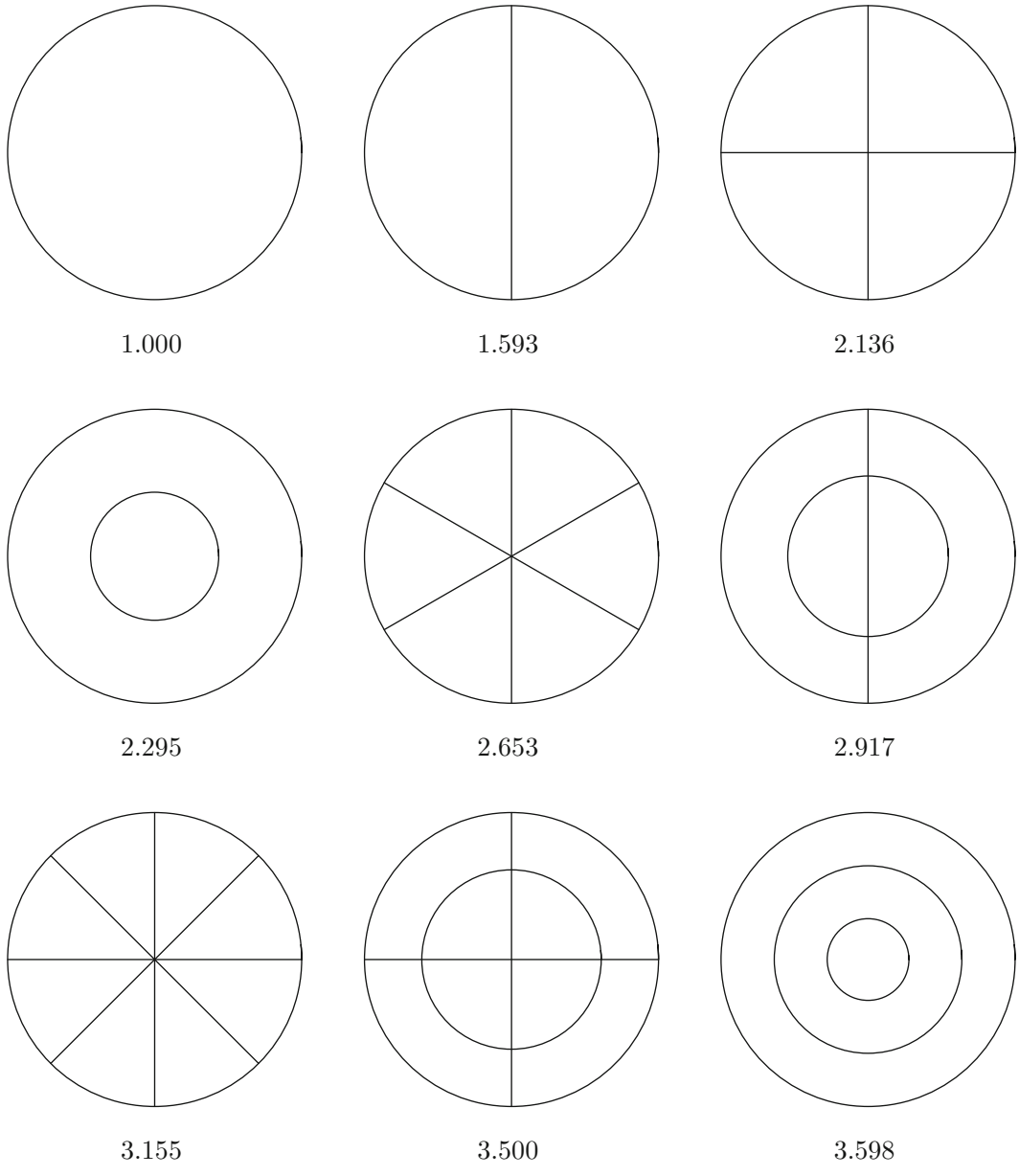


Figure 11.12. Nodal curves and relative vibrational frequencies of a circular membrane.

particles (e.g., fine sand) over a membrane or plate vibrating in an eigenmode will enable us to visualize the Chladni figure, because the particles will tend to accumulate along the stationary nodal curves. Adjacent nodal regions, lying on the opposite sides of a nodal curve, move in opposing directions — when one is up, its neighbors are down, and then they switch roles as the membrane becomes momentarily flat. Let us look at a couple of examples where the Chladni figures can be readily determined.

Example 11.15. *Circular Drums.* Since the eigenfunctions (11.156) for a disk are products of trigonometric functions in the angular variable and Bessel functions of the radius, the nodal curves for the normal modes of vibrations of a circular membrane are rays emanating from and circles centered at the origin. Consequently, the nodal regions are annular sectors. Chladni figures associated with the first nine normal modes, indexed by their relative frequencies, are plotted in Figure 11.12. Representative displacements of the membrane in each of the first twelve modes can be found earlier, in Figure 11.6. The dominant (lowest frequency) mode is the only one that has no nodal curves; it has the form of a radially symmetric bump where the entire membrane flexes up and down. The next lowest modes vibrate proportionally faster at a relative frequency $\rho_{1,1} \approx 1.593$. The most general solution with this vibrational frequency is a linear combination of the two eigensolutions: $\alpha\, u_{1,1} + \beta\, \widetilde{u}_{1,1}$. Each such combination has a single diameter as a nodal curve, whose angle with the horizontal depends on the ratio β/α. The two semicircular halves of the drum vibrate in opposing directions — when the top half is up, the bottom half is down and vice versa. The next set of modes have two perpendicular diameters as nodal curves; the four quadrants of the drum vibrate in tandem, with opposite quadrants moving in the same direction. Next in increasing order of vibrational frequency is a single mode, which has a circular nodal curve whose (relative) radius equals the ratio of the first two roots of the order zero Bessel function, $\zeta_{0,2}/\zeta_{0,1} \approx .43565$; see Exercise 11.6.39 for a justification. In this case, the inner disk and the outer annulus vibrate in opposing directions. And so on . . . .

Example 11.16. *Rectangular Drums.* For most rectangular drums, the Chladni figures are relatively uninteresting. Since the normal modes (11.151) are separable products of trigonometric functions in the coordinate variables x, y, the nodal curves are equally spaced straight lines parallel to the sides of the rectangle. The internodal regions are smaller rectangles, of identical size and shape, with adjacent rectangles vibrating in opposite directions.

More interesting figures appear when the rectangle admits multiple eigenvalues — so-called *accidental degeneracies*. Note that two of the eigenvalues (11.149) coincide, $\lambda_{m,n} = \lambda_{k,l}$, if and only if

$$\frac{m^2}{a^2} + \frac{n^2}{b^2} = \frac{k^2}{a^2} + \frac{l^2}{b^2}\,, \tag{11.170}$$

where $(m,n) \neq (k,l)$ are distinct pairs of positive integers. In such situations, the two eigenmodes happen to vibrate with a common frequency $\omega = \omega_{m,n} = \omega_{k,l}$. Consequently, any linear combination of the eigenmodes, e.g.,

$$\cos(\omega\, t)\left(\alpha \sin\frac{m\,\pi\, x}{a}\,\sin\frac{n\,\pi\, y}{b} + \beta \sin\frac{k\,\pi\, x}{a}\,\sin\frac{l\,\pi\, y}{b}\right), \qquad \alpha, \beta \in \mathbb{R},$$

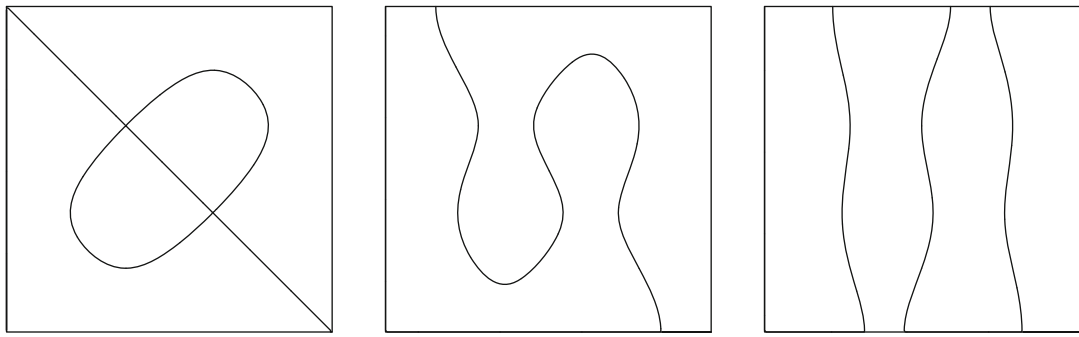
is also a pure vibration, and hence qualifies as a normal mode. The associated nodal curves,

$$\alpha \sin\frac{m\,\pi\, x}{a}\,\sin\frac{n\,\pi\, y}{b} + \beta \sin\frac{k\,\pi\, x}{a}\,\sin\frac{l\,\pi\, y}{b} = 0, \qquad \begin{array}{l} 0 \le x \le a, \\ 0 \le y \le b, \end{array} \tag{11.171}$$

have a more intriguing geometry, which can change dramatically as the coefficients α, β vary.

For example, on the unit square $R = \big\{\, 0 < x, y < 1 \,\big\}$, an accidental degeneracy occurs whenever

$$m^2 + n^2 = k^2 + l^2 \tag{11.172}$$



$\alpha = \beta = 1$ $\alpha = 2, \ \beta = 1$ $\alpha = 5, \ \beta = 1$

Figure 11.13. Some Chladni figures for a square membrane.

for distinct pairs of positive integers $(m, n) \neq (k, l)$. The simplest possibility arises whenever $m \neq n$, in which case we can merely reverse the order, setting $k = n$, $l = m$. In Figure 11.13 we plot three sample nodal curves

$$\alpha \sin 4\pi x \ \sin \pi y + \sin \pi x \ \sin 4\pi y = 0,$$

corresponding to three different linear combinations of the eigenfunctions with $m = l = 4$, $n = k = 1$. The associated vibrational frequency is, in all cases, $\omega_{4,1} = c\sqrt{17}\,\pi$, where c is the wave speed.

Classifying accidental degeneracies of rectangles takes us into the realm of number theory, [**9**, **29**]. In the case of a square, equation (11.172) is asking us to locate all integer points $(m, n) \in \mathbb{Z}^2$ that lie on a common circle.

Remark: Bourget's hypothesis, mentioned after (11.157), implies that $\zeta_{m,n} \neq \zeta_{k,l}$ whenever $(m, n) \neq (k, l)$. This implies that a disk has no accidental degeneracies, and hence all its nodal curves are concentric circles and diameters.

Exercises

♢ 11.6.38. Suppose that a membrane is vibrating in a normal mode. Prove that the membrane lies instantaneously completely flat at regular time intervals.

♢ 11.6.39. For a vibrating disk of unit radius, determine the radius of the circular nodal curve for the next-to-lowest circular mode.

11.6.40. Order the five nodal circles displayed in Figure 11.12 according to their size.

11.6.41. Sketch the Chladni figures in a unit disk corresponding to the following vibrational frequencies. Determine numerical values for the radii of any circular nodal curves.
(*a*) $\omega_{4,0}$, (*b*) $\omega_{4,2}$, (*c*) $\omega_{2,4}$, (*d*) $\omega_{3,3}$, (*e*) $\omega_{1,5}$.

11.6.42. *True or false*: Any diameter of a circular disk is a nodal curve for some normal mode.

11.6.43. *True or false*: The nodal curves on a semicircular disk are all semicircles and rays emanating from the center.

11.6.44. (*a*) Find the smallest distinct pair of positive integers $(k, l) \neq (m, n)$ satisfying (11.172) that are not obtained by simply reversing the order, i.e., $(k, l) \neq (n, m)$. (*b*) Find the next-smallest example. (*c*) Plot two or three Chladni figures arising from such degenerate eigenfunctions.

♡ 11.6.45. Let R be a rectangle all of whose sides are fixed to the (x, y)–plane. Suppose that all its nodal curves are straight lines. What can you say about its side lengths a, b?

11.6.46. *True or false*: The nodal regions of a vibrating rectangle are similarly shaped rectangles.

♢ 11.6.47. Prove that any point of intersection (x_0, y_0) of two nodal curves associated with the same normal mode is a critical point of the associated eigenfunction: $\nabla v(x_0, y_0) = \mathbf{0}$.

11.6.48. *True or false*: The nodal curves on a domain do not depend on the choice of boundary conditions.

Chapter 12
Partial Differential Equations in Space

At last we have ascended to the ultimate rung of the dimensional ladder (at least for those of us living in a three-dimensional universe): partial differential equations in physical space. As in the one- and two-dimensional settings developed in the preceding chapters, the main protagonists are the Laplace and Poisson equations, modeling equilibrium configurations of solid bodies; the three-dimensional wave equation, governing vibrations of solids, liquids, and electromagnetic waves; and the three-dimensional heat equation, modeling spatial diffusion processes. To conclude this chapter — and the book — we will also analyze the particular three-dimensional Schrödinger equation that governs the hydrogen atom, and thereby characterizes atomic orbitals.

Fortunately, almost everything of importance has already appeared in the previous chapters, and appending a third dimension is, for the most part, simply a matter of appropriately adapting the constructions. We have already developed the principal solution techniques: separation of variables, Green's functions, and fundamental solutions. In three-dimensional problems, separation of variables is applicable in a variety of coordinate systems, including the usual rectangular, cylindrical, and spherical coordinates. The first two do not lead to anything fundamentally new, and are therefore relegated to the exercises. Separation in spherical coordinates requires spherical Bessel functions and spherical harmonics, which play essential roles in a wide variety of physical systems, both classical and quantum.

The Green's function for the three-dimensional Poisson equation in space can be identified as the classic Newton (Coulomb) $1/r$ gravitational (electrostatic) potential. The fundamental solution for the three-dimensional heat equation can be easily guessed from its one- and two-dimensional forms. The three-dimensional wave equation, surprisingly, has an explicit solution formula, named after Kirchhoff, of electrical fame, but originally due to Poisson. Counterintuitively, the best way to handle the two-dimensional wave equation is by "descending" from the simpler(!) three-dimensional Kirchhoff formula. Descent reveals a remarkable difference between waves in planar and spatial media. Huygens' Principle states that three-dimensional waves emanating from a localized initial disturbance remain localized as they propagate through space. In contrast, initially concentrated two-dimensional disturbances leave a slowly decaying remnant that never entirely disappears.

The final section is concerned with the Schrödinger equation for a hydrogen atom, that is, the quantum-dynamical system governing the spatial motion of a single electron around a positively charged nucleus. As we will see, the spherical harmonic eigensolutions account for the observed quantum energy levels of atoms that underly the periodic table and hence the foundations of molecular chemistry.

P.J. Olver, *Introduction to Partial Differential Equations*, Undergraduate Texts in Mathematics,
DOI 10.1007/978-3-319-02099-0_12,

12.1 The Three–Dimensional Laplace and Poisson Equations

We begin our investigations, as usual, with systems in equilibrium, deferring dynamics until later. The prototypical equilibrium system is the three-dimensional Laplace equation

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0, \tag{12.1}$$

in which $\mathbf{x} = (\, x, y, z \,)^T$ represents rectangular coordinates on $\mathbb{R}^3$. The solutions $u(x, y, z)$ continue to be known as *harmonic functions*. The Laplace equation models unforced equilibria; *Poisson's equation* is the inhomogeneous version

$$-\Delta u = f(x, y, z), \tag{12.2}$$

whose right-hand side represents some form of external forcing.

The basic boundary value problem for the Laplace and Poisson equations seeks a solution inside a bounded domain $\Omega \subset \mathbb{R}^3$ subject to either *Dirichlet boundary conditions*, prescribing the function values on the domain's boundary:

$$u = h \qquad \text{on} \qquad \partial\Omega, \tag{12.3}$$

or *Neumann boundary conditions*, prescribing its normal derivative or flux through the boundary:

$$\frac{\partial u}{\partial \mathbf{n}} = k \qquad \text{on} \qquad \partial\Omega, \tag{12.4}$$

or *mixed boundary conditions*, in which one imposes Dirichlet conditions on part of the boundary and Neumann conditions on the remainder. Keep in mind that the boundary of the solid domain Ω consists of one or more piecewise smooth closed surfaces, which will be oriented by use of the outward — meaning exterior to the domain — unit normal $\mathbf{n}$.

The boundary value problems for the three-dimensional Laplace and Poisson equations govern a wide variety of physical systems, including:

- *Heat conduction*: The solution u represents the equilibrium temperature in a solid body. The inhomogeneity f represents some form of internal heat source or sink. Dirichlet conditions correspond to fixing the temperature on the bounding surface(s), whereas homogeneous Neumann conditions correspond to an insulated boundary, i.e., one that does not allow any heat flux.
- *Ideal fluid flow*: Here the solution u to the Laplace equation represents the velocity potential for an incompressible, irrotational steady-state fluid flow inside a container governed by the velocity vector field $\mathbf{v} = \nabla u$. Homogeneous Neumann boundary conditions correspond to a solid boundary that the fluid cannot penetrate.
- *Elasticity*: In certain restricted contexts, u represents an equilibrium deformation of a solid body, e.g., the radial deformation of an elastic ball.
- *Electrostatics*: In applications to electromagnetism, u is the electric potential in a conducting medium; its gradient ∇u prescribes the electromotive force on a charged particle. The inhomogeneity f represents an external electrostatic force field.
- *Gravitation*: The Newtonian gravitational potential in flat empty space is also prescribed by the Laplace equation. (In contrast, Einstein's theory of general relativity requires a vastly more complicated nonlinear system of partial differential equations, [**75**].)

Self–Adjoint Formulation and Minimum Principle

The Laplace and Poisson equations naturally fit into the general self-adjoint equilibrium framework summarized in Chapter 9. We introduce the L^2 inner products

$$\begin{aligned} \langle\, u\,, \widetilde{u}\,\rangle &= \iiint_\Omega u(x,y,z)\, \widetilde{u}(x,y,z)\, dx\, dy\, dz, \\ \langle\, \mathbf{v}\,, \widetilde{\mathbf{v}}\,\rangle &= \iiint_\Omega \mathbf{v}(x,y,z)\cdot \widetilde{\mathbf{v}}(x,y,z)\, dx\, dy\, dz, \end{aligned} \tag{12.5}$$

between, respectively, scalar fields $u, \widetilde{u}$, and vector fields $\mathbf{v}, \widetilde{\mathbf{v}}$, which are defined on the domain $\Omega \subset \mathbb{R}^3$. We assume that the functions in question are sufficiently nice in order that these inner products be well defined; if Ω is unbounded, this, in essence, requires that they decay reasonably rapidly to zero at large distances.

When subject to suitable homogeneous boundary conditions, the three-dimensional Laplace equation can be placed in our standard self-adjoint form

$$-\Delta u = -\nabla\cdot\nabla u = \nabla^* \circ \nabla u. \tag{12.6}$$

This relies on the fact that the adjoint of the gradient operator with respect to the L^2 inner products (12.5) is minus the divergence operator:

$$\nabla^* \mathbf{v} = -\nabla\cdot\mathbf{v}. \tag{12.7}$$

As usual, the determination of the adjoint rests on an integration by parts formula, which, in three-dimensional space, is a consequence of the Divergence Theorem from multivariable calculus, [**8**, **108**]:

Theorem 12.1. *Let $\Omega \subset \mathbb{R}^3$ be a bounded domain whose boundary $\partial\Omega$ consists of one or more piecewise smooth simple closed surfaces. Let $\mathbf{n}$ denote the unit* outward *normal to the boundary of Ω. Let $\mathbf{v}$ be a C^1 vector field defined on Ω and continuous up to its boundary. Then the surface integral, with respect to surface area, of the normal component of $\mathbf{v}$ over the boundary of the domain equals the triple integral of its divergence over the domain*:

$$\iint_{\partial\Omega} \mathbf{v}\cdot\mathbf{n}\; dS = \iiint_\Omega \nabla\cdot\mathbf{v}\; dx\, dy\, dz. \tag{12.8}$$

Replacing $\mathbf{v}$ by the product $u\,\mathbf{v}$ of a scalar field u and a vector field $\mathbf{v}$ yields

$$\iiint_\Omega (u\,\nabla\cdot\mathbf{v} + \nabla u\cdot\mathbf{v})\, dx\, dy\, dz = \iiint_\Omega \nabla\cdot(u\,\mathbf{v})\, dx\, dy\, dz = \iint_{\partial\Omega} u\,(\mathbf{v}\cdot\mathbf{n})\, dS. \tag{12.9}$$

Rearranging the terms produces the desired *integration by parts formula* for triple integrals:

$$\iiint_\Omega (\nabla u\cdot\mathbf{v})\, dx\, dy\, dz = \iint_{\partial\Omega} u\,(\mathbf{v}\cdot\mathbf{n})\, dS - \iiint_\Omega u\,(\nabla\cdot\mathbf{v})\, dx\, dy\, dz. \tag{12.10}$$

The boundary surface integral will vanish, provided either $u = 0$ or $\mathbf{v}\cdot\mathbf{n} = \mathbf{0}$ at each point on $\partial\Omega$. When $u = 0$ on all of $\partial\Omega$, we have homogeneous Dirichlet conditions. Setting $\mathbf{v}\cdot\mathbf{n} = \mathbf{0}$ everywhere on $\partial\Omega$ results in the homogeneous Neumann boundary value problem owing to the identification of $\mathbf{v} = \nabla u$. Finally, the mixed boundary value problem takes $u = 0$ on part of $\partial\Omega$ and $\mathbf{v}\cdot\mathbf{n} = \mathbf{0}$ on the rest. Thus, subject to one of these choices, the integration by parts formula (12.10) reduces to

$$\langle\, \nabla u\,, \mathbf{v}\,\rangle = \langle\, u\,, -\nabla\cdot\mathbf{v}\,\rangle, \tag{12.11}$$

which suffices to establish the adjoint formula (12.7).

Remark: Adopting more general weighted inner products results in a more general elliptic boundary value problem. See Exercise 12.1.9 for details.

According to Theorem 9.20, the self-adjoint formulation (12.6) automatically implies positive semi-definiteness of the boundary value problem, with positive definiteness if $\ker \nabla = \{0\}$. Since, on a connected domain, only constant functions are annihilated by the gradient operator — see Lemma 6.16, which also applies to three-dimensional domains — both the Dirichlet and mixed boundary value problems are positive definite, while the Neumann boundary value problem is only positive semi-definite.

Finally, in the positive definite cases, Theorem 9.26 implies that the solution can be characterized by the three-dimensional version of the Dirichlet minimization principle (9.82).

Theorem 12.2. *The solution $u(x, y, z)$ to the Poisson equation* (12.2) *subject to homogeneous Dirichlet or mixed boundary conditions* (12.3) *is the unique function that minimizes the* Dirichlet integral

$$\tfrac{1}{2} \,|||\, \nabla u \,|||^2 - \langle\, u \,,\, f \,\rangle = \iiint_\Omega \left[\, \tfrac{1}{2}\,(u_x^2 + u_y^2 + u_z^2) - f\,u \,\right] dx\, dy\, dz \tag{12.12}$$

among all C^2 *functions that satisfy the prescribed boundary conditions.*

As in the two-dimensional version discussed in Chapter 9, the Dirichlet minimization principle continues to hold in the case of the inhomogeneous Dirichlet boundary value problem. Modifications for the inhomogeneous mixed boundary value problem appear in Exercise 12.1.13.

Exercises

12.1.1. Find bases for the following: (*a*) the space of harmonic polynomials $u(x, y, z)$ of degree ≤ 2; (*b*) the space of homogeneous cubic harmonic polynomials $u(x, y, z)$.

12.1.2. *True or false*: (a) Every harmonic polynomial is homogeneous.
(*b*) Every homogeneous polynomial is harmonic.

12.1.3. Solve the Poisson boundary value problem $-\Delta u = 1$ on the unit ball $x^2 + y^2 + z^2 < 1$ with homogeneous Dirichlet boundary conditions. *Hint*: Look for a polynomial solution.

♢ 12.1.4. Prove that if $u(x, y, z)$ solves the Laplace equation, then so does the translated function $U(x, y, z) = u(x - a, y - b, z - c)$ for constants a, b, c.

♢ 12.1.5. (*a*) Prove that if $u(x, y, z)$ solves Laplace's equation, so does the rescaled function $U(x, y, z) = u(\lambda x, \lambda y, \lambda z)$ for any constant λ. (*b*) More generally, show that $U(x, y, z) = \mu\, u(\lambda x, \lambda y, \lambda z) + c$ solves Laplace's equation for any constants λ, μ, c.

♢ 12.1.6. Let A be a constant nonsingular 3×3 matrix, $u(\mathbf{x})$ a C^1 scalar field, and $\mathbf{v}(\mathbf{x})$ a C^1 vector field. Set $U(\mathbf{x}) = u(A\,\mathbf{x})$ and $\mathbf{V}(\mathbf{x}) = \mathbf{v}(A\,\mathbf{x})$. Prove that
(*a*) $\nabla U(\mathbf{x}) = A^T \nabla u(A\,\mathbf{x})$, (*b*) $\nabla \cdot \mathbf{V}(\mathbf{x}) = w(A\,\mathbf{x})$, where $w(\mathbf{x}) = \nabla \cdot (A\,\mathbf{v})(\mathbf{x})$.

♢ 12.1.7. Prove that every rotation and reflection is a symmetry of the Laplace equation. In other words, if Q is any 3×3 orthogonal matrix, so $Q^T Q = \mathrm{I}$, and $u(\mathbf{x})$ is a harmonic function, then so is $U(\mathbf{x}) = u(Q\,\mathbf{x})$. *Hint*: Use Exercise 12.1.6.

♢ 12.1.8. *The Weak Maximum Principle*: Let $\Omega \subset \mathbb{R}^2$ be a bounded domain. Let $u(x, y, z)$ solve the Poisson equation $-\Delta u = f(x, y, z)$, where $f(x, y, z) < 0$ for all $(x, y, z) \in \Omega$.
(*a*) Prove that the maximum value of u occurs on the boundary $\partial\Omega$.
Hint: Explain why u cannot have a local maximum at any interior point in Ω.
(*b*) Generalize your result to the case $f(x, y, z) \leq 0$.
Hint: Look at $v_\varepsilon(x, y, z) = u(x, y, z) + \varepsilon\,(x^2 + y^2 + z^2)$ and let $\varepsilon \to 0^+$.

♢ 12.1.9. Find the equilibrium equations corresponding to minimizing $|||\, \nabla u \,|||^2$ subject to homogeneous Dirichlet boundary conditions, where the indicated norm is based on the weighted inner product

$$\langle\!\langle\, \mathbf{v}\,,\, \mathbf{w}\,\rangle\!\rangle = \iiint_\Omega \mathbf{v}(x, y, z) \cdot \mathbf{w}(x, y, z)\, \sigma(x, y, z)\, dx\, dy\, dz,$$

with $\sigma(x, y, z) > 0$ a positive scalar function.

♢ 12.1.10. Prove the following vector calculus identities:
(*a*) $\nabla \cdot (u\,\mathbf{v}) = \nabla u \cdot \mathbf{v} + u\,\nabla \cdot \mathbf{v}$, (*b*) $\nabla \times (u\,\mathbf{v}) = \nabla u \times \mathbf{v} + u\,\nabla \times \mathbf{v}$,
(*c*) $\nabla \cdot (\mathbf{v} \times \mathbf{w}) = (\nabla \times \mathbf{v}) \cdot \mathbf{w} - \mathbf{v} \cdot (\nabla \times \mathbf{w})$, (*d*) $\nabla \times (\nabla \times \mathbf{v}) = \nabla(\nabla \cdot \mathbf{v}) - \Delta\mathbf{v}$.
(In the final term, the Laplacian Δ acts component-wise on the vector field $\mathbf{v}$.)

♢ 12.1.11. Let Ω be a bounded domain with piecewise smooth boundary $\partial\Omega$. Prove the following identities: (*a*) $\displaystyle\iiint_\Omega \Delta u\, dx\, dy\, dz = \iint_{\partial\Omega} \frac{\partial u}{\partial \mathbf{n}}\, dS$,
(*b*) $\displaystyle\iiint_\Omega u\, \Delta u\, dx\, dy\, dz = \iint_{\partial\Omega} u\, \frac{\partial u}{\partial \mathbf{n}}\, dS - \iiint_\Omega |||\, \nabla u \,|||^2\, dx\, dy\, dz$.

12.1.12. Suppose the inhomogeneous Neumann boundary value problem (12.1, 4) has a solution. (*a*) Prove that $\displaystyle\iint_{\partial\Omega} k\, dS = 0$. (*b*) Is the solution unique? If not, what is the most general solution? (*c*) State and prove an analogous result for the inhomogeneous Poisson equation $-\Delta u = f(x, y, z)$. (*d*) Provide a physical explanation for your answers.

♢ 12.1.13. Find a minimization principle that characterizes the solution to the inhomogeneous mixed boundary value problem $-\Delta u = f$ on Ω, with $u = g$ on $D \subsetneq \partial\Omega$, and $\partial u/\partial \mathbf{n} = h$ on $N = \partial\Omega \setminus D$.

♡ 12.1.14. (*a*) Prove that, subject to suitable boundary conditions, the curl $\nabla\times$ defines a self-adjoint operator with respect to the L^2 inner product between vector fields. What kinds of boundary conditions do you need to impose for your integration by parts argument to be valid? *Hint*: Use the identity in Exercise 12.1.10(*c*). (*b*) What operator on vector fields is given by the self-adjoint composition $S = (\nabla\times)^* \circ (\nabla\times)$? (*c*) Choose a set of homogeneous boundary conditions that make S self-adjoint. Is the resulting boundary value problem $S[\mathbf{v}] = \mathbf{f}$ positive definite? If not, what does the Fredholm Alternative say about its solvability?

12.2 Separation of Variables for the Laplace Equation

In this section, we revisit the method of separation of variables in the context of the three-dimensional Laplace equation. As always, its applicability is unfortunately restricted to rather special, but important, geometric configurations, the simplest being rectangular, cylindrical, and spherical domains. Since the first two are straightforward extensions of their two-dimensional counterparts, we will discuss only spherically separable solutions in any detail.

The simplest domain to which the separation of variables method applies is a rectan-

gular box:

$$B = \{\, 0 < x < a,\ 0 < y < b,\ 0 < z < c \,\}.$$

For functions of three variables, one begins the separation process by splitting off one of them, by setting $u(x, y, z) = v(x)\, w(y, z)$, say. The function $v(x)$ satisfies a simple second-order ordinary differential equation, while $w(y, z)$ solves the two-dimensional Helmholtz equation (11.21), which is further separated by writing $w(y, z) = p(y)\, q(z)$. The resulting fully separated solutions $u(x, y, z) = v(x)\, p(y)\, q(z)$ are (mostly) products of trigonometric and hyperbolic functions. Implementation of the technique and analysis of the resulting series solutions are relegated to Exercise 12.2.34.

In the case that the domain is a cylinder, one passes to *cylindrical coordinates* r, θ, z, where

$$x = r\cos\theta, \qquad y = r\sin\theta, \qquad z = z, \tag{12.13}$$

to effect the separation. Writing $u(r, \theta, z) = v(r, \theta)\, w(z)$, one finds that $w(z)$ satisfies a simple second-order ordinary differential equation, while $v(r, \theta)$ solves the two-dimensional polar Helmholtz equation (11.51) on a disk. Applying a further separation to $v(r, \theta)$, as in Chapter 11, produces fully separable solutions $u(r, \theta, z) = p(r)\, q(\theta)\, w(z)$ as products of Bessel functions of the cylindrical radius r, trigonometric functions of the polar angle θ, and hyperbolic functions of z; see Exercise 12.2.40.

The most interesting case is that of spherical coordinates, which we proceed to analyze in detail in the following subsection.

Remark: These are just three of the many coordinate systems in which the three-dimensional Laplace equation separates. See [**78**, **79**] for 37 additional exotic types, including ellipsoidal, toroidal, and parabolic spheroidal coordinates. The resulting separable solutions are written in terms of new classes of special functions that solve interesting second-order ordinary differential equations, all of Sturm–Liouville form (9.71).

Laplace's Equation in a Ball

Suppose a solid ball (e.g., the Earth) is subject to a specified steady temperature distribution on its spherical boundary. Our task is to determine the equilibrium temperature within the ball. We assume that the body is composed of an isotropic, uniform medium and, to slightly simplify the analysis, choose units in which its radius equals 1.

To find the equilibrium temperature within the ball, we must solve the Dirichlet boundary value problem

$$\begin{aligned} \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} &= 0, & x^2 + y^2 + z^2 &< 1, \\ u(x, y, z) &= h(x, y, z), & x^2 + y^2 + z^2 &= 1, \end{aligned} \tag{12.14}$$

where h is prescribed on the bounding unit sphere. Problems in spherical geometries are most naturally analyzed in *spherical coordinates* r, φ, θ. Our convention is to set

$$x = r\sin\varphi\,\cos\theta, \qquad y = r\sin\varphi\,\sin\theta, \qquad z = r\cos\varphi, \tag{12.15}$$

where $-\pi < \theta \le \pi$ is the *azimuthal angle* or *longitude*, while $0 \le \varphi \le \pi$ is the *zenith angle* or *latitude* on the sphere of radius $r = \sqrt{x^2 + y^2 + z^2}$. In other words, φ measures

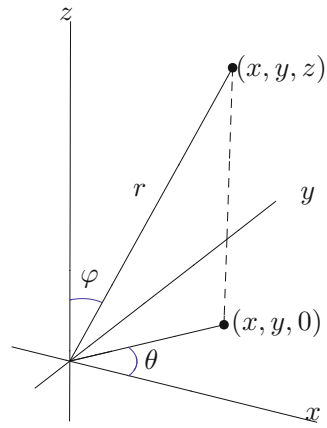


Figure 12.1. Spherical coordinates.

the angle between the vector $(x, y, z)^T$ and the positive z–axis, while θ measures the angle between its projection $(x, y, 0)^T$ on the (x, y)–plane and the positive x–axis; see Figure 12.1. On Earth, longitude θ is measured from the Greenwich prime meridian, while latitude is measured from the equator, and so equals $\frac{1}{2}\pi - \varphi$ (although the everyday units are degrees, not radians).

Warning: In many books, particularly those in physics, the roles of θ and φ are *reversed*, leading to much confusion when one is perusing the literature. We prefer the mathematical convention, since the azimuthal angle θ coincides with the cylindrical angle coordinate (and the polar coordinate on the (x, y)–plane), thus avoiding unnecessary confusion when going from one coordinate system to the other. You must be attentive to the convention being used when consulting any reference!

In spherical coordinates, the Laplace equation for $u(r, \varphi, \theta)$ takes the form

$$\Delta u = \frac{\partial^2 u}{\partial r^2} + \frac{2}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \varphi^2} + \frac{\cos\varphi}{r^2 \sin\varphi}\frac{\partial u}{\partial \varphi} + \frac{1}{r^2 \sin^2\varphi}\frac{\partial^2 u}{\partial \theta^2} = 0. \tag{12.16}$$

This important formula is the final result of a fairly nasty chain rule computation, whose details are left to the motivated reader. (Set aside lots of paper and keep an eraser handy!)

To construct separable solutions to the spherical coordinate form (12.16) of the Laplace equation, we begin by separating off the radial part of the solution, setting

$$u(r, \varphi, \theta) = v(r)\, w(\varphi, \theta). \tag{12.17}$$

Substituting this ansatz into (12.16), multiplying the resulting equation through by $\dfrac{r^2}{v\, w}$, and then placing all the terms involving r on one side yields

$$\frac{1}{v}\left(r^2 \frac{d^2 v}{dr^2} + 2r \frac{dv}{dr} \right) = -\frac{1}{w}\, \Delta_S[w], \tag{12.18}$$

where

$$\Delta_S[w] = \frac{\partial^2 w}{\partial \varphi^2} + \frac{\cos\varphi}{\sin\varphi}\frac{\partial w}{\partial \varphi} + \frac{1}{\sin^2\varphi}\frac{\partial^2 w}{\partial \theta^2}. \tag{12.19}$$

The second-order differential operator Δ_S, which involves only the angular components of the full Laplacian operator Δ, is of particular significance. It is known as the *spherical Laplacian*, and governs the equilibrium and dynamics of thin spherical shells — see Example 12.15 below.

Returning to equation (12.18), our usual separation argument applies. The left-hand side depends only on r, while the right-hand side depends only on the angles φ, θ. This can occur only when both sides are equal to a common separation constant, which we denote by μ. As a consequence, the radial component $v(r)$ satisfies the ordinary differential equation

$$r^2\, v'' + 2\, r\, v' - \mu\, v = 0, \tag{12.20}$$

which is of Euler type (11.89), and hence can be readily solved. However, let us put this equation aside for the time being, and concentrate our efforts on the more complicated angular components.

The second equation in (12.18) assumes the form

$$\Delta_S[w] + \mu\, w = \frac{\partial^2 w}{\partial \varphi^2} + \frac{\cos\varphi}{\sin\varphi}\,\frac{\partial w}{\partial \varphi} + \frac{1}{\sin^2\varphi}\,\frac{\partial^2 w}{\partial \theta^2} + \mu\, w = 0. \tag{12.21}$$

This second-order partial differential equation can be regarded as the eigenvalue equation for the spherical Laplacian operator Δ_S and is known as the *spherical Helmholtz equation*. To find explicit solutions, we adopt a further separation of angular variables,

$$w(\varphi, \theta) = p(\varphi)\, q(\theta), \tag{12.22}$$

which we substitute into (12.21). Dividing the result by the product $w = p\, q$, multiplying by $\sin^2\varphi$, and then rearranging terms, we are led to the separated system

$$\frac{1}{p}\left(\sin^2\varphi\,\frac{d^2 p}{d\varphi^2} + \cos\varphi\sin\varphi\,\frac{dp}{d\varphi}\right) + \mu\sin^2\varphi = -\,\frac{1}{q}\,\frac{d^2 q}{d\theta^2} = \nu,$$

where, by our usual argument, ν is another separation constant. The spherical Helmholtz equation thereby splits into a pair of ordinary differential equations

$$\sin^2\varphi\,\frac{d^2 p}{d\varphi^2} + \cos\varphi\,\sin\varphi\,\frac{dp}{d\varphi} + (\mu\sin^2\varphi - \nu)\, p = 0, \qquad \frac{d^2 q}{d\theta^2} + \nu\, q = 0.$$

The equation for $q(\theta)$ is easy to solve. As one circumnavigates the sphere, the azimuthal angle θ increases from $-\pi$ to π, so $q(\theta)$ must be a 2π–periodic function. Thus, $q(\theta)$ solves the well-studied periodic boundary value problem treated, for instance, in (4.109). Up to a constant multiple, nonzero periodic solutions occur only when the separation constant assumes one of the values $\nu = m^2$, where $m = 0, 1, 2, \dots$ is an integer, with

$$q(\theta) = \cos m\theta \qquad \text{or} \qquad \sin m\theta, \qquad\qquad m = 0, 1, 2, \dots. \tag{12.23}$$

Each positive $\nu = m^2 > 0$ admits two linearly independent 2π–periodic solutions, while when $\nu = 0$, only the constant solutions are periodic.

The Legendre Equation and Ferrers Functions

With this information, we endeavor to solve the zenith differential equation

$$\sin^2\varphi\,\frac{d^2 p}{d\varphi^2} + \cos\varphi\,\sin\varphi\,\frac{dp}{d\varphi} + (\mu\sin^2\varphi - m^2)\, p = 0. \tag{12.24}$$

This is not so easy, and constructing analytic formulas for its solutions requires some ingenuity. The motivation behind the following steps may not be so apparent; indeed, they are the culmination of a long, detailed study of this important differential equation by mathematicians over the last 200 years.

As an initial simplification, let us get rid of the trigonometric functions, by invoking the change of variables

$$t = \cos\varphi, \qquad \text{with} \qquad p(\varphi) = P(\cos\varphi) = P(t). \tag{12.25}$$

Since

$$0 \le \varphi \le \pi, \qquad \text{we have} \qquad 0 \le \sqrt{1-t^2} = \sin\varphi \le 1.$$

According to the chain rule,

$$\frac{dp}{d\varphi} = \frac{dP}{dt}\,\frac{dt}{d\varphi} = -\sin\varphi\,\frac{dP}{dt} = -\sqrt{1-t^2}\;\frac{dP}{dt}\,,$$
$$\frac{d^2p}{d\varphi^2} = -\sin\varphi\,\frac{d}{dt}\left(-\sqrt{1-t^2}\;\frac{dP}{dt}\right) = (1-t^2)\,\frac{d^2P}{dt^2} - t\;\frac{dP}{dt}\,.$$

Substituting these expressions into (12.24), we conclude that $P(t)$ must satisfy

$$(1-t^2)^2\;\frac{d^2P}{dt^2} - 2\,t\,(1-t^2)\,\frac{dP}{dt} + \big[\,\mu\,(1-t^2) - m^2\,\big]\,P = 0. \tag{12.26}$$

Unfortunately, the resulting differential equation is still not elementary, but at least its coefficients are polynomials. It is known as the *Legendre differential equation* of *order* m, having first been employed by Adrien–Marie Legendre to study the gravitational attraction of *ellipsoid*al bodies. In the cases of interest to us, the order parameter m is an integer, while the separation constant μ plays the role of an eigenvalue.

Power series solutions to the Legendre equation can be constructed by the standard techniques presented in Section 11.3. The most general solution is a new type of special function, called a *Legendre function*, [**86**]. However, it turns out that the solutions we are actually interested in can all be written in terms of elementary algebraic functions. First of all, since $t = \cos\varphi$, the solution only needs to be defined on the interval $-1 \le t \le 1$, the so-called *cut locus*. The endpoints of the cut locus, $t = 1$ and $t = -1$, correspond to the sphere's north pole, $\varphi = 0$, and south pole, $\varphi = \pi$, respectively. Both endpoints are singular points for the Legendre equation, since the coefficient $(1-t^2)^2$ of the leading-order derivative vanishes when $t = \pm 1$. In fact, both are regular singular points, as you are asked to show in Exercise 12.2.11. Since ultimately we need the separable solution (12.17) to be a well-defined function of x, y, z (even at points where the spherical coordinates degenerate, i.e., on the z–axis), we need $p(\varphi)$ to be well defined at $\varphi = 0$ and π, and this requires $P(t)$ to be bounded at the singular points:

$$|\,P(-1)\,| < \infty, \qquad |\,P(+1)\,| < \infty. \tag{12.27}$$

Let us begin our analysis with the Legendre equation of order $m = 0$

$$(1-t^2)\;\frac{d^2P}{dt^2} - 2\,t\;\frac{dP}{dt} + \mu\,P = 0. \tag{12.28}$$

In this case, the eigenfunctions, i.e., solutions to the Legendre boundary value problem (12.27–28), are the *Legendre polynomials*

$$P_n(t) = \frac{(-1)^n}{2^n\,n!}\;\frac{d^n}{dt^n}\;(1-t^2)^n. \tag{12.29}$$

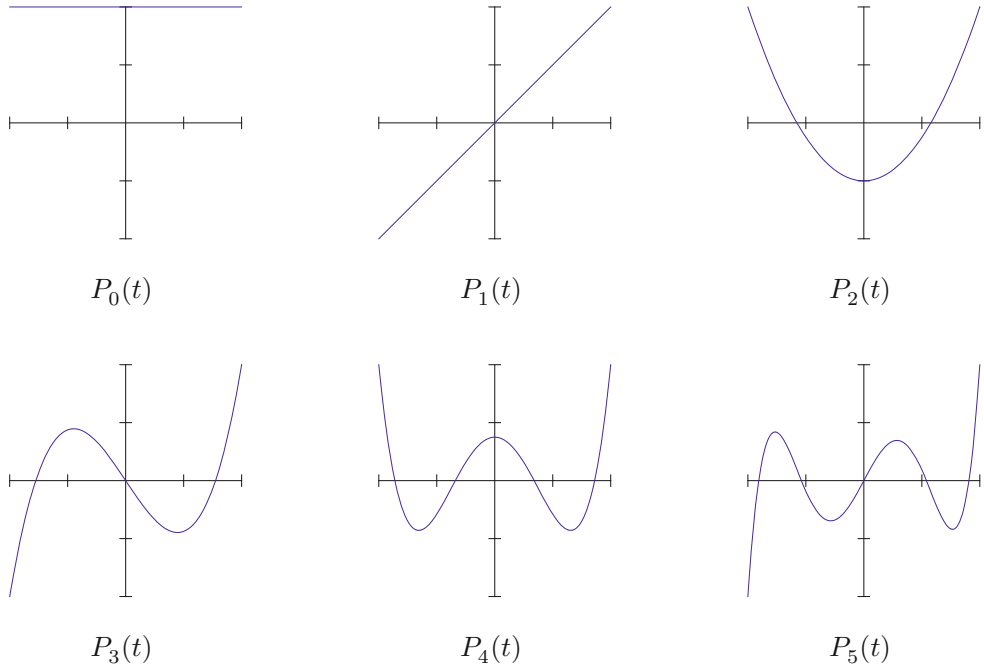


Figure 12.2. Legendre polynomials.

(The initial factor is by common convention, [**86**]; see (12.64) for the explicit formula.) The first few are

$$P_0(t) = 1, \qquad P_1(t) = t, \qquad P_2(t) = \tfrac{3}{2}t^2 - \tfrac{1}{2}, \qquad P_3(t) = \tfrac{5}{2}t^3 - \tfrac{3}{2}t,$$
$$P_4(t) = \tfrac{35}{8}t^4 - \tfrac{15}{4}t^2 + \tfrac{3}{8}, \qquad P_5(t) = \tfrac{63}{8}t^5 - \tfrac{35}{4}t^3 + \tfrac{15}{8}t,$$

and are graphed in Figure 12.2.

Each Legendre polynomial clearly satisfies the boundary conditions (12.27). To verify that they are indeed solutions to the differential equation (12.28), we set

$$Q_n(t) = (1 - t^2)^n.$$

By the chain rule, the derivative of $Q_n(t)$ is

$$Q_n' = -2nt(1 - t^2)^{n-1}, \qquad \text{and hence} \qquad (1 - t^2)Q_n' = -2nt(1 - t^2)^n = -2ntQ_n.$$

Differentiating the latter formula yields

$$(1 - t^2)Q_n'' - 2tQ_n' = -2ntQ_n' - 2nQ_n, \qquad \text{or} \qquad (1 - t^2)Q_n'' = -2(n-1)tQ_n' - 2nQ_n.$$

A simple induction proves that the k^{th} order derivative $Q_n^{(k)}(t) = \dfrac{d^k Q_n}{dt^k}$ satisfies

$$\begin{aligned}(1 - t^2)\,Q_n^{(k+2)} &= -2(n-k-1)\,tQ_n^{(k+1)} - 2[\,n + (n-1) + \cdots + (n-k)\,]\,Q_n^{(k)} \\ &= -2(n-k-1)\,tQ_n^{(k+1)} - (k+1)(2n-k)\,Q_n^{(k)}.\end{aligned} \tag{12.30}$$

In particular, when $k = n$, this reduces to

$$(1 - t^2)Q_n^{(n+2)} = 2\,t\,Q_n^{(n+1)} - n(n+1)\,Q_n^{(n)} = 0,$$

and so $\widehat{P}_n(t) = Q_n^{(n)}(t)$ satisfies

$$(1 - t^2)\widehat{P}_n'' - 2\,t\,\widehat{P}_n' + n(n+1)\,\widehat{P}_n = 0,$$

which is precisely the order 0 Legendre equation (12.28) with eigenvalue parameter $\mu = n(n+1)$. The Legendre polynomial P_n is a constant multiple of $\widehat{P}_n$, and hence it too satisfies the order 0 Legendre equation. According to Theorem 12.3 below, the Legendre polynomials form a complete system of eigenfunctions for the order 0 Legendre boundary value problem.

When the order $m > 0$, the eigenfunctions of the Legendre boundary value problem (12.26–27) are not always polynomials. They are known as the *Ferrers functions*, named after the nineteenth-century British mathematician Norman Ferrers, or, more generally, as *associated Legendre functions*. They have the explicit formula†

$$\begin{aligned} P_n^m(t) &= (1 - t^2)^{m/2}\,\frac{d^m}{dt^m}\,P_n(t) \\ &= (-1)^n\,\frac{(1 - t^2)^{m/2}}{2^n\,n!}\,\frac{d^{n+m}}{dt^{n+m}}\,(1 - t^2)^n, \end{aligned} \qquad n = m, m+1, \ldots, \tag{12.31}$$

which generalizes the formula (12.29) for the Legendre polynomials. In particular $P_n^0(t) = P_n(t)$. Here is a list of the first few Ferrers functions, which, for completeness, includes Legendre polynomials:

$$\begin{aligned} &P_0^0(t) = 1, && P_1^0(t) = t, && P_1^1(t) = \sqrt{1 - t^2}\,, \\ &P_2^0(t) = -\tfrac{1}{2} + \tfrac{3}{2}\,t^2, && P_2^1(t) = 3\,t\,\sqrt{1 - t^2}\,, && P_2^2(t) = 3\,(1 - t^2), \\ &P_3^0(t) = -\tfrac{3}{2}\,t + \tfrac{5}{2}\,t^3, && P_3^1(t) = \big(-\tfrac{3}{2} + \tfrac{15}{2}\,t^2\big)\,\sqrt{1 - t^2}, \\ &P_3^2(t) = 15\,t\,(1 - t^2), && P_3^3(t) = 15\,(1 - t^2)^{3/2}\,, \\ &P_4^0(t) = \tfrac{3}{8} - \tfrac{15}{4}\,t^2 + \tfrac{35}{8}\,t^4, && P_4^1(t) = \big(-\tfrac{15}{2}\,t + \tfrac{35}{2}\,t^3\big)\,\sqrt{1 - t^2}, \\ &P_4^2(t) = \big(-\tfrac{15}{2} + \tfrac{105}{2}\,t^2\big)(1 - t^2), && P_4^3(t) = 105\,t\,(1 - t^2)^{3/2}\,, && P_4^4(t) = 105\,(1 - t^2)^2. \end{aligned} \tag{12.32}$$

When $m = 2\,k \le n$ is an even integer, $P_n^m(t)$ is a polynomial function, while when $m = 2\,k + 1 \le n$ is odd, there is an extra factor of $\sqrt{1 - t^2}$. Keep in mind that the square root is real and positive, since we are restricting our attention to the interval $-1 \le t \le 1$. If $m > n$, formula (12.31) reduces to the zero function and so is not included in the final tally.

Warning: Even though half of the Ferrers functions are polynomials, only those with $m = 0$, i.e., $P_n(t) = P_n^0(t)$, are called *Legendre polynomials*.

---

† *Warning*: Some authors include a $(-1)^m$ factor in the formula, resulting in the opposite sign when m is odd. Another source of confusion is that many tables define the associated Legendre functions using the alternative initial factor $(t^2 - 1)^{m/2}$. But this is unsuitable, since we are solely interested in values of t lying in the interval $-1 \le t \le 1$, and this convention would result in a complex-valued function when m is odd. Following [**86**], we use the term "Ferrers function" to refer to the restriction of the associated Legendre function to the cut locus $-1 \le t \le 1$.

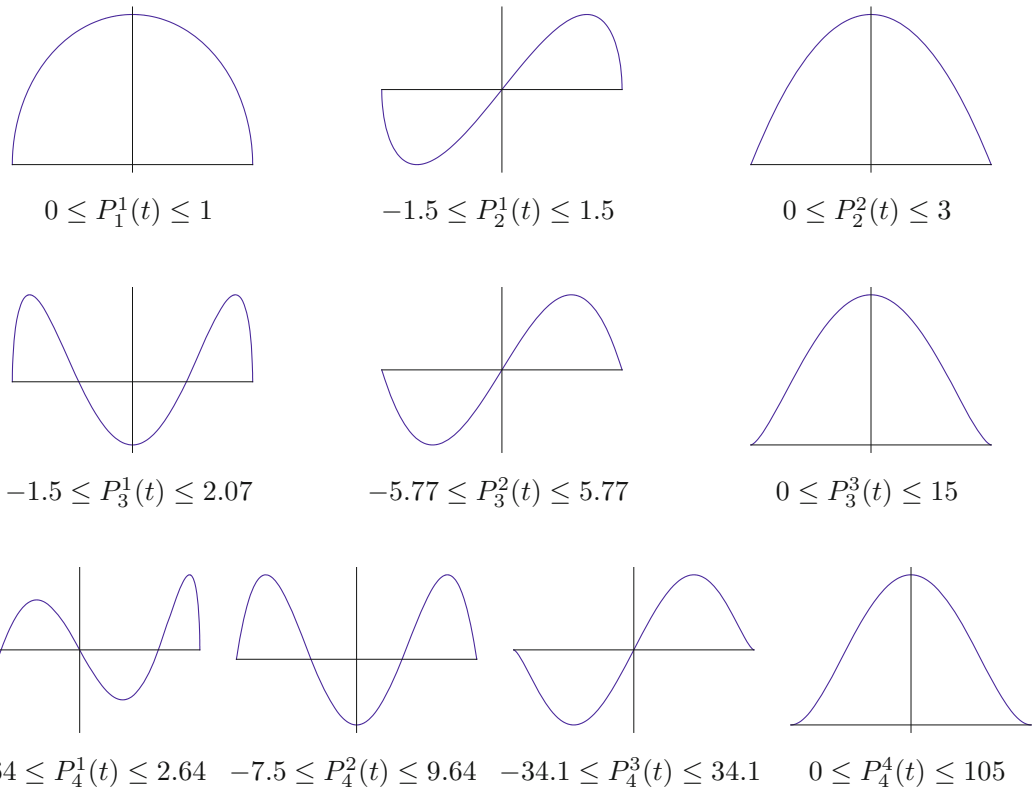


Figure 12.3. Ferrers functions.

Figure 12.3 displays graphs of the Ferrers functions $P_n^m(t)$ for $1 \le m \le n \le 4$. Pay particular attention to the fact that, owing to the choice of normalization factor, the graphs have very different vertical scales, as indicated by their minimum and maximum values (rounded to two decimal places) written below each — although one always has the freedom to rescale the eigenfunctions as desired, e.g., so as to be orthonormal.

To show that the Ferrers functions $P_n^m(t)$ satisfy the Legendre differential equation (12.26) of order m, we substitute $k = m + n$ in (12.30):

$$(1 - t^2)\,\frac{d^2 R_n^m}{dt^2} - 2\,(m+1)\,t\,\frac{dR_n^m}{dt} + (m + n + 1)\,(n - m)\,R_n^m = 0, \tag{12.33}$$

where

$$R_n^m(t) = Q_n^{(m+n)}(t).$$

This is *not* the order m Legendre equation, but it can be converted into it by setting

$$R_n^m(t) = (1 - t^2)^{-m/2}\,S_n^m(t).$$

Differentiating, we obtain

$$\frac{dR_n^m}{dt} = (1-t^2)^{-m/2}\,\frac{dS_n^m}{dt} - m\,t\,(1-t^2)^{-m/2-1}\,S_n^m,$$
$$\frac{d^2R_n^m}{dt^2} = (1-t^2)^{-m/2}\,\frac{d^2S_n^m}{dt^2} - 2\,m\,t\,(1-t^2)^{-m/2-1}\,\frac{dS_n^m}{dt} + \big[\,m + m(m+1)t^2\,\big]\,(1-t^2)^{-m/2-2}\,S_n^m.$$

Therefore, after a little algebra, equation (12.33) takes the alternative form

$$(1-t^2)^{-m/2+1}\,\frac{d^2S_n^m}{dt^2} - 2\,t\,(1-t^2)^{-m/2}\,\frac{dS_n^m}{dt} + \big[\,n\,(n+1)\,(1-t^2) - m^2\,\big]\,(1-t^2)^{-m/2-1}\,S_n^m = 0,$$

which, when multiplied by $(1-t^2)^{m/2+1}$, is precisely the order m Legendre equation (12.26) with eigenvalue parameter $\mu = n(n+1)$. Thus,

$$S_n^m(t) = (1-t^2)^{m/2}R_n^m(t) = (1-t^2)^{m/2}\,\frac{d^{n+m}}{dt^{n+m}}\,(1-t^2)^n,$$

which is a constant multiple of the Ferrers function $P_n^m(t)$, is a solution to the order m Legendre equation. Moreover, we note that

$$P_n^m(1) = P_n^m(-1) = 0, \qquad \text{when} \qquad m > 0, \tag{12.34}$$

and we conclude that $P_n^m(t)$ is an eigenfunction for the order m Legendre boundary value problem.

The following result states that the Ferrers functions provide a complete list of solutions to the Legendre boundary value problem (12.26–27).

Theorem 12.3. *Let $m \geq 0$ be a nonnegative integer. Then the order m Legendre boundary value problem prescribed by* (12.26–27) *has eigenvalues $\mu_n = n(n+1)$ for $n = 0, 1, 2, \ldots$, and associated eigenfunctions $P_n^m(t)$, where $m = 0, \ldots, n$. Moreover, the Ferrers eigenfunctions form a complete orthogonal system relative to the* L^2 *inner product on the cut locus $[-1, 1]$.*

Returning to the zenith variable φ via (12.25), Theorem 12.3 implies that our original boundary value problem

$$\sin^2\varphi\,\frac{d^2p}{d\varphi^2} + \cos\varphi\,\sin\varphi\,\frac{dp}{d\varphi} + (\mu\,\sin^2\varphi - m^2)\,p = 0, \qquad |\,p(0)\,|,\ \ |\,p(\pi)\,| < \infty, \tag{12.35}$$

has its eigenvalues and eigenfunctions expressed in terms of the Ferrers functions:

$$\mu_n = n(n+1), \qquad p_n^m(\varphi) = P_n^m(\cos\varphi), \qquad \text{for} \qquad 0 \leq m \leq n. \tag{12.36}$$

Since $P_n^m(t)$ is either a polynomial or a polynomial multiplied by a power of $\sqrt{1-t^2}$, the eigenfunction $p_n^m(\varphi)$ is a trigonometric polynomial of degree n, which we call a *trigono-*

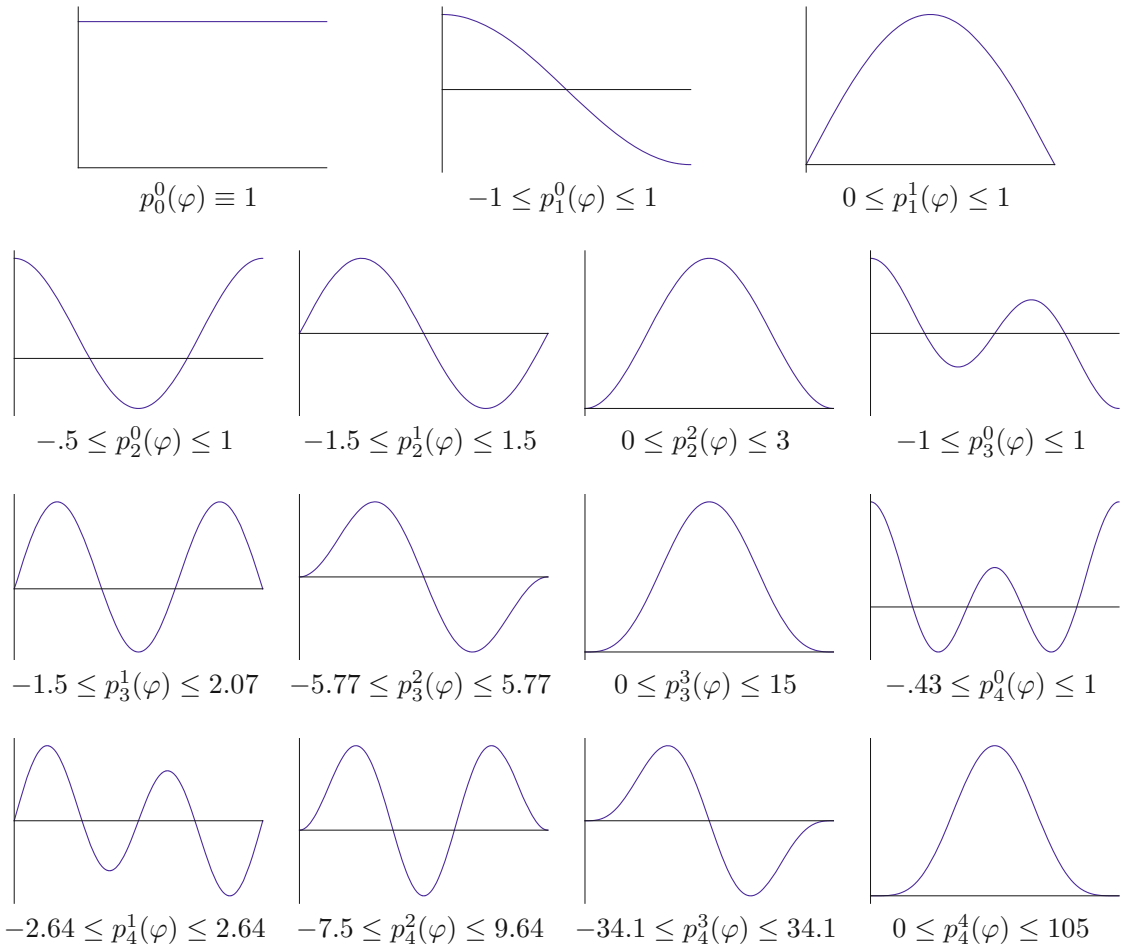


Figure 12.4. Trigonometric Ferrers functions.

metric Ferrers function. Here are the first few, written in Fourier form, as in (3.38):

$$
\begin{aligned}
&p_0^0(\varphi) = 1, && p_1^0(\varphi) = \cos\varphi, && p_1^1(\varphi) = \sin\varphi,\\
&p_2^0(\varphi) = \tfrac{1}{4} + \tfrac{3}{4}\cos 2\varphi, && p_2^1(\varphi) = \tfrac{3}{2}\sin 2\varphi, && p_2^2(\varphi) = \tfrac{3}{2} - \tfrac{3}{2}\cos 2\varphi,\\
&p_3^0(\varphi) = \tfrac{3}{8}\cos\varphi + \tfrac{5}{8}\cos 3\varphi, && && p_3^1(\varphi) = \tfrac{3}{8}\sin\varphi + \tfrac{15}{8}\sin 3\varphi,\\
&p_3^2(\varphi) = \tfrac{15}{4}\cos\varphi - \tfrac{15}{4}\cos 3\varphi, && && p_3^3(\varphi) = \tfrac{45}{4}\sin\varphi - \tfrac{15}{4}\sin 3\varphi,\\
&p_4^0(\varphi) = \tfrac{9}{64} + \tfrac{5}{16}\cos 2\varphi + \tfrac{35}{64}\cos 4\varphi, && && p_4^1(\varphi) = \tfrac{5}{8}\sin 2\varphi + \tfrac{35}{16}\sin 4\varphi,\\
&p_4^2(\varphi) = \tfrac{45}{16} + \tfrac{15}{4}\cos 2\varphi - \tfrac{105}{16}\cos 4\varphi, && && p_4^3(\varphi) = \tfrac{105}{4}\sin 2\varphi - \tfrac{105}{8}\sin 4\varphi,\\
&p_4^4(\varphi) = \tfrac{315}{8} - \tfrac{105}{2}\cos 2\varphi + \tfrac{105}{8}\cos 4\varphi. && &&
\end{aligned}
\tag{12.37}
$$

It is also instructive to plot the eigenfunctions in terms of the zenith angle φ; see Figure 12.4. As in Figure 12.3, the vertical scales are not the same, as indicated by the listed minimum and maximum values.

Spherical Harmonics

At this stage, we have determined both angular components of our separable solutions (12.22). Multiplying the two parts together results in the spherical angle functions

$$\begin{aligned} Y_n^m(\varphi,\theta) &= p_n^m(\varphi)\cos m\theta, & & n = 0,1,2,\dots, \\ \widetilde{Y}_n^m(\varphi,\theta) &= p_n^m(\varphi)\sin m\theta, & & m = 0,1,\dots,n, \end{aligned} \tag{12.38}$$

known as *spherical harmonics*. They satisfy the *spherical Helmholtz equation*

$$\Delta_S\, Y_n^m + n(n+1)\, Y_n^m \;=\; 0 \;=\; \Delta_S\, \widetilde{Y}_n^m + n(n+1)\, \widetilde{Y}_n^m, \tag{12.39}$$

and so are eigenfunctions for the spherical Laplacian operator, (12.19), with associated eigenvalues $\mu_n = n(n+1)$ for $n = 0,1,2,\dots$. The n^{th} eigenvalue μ_n admits a $(2n+1)$–dimensional eigenspace, spanned by the spherical harmonics

$$Y_n^0(\varphi,\theta),\quad Y_n^1(\varphi,\theta),\quad \dots\ ,\ Y_n^n(\varphi,\theta),\quad \widetilde{Y}_n^1(\varphi,\theta),\quad \dots\ ,\ \widetilde{Y}_n^n(\varphi,\theta).$$

(The omitted function $\widetilde{Y}_n^0(\varphi,\theta) \equiv 0$ is trivial, and so does not contribute.) In Figure 12.5 we plot the first few spherical harmonic surfaces $r = Y_n^m(\varphi,\theta)$. In these graphs, in view of the spherical coordinate formulae (12.15), points with a negative r coordinate appear on the opposite side of the origin from their positive r counterparts. Incidentally, the graphs of the other spherical harmonic surfaces $r = \widetilde{Y}_n^m(\varphi,\theta)$, when $m > 0$, are obtained by rotation around the z–axis by 90°; see Exercise 12.2.20. On the other hand, the graphs of Y_n^0 are cylindrically symmetric (why?), and hence unaffected by such a rotation.

Self-adjointness of the spherical Laplacian, as per Exercise 12.2.21, implies that the spherical harmonics are orthogonal with respect to the L^2 inner product

$$\langle\, f\,,g\,\rangle = \iint_{S_1} f\, g\, dS = \int_{-\pi}^{\pi}\int_0^{\pi} f(\varphi,\theta)\, g(\varphi,\theta)\sin\varphi\; d\varphi\, d\theta \tag{12.40}$$

given by integrating the product of the functions with respect to the surface area element $dS = \sin\varphi\; d\varphi\, d\theta$ on the unit sphere $S_1 = \{\, \|\,\mathbf{x}\,\| = 1\,\}$. More correctly, self-adjointness only guarantees orthogonality of the harmonics corresponding to distinct eigenvalues: $\mu_n \neq \mu_l$. However, the orthogonality relations

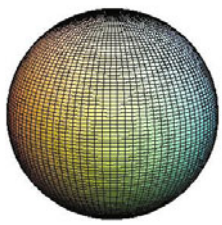
$$\begin{aligned} \langle\, Y_n^m\,, Y_l^k\,\rangle &= \iint_{S_1} Y_n^m\, Y_l^k\, dS = 0, & &\text{for} & &(m,n) \neq (k,l), \\ \langle\, Y_n^m\,, \widetilde{Y}_l^k\,\rangle &= \iint_{S_1} Y_n^m\, \widetilde{Y}_l^k\, dS = 0, & &\text{for all} & &(m,n),\ (k,l), \\ \langle\, \widetilde{Y}_n^m\,, \widetilde{Y}_l^k\,\rangle &= \iint_{S_1} \widetilde{Y}_n^m\, \widetilde{Y}_l^k\, dS = 0, & &\text{for} & &(m,n) \neq (k,l), \end{aligned} \tag{12.41}$$

do, in fact, hold in full generality; Exercise 12.2.22 asks you to supply the details. Moreover, their norms can be explicitly computed:

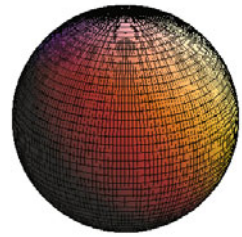
$$\|\, Y_n^0\,\|^2 = \frac{4\,\pi}{2\,n+1}\,, \qquad \|\, Y_n^m\,\|^2 = \|\, \widetilde{Y}_n^m\,\|^2 = \frac{2\,\pi(n+m)\,!}{(2\,n+1)(n-m)\,!}\,, \qquad m = 1,\dots,n. \tag{12.42}$$

Proofs of the latter formulae are outlined in Exercise 12.2.24.

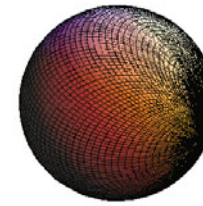
With some further work, it can be shown that the spherical harmonics form a complete orthogonal system of functions on the unit sphere. This means that any reasonable (e.g.,



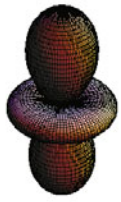
$Y_0^0(\varphi,\theta)$



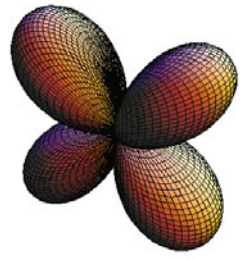
$Y_1^0(\varphi,\theta)$



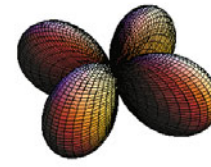
$Y_1^1(\varphi,\theta)$



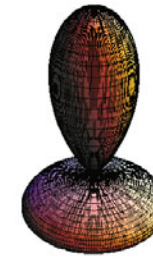
$Y_2^0(\varphi,\theta)$



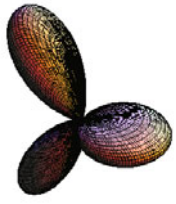
$Y_2^1(\varphi,\theta)$



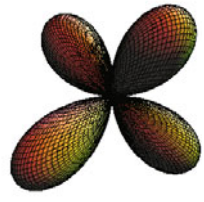
$Y_2^2(\varphi,\theta)$



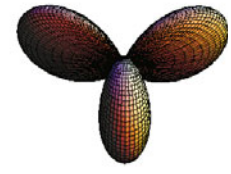
$Y_3^0(\varphi,\theta)$



$Y_3^1(\varphi,\theta)$

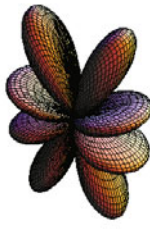


$Y_3^2(\varphi,\theta)$

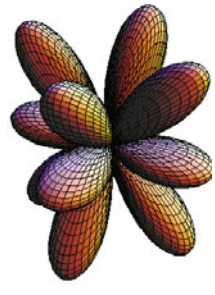


$Y_3^3(\varphi,\theta)$

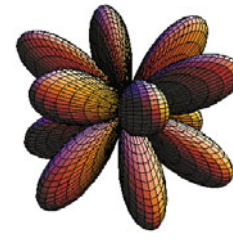
$Y_4^0(\varphi,\theta)$



$Y_4^1(\varphi,\theta)$



$Y_4^2(\varphi,\theta)$



$Y_4^3(\varphi,\theta)$

$Y_4^4(\varphi,\theta)$

Figure 12.5. Spherical harmonics.

piecewise C^1 or even L^2) function $h\colon S_1 \to \mathbb{R}$, can be expanded into a convergent *spherical harmonic series*

$$h(\varphi,\theta) = \frac{c_{0,0}}{2} + \sum_{n=1}^{\infty} \left(\frac{c_{0,n}}{2}\, Y_n^0(\varphi) + \sum_{m=1}^{n} \Big[c_{m,n} Y_n^m(\varphi,\theta) + \widetilde{c}_{m,n} \widetilde{Y}_n^m(\varphi,\theta) \Big] \right). \tag{12.43}$$

Applying the orthogonality relations (12.41), we find that the spherical harmonic coefficients are given by the inner products

$$c_{0,n} = \frac{2\,\langle\, h\,, Y_n^0 \,\rangle}{\| Y_n^0 \|^2}, \qquad c_{m,n} = \frac{\langle\, h\,, Y_n^m \,\rangle}{\| Y_n^m \|^2}, \qquad \widetilde{c}_{m,n} = \frac{\langle\, h\,, \widetilde{Y}_n^m \,\rangle}{\| \widetilde{Y}_n^m \|^2}, \qquad \begin{array}{l} 0 \le n, \\ 1 \le m \le n, \end{array}$$

or, explicitly, using (12.40) and the formulae (12.42) for the norms,

$$\begin{aligned} c_{m,n} &= \frac{(2n+1)(n-m)!}{2\pi\,(n+m)!} \int_{-\pi}^{\pi}\int_{0}^{\pi} h(\varphi,\theta)\, p_n^m(\varphi)\, \cos m\theta\, \sin\varphi\; d\varphi\, d\theta, \\ \widetilde{c}_{m,n} &= \frac{(2n+1)(n-m)!}{2\pi\,(n+m)!} \int_{-\pi}^{\pi}\int_{0}^{\pi} h(\varphi,\theta)\, p_n^m(\varphi)\, \sin m\theta\, \sin\varphi\; d\varphi\, d\theta. \end{aligned} \tag{12.44}$$

As with an ordinary Fourier series, the extra $\frac{1}{2}$ was appended to the $c_{0,n}$ terms in (12.43) so that equations (12.44) remain valid for all values of m, n. In particular, the constant term in the spherical harmonic series is the mean of the function h over the unit sphere:

$$\frac{c_{0,0}}{2} = \frac{1}{4\pi} \iint_{S_1} h\, dS = \frac{1}{4\pi} \int_{-\pi}^{\pi}\int_{0}^{\pi} h(\varphi,\theta)\, \sin\varphi\, d\varphi\, d\theta. \tag{12.45}$$

Remark: Establishing uniform convergence of a spherical harmonic series (12.43) is more challenging than in the Fourier series case, because, unlike the trigonometric functions, the orthonormal spherical harmonics are not uniformly bounded. A recent survey of what is known in this regard can be found in [**10**].

Remark: An alternative approach is to replace the real trigonometric functions by complex exponentials, and work with the *complex spherical harmonics*[†]

$$\mathcal{Y}_n^m(\varphi,\theta) = Y_n^m(\varphi,\theta) + \mathrm{i}\,\widetilde{Y}_n^m(\varphi,\theta) = p_n^m(\varphi)\, e^{\,\mathrm{i}\,m\theta}, \qquad \begin{array}{l} n = 0, 1, 2, \dots, \\ m = -n, -n+1, \dots, n. \end{array} \tag{12.46}$$

The associated orthogonality and expansion formulas are relegated to the exercises.

Harmonic Polynomials

To complete our solution to the Laplace equation on the solid ball, we still need to solve the ordinary differential equation (12.20) for the radial component $v(r)$. In view of our analysis of the spherical Helmholtz equation, the original separation constant is $\mu = n(n+1)$ for some nonnegative integer $n \ge 0$, and so the radial equation takes the form

$$r^2 v'' + 2r v' - n(n+1)\, v = 0. \tag{12.47}$$

[†] Here we use the convention that $Y_n^m = Y_n^{-m}$, $\widetilde{Y}_n^m = -\widetilde{Y}_n^{-m}$, and $\widetilde{Y}_n^0 \equiv 0$, which is compatible with their defining formulas (12.38).

To solve this Euler equation, we substitute the power ansatz $v(r) = r^\alpha$, and find that the exponent α must satisfy the quadratic indicial equation

$$\alpha^2 + \alpha - n(n+1) = 0, \qquad \text{and hence} \qquad \alpha = n \qquad \text{or} \qquad \alpha = -(n+1).$$

Therefore, the two linearly independent solutions are

$$v_1(r) = r^n \qquad \text{and} \qquad v_2(r) = r^{-n-1}. \tag{12.48}$$

Since we are currently interested only in solutions that remain bounded at $r = 0$ — the center of the ball — we will retain just the first solution $v(r) = r^n$ for our subsequent analysis.

At this stage, we have solved all three ordinary differential equations for the separable solutions. We combine (12.23, 38, 48) to produce the following spherically separable solutions to the Laplace equation:

$$\begin{aligned} H_n^m &= r^n\, Y_n^m(\varphi,\theta) = r^n\, p_n^m(\varphi)\cos m\theta, & & n = 0,1,2,\dots, \\ \widetilde{H}_n^m &= r^n\, \widetilde{Y}_n^m(\varphi,\theta) = r^n\, p_n^m(\varphi)\sin m\theta, & & m = 0,1,\dots,n. \end{aligned} \tag{12.49}$$

Although apparently complicated, these solutions are, perhaps surprisingly, elementary polynomial functions of the rectangular coordinates x, y, z, and hence are *harmonic polynomials*. The first few are

$$\begin{aligned} H_0^0 &= 1, & H_1^0 &= z, & H_2^0 &= z^2 - \tfrac{1}{2}x^2 - \tfrac{1}{2}y^2, & H_3^0 &= z^3 - \tfrac{3}{2}x^2 z - \tfrac{3}{2}y^2 z, \\ & & H_1^1 &= x, & H_2^1 &= 3\,x\,z, & H_3^1 &= 6\,x\,z^2 - \tfrac{3}{2}x^3 - \tfrac{3}{2}x\,y^2, \\ & & \widetilde{H}_1^1 &= y, & \widetilde{H}_2^1 &= 3\,y\,z, & \widetilde{H}_3^1 &= 6\,y\,z^2 - \tfrac{3}{2}x^2 y - \tfrac{3}{2}y^3, \\ & & & & H_2^2 &= 3\,x^2 - 3\,y^2, & H_3^2 &= 15\,x^2 z - 15\,y^2 z, \\ & & & & \widetilde{H}_2^2 &= 6\,x\,y, & \widetilde{H}_3^2 &= 30\,x\,y\,z, \\ & & & & & & H_3^3 &= 15\,x^3 - 45\,x\,y^2, \\ & & & & & & \widetilde{H}_3^3 &= 45\,x^2 y - 15\,y^3. \end{aligned} \tag{12.50}$$

The polynomials

$$H_n^0, \quad H_n^1, \ \dots\ , \ H_n^n, \quad \widetilde{H}_n^1, \ \dots\ , \ \widetilde{H}_n^n$$

are homogeneous of degree n. Orthogonality of the spherical harmonics implies that they form a basis for the vector space comprised of all homogeneous harmonic polynomials of degree n, which hence has dimension $2n + 1$.

The harmonic polynomials (12.49) form a complete system, and therefore the general solution to the Laplace equation inside the unit ball can be written as a harmonic polynomial series:

$$u(x,y,z) = \frac{c_{0,0}}{2} + \sum_{n=1}^{\infty} \left(\frac{c_{0,n}}{2}\, H_n^0(x,y,z) + \sum_{m=1}^{n} \big[\, c_{m,n} H_n^m(x,y,z) + \widetilde{c}_{m,n} \widetilde{H}_n^m(x,y,z) \,\big] \right), \tag{12.51}$$

or equivalently, in spherical coordinates,

$$u(r,\varphi,\theta) = \frac{c_{0,0}}{2} + \sum_{n=1}^{\infty} \left(\frac{c_{0,n}}{2}\, r^n\, Y_n^0(\varphi) + \sum_{m=1}^{n} \big[\, c_{m,n}\, r^n\, Y_n^m(\varphi,\theta) + \widetilde{c}_{m,n}\, r^n\, \widetilde{Y}_n^m(\varphi,\theta) \,\big] \right). \tag{12.52}$$

The coefficients $c_{m,n}, \widetilde{c}_{m,n}$ are uniquely prescribed by the boundary conditions. Indeed, substituting (12.52) into the Dirichlet boundary conditions on the unit sphere $r = 1$ yields

$$u(1,\varphi,\theta) = \frac{c_{0,0}}{2} + \sum_{n=1}^{\infty} \left(\frac{c_{0,n}}{2}\, Y_n^0(\varphi) + \sum_{m=1}^{n} \big[\, c_{m,n} Y_n^m(\varphi,\theta) + \widetilde{c}_{m,n} \widetilde{Y}_n^m(\varphi,\theta) \,\big] \right) = h(\varphi,\theta). \tag{12.53}$$

Thus, the coefficients $c_{m,n}, \widetilde{c}_{m,n}$ are given by the inner product formulae (12.44). If the terms in the resulting series are uniformly bounded — which occurs for all piecewise continuous functions h, as well as all L^2 functions and many generalized functions such as the delta function — then the harmonic polynomial series (12.52) converges everywhere, and, in fact, uniformly on any smaller ball $\| \mathbf{x} \| = r \le r_0 < 1$.

Averaging, the Maximum Principle, and Analyticity

In rectangular coordinates, the n^{th} summand of the series (12.51) is a homogeneous polynomial of degree n. Therefore, repeating the argument used in the two-dimensional situation (4.115), we conclude that the harmonic polynomial series is, in fact, a power series, and hence provides the *Taylor expansion for the harmonic function $u(x,y,z)$ at the origin*! In particular, its convergence for all $r < 1$ implies that the harmonic function $u(x,y,z)$ is analytic at $x = y = z = 0$.

The constant term in such a Taylor series can be identified with the value of the function at the origin: $u(0,0,0) = \frac{1}{2}\, c_{0,0}$. On the other hand, since $u = h$ on $S_1 = \partial\Omega$, the coefficient formula (12.45) tells us that

$$u(0,0,0) = \frac{c_{0,0}}{2} = \frac{1}{4\,\pi} \iint_{S_1} u \, dS. \tag{12.54}$$

Therefore, we have established the three-dimensional counterpart of Theorem 4.8: the value of a harmonic function u at the center of the sphere is equal to the average of its values on the sphere's surface. Moreover, each partial derivative $\dfrac{\partial^{i+j+k} u}{\partial x^i \partial y^j \partial z^k}(0,0,0)$ appears, up to a factor, as the coefficient of the terms $x^i y^j z^k$ in the Taylor series, and hence can be expressed as a certain linear combination of the coefficients $c_{m,n}, \widetilde{c}_{m,n}$, which are in turn given by the integral formulae (12.44).

So far, we have restricted our attention to a ball of unit radius. A simple scaling argument serves to establish the general result.

Theorem 12.4. *If $u(\mathbf{x})$ is a harmonic function defined on a domain $\Omega \subset \mathbb{R}^3$, then u is analytic inside Ω. Moreover, its value at any $\mathbf{x}_0 \in \Omega$ is obtained by averaging its values on any sphere centered at $\mathbf{x}_0$:*

$$u(\mathbf{x}_0) \;=\; \frac{1}{4\,\pi a^2} \iint_{\| \mathbf{x} - \mathbf{x}_0 \| = a} u \; dS, \tag{12.55}$$

provided the enclosed ball lies within its domain of analyticity: $\{ \| \mathbf{x} - \mathbf{x}_0 \| \le a \} \subset \Omega$.

Proof: It is easily checked that, under the hypothesis of the theorem, the rescaled and translated function

$$U(\mathbf{y}) = u(a\,\mathbf{y} + \mathbf{x}_0) = u(\mathbf{x}), \qquad \text{where} \qquad \mathbf{y} = \frac{\mathbf{x} - \mathbf{x}_0}{a}, \tag{12.56}$$

is harmonic on the unit ball $\| \mathbf{y} \| \leq 1$, and hence solves the boundary value problem (12.14) with boundary values $h(\mathbf{y}) = U(\mathbf{y}) = u(a\,\mathbf{y} + \mathbf{x}_0)$ on $\| \mathbf{y} \| = 1$. By the preceding remarks, $U(\mathbf{y})$ is analytic at $\mathbf{y} = \mathbf{0}$, and so $u(\mathbf{x}) = U\left(\dfrac{\mathbf{x} - \mathbf{x}_0}{a}\right)$ is analytic at $\mathbf{x} = \mathbf{x}_0$. Since $\mathbf{x}_0$ can be any point inside Ω, this establishes the analyticity of u everywhere in Ω. Moreover, according to (12.54),

$$u(\mathbf{x}_0) = U(\mathbf{0}) = \frac{1}{4\,\pi} \iint_{\| \mathbf{y} \| = 1} U \, dS = \frac{1}{4\,\pi\, a^2} \iint_{\| \mathbf{x} - \mathbf{x}_0 \| = a} u \, dS,$$

since the effect of the change of variables (12.56) is just to rescale the spherical surface integral. *Q.E.D.*

Arguing as in the planar case of Theorem 4.9, we readily establish the corresponding *Strong Maximum Principle* for harmonic functions of three variables.

Theorem 12.5. *A nonconstant harmonic function cannot have a local maximum or minimum at any interior point of its domain of definition. Moreover, its global maximum or minimum (if any) is located on the boundary of the domain.*

For instance, the Maximum Principle implies that the maximum and minimum temperatures in a solid body in thermal equilibrium are to be found only on its boundary. In physical terms, since heat energy must flow away from an internal maximum and towards an internal minimum, any local temperature extremum inside the body would preclude it from being in thermal equilibrium.

Example 12.6. In this example, we shall determine the electrostatic potential inside a hollow sphere when the upper and lower hemispheres are held at different constant potentials. This device is called a *spherical capacitor* and is realized experimentally by separating the two charged conducting hemispherical shells by a thin insulating ring at the equator. A straightforward scaling argument allows us to choose our units so that the sphere has unit radius, while the potential is set equal to 1 on the upper hemisphere and equal to 0, i.e., grounded, on the lower hemisphere. The resulting electrostatic potential satisfies the Laplace equation

$$\Delta u = 0 \qquad \text{inside a solid ball} \qquad \| \mathbf{x} \| < 1,$$

and is subject to Dirichlet boundary conditions

$$u(x, y, z) = h(x, y, z) \equiv \begin{cases} 1, & z > 0, \\ 0, & z < 0, \end{cases} \qquad \text{on the unit sphere} \qquad \| \mathbf{x} \| = 1. \tag{12.57}$$

The solution will be prescribed by a harmonic polynomial series (12.51) whose coefficients are fixed by the boundary values (12.57). Before tackling the required computation, let us first note that since the boundary data does not depend upon the azimuthal angle θ, the solution $u = u(r, \varphi)$ will also be independent of θ. Therefore, we need only consider the θ-independent spherical harmonic polynomials (12.38), which are those with $m = 0$. Thus,

$$u(x, y, z) \;=\; \frac{1}{2} \sum_{n=0}^{\infty} c_n \, H_n^0(x, y, z) \;=\; \frac{1}{2} \sum_{n=0}^{\infty} c_n \, r^n \, P_n(\cos \varphi), \tag{12.58}$$

where we abbreviate $c_n = c_{0,n}$. The boundary conditions (12.57) require

$$u|_{r=1} = \frac{1}{2}\sum_{n=0}^{\infty} c_n P_n(\cos\varphi) = h(\varphi) = \begin{cases} 1, & 0 \le \varphi < \frac{1}{2}\pi, \\ 0, & \frac{1}{2}\pi < \varphi \le \pi. \end{cases}$$

The coefficients are given by (12.44), which, in the case $m = 0$, reduce to

$$c_n = \frac{2n+1}{2\pi}\iint_{S_1} h\, Y_n^0\, dS = (2n+1)\int_0^{\pi/2} P_n(\cos\varphi)\sin\varphi\; d\varphi = (2n+1)\int_0^1 P_n(t)\, dt, \tag{12.59}$$

since $h = 0$ when $\frac{1}{2}\pi < \varphi \le \pi$. The first few are

$$c_0 = 1, \qquad c_1 = \tfrac{3}{2}, \qquad c_2 = 0, \qquad c_3 = -\tfrac{7}{8}, \qquad c_4 = 0, \qquad \dots .$$

Therefore, the solution has the explicit Taylor expansion

$$\begin{aligned} u(x,y,z) &= \tfrac{1}{2} + \tfrac{3}{4}\, r\cos\varphi - \tfrac{21}{128}\, r^3\cos\varphi - \tfrac{35}{128}\, r^3\cos 3\varphi + \cdots \\ &= \tfrac{1}{2} + \tfrac{3}{4}\, z + \tfrac{21}{32}\,(x^2+y^2)\, z - \tfrac{7}{16}\, z^3 + \cdots . \end{aligned} \tag{12.60}$$

Note in particular that the value $u(0,0,0) = \frac{1}{2}$ at the center of the sphere is the average of its boundary values, in accordance with Theorem 12.4. The solution depends only on the cylindrical coordinates r, z, which is a consequence of the invariance of the Laplace equation under general rotations, coupled with the invariance of the boundary data under rotations around the z–axis.

Remark: The same solution $u(x,y,z)$ describes the thermal equilibrium in a solid sphere whose upper hemisphere is held at temperature 1° and lower hemisphere at 0°.

Example 12.7. A closely related problem is to determine the electrostatic potential *outside* a spherical capacitor. As in the preceding example, we take our capacitor of radius 1, with electrostatic charge of 1 on the upper hemisphere and 0 on the lower hemisphere. Here, we need to solve the Laplace equation $\Delta u = 0$ in the unbounded domain $\Omega = \{\, \|\,\mathbf{x}\,\| > 1 \,\}$ — the exterior of the unit sphere — subject to the same Dirichlet boundary conditions (12.57). We anticipate that the potential will be vanishingly small at large distances away from the capacitor: $r = \|\,\mathbf{x}\,\| \gg 1$. Therefore, the harmonic polynomial solutions (12.49) will not help us solve this problem, since (except for the constant case) they become unboundedly large far away from the origin.

However, revisiting our original separation of variables argument will produce a different class of solutions having the desired decay properties. When we solved the radial equation (12.47), we discarded the solution $v_2(r) = r^{-n-1}$ because it had a singularity at the origin. In the present situation, the behavior of the function at $r = 0$ is irrelevant; our requirement is that the solution decay as $r \to \infty$, and $v_2(r)$ has this property. Therefore, we will utilize the *complementary harmonic functions*

$$\begin{aligned} K_n^m(x,y,z) &= r^{-2n-1}\, H_n^m(x,y,z) = r^{-n-1}\, Y_n^m(\varphi,\theta) = r^{-n-1} p_n^m(\varphi)\cos m\theta, \\ \widetilde{K}_n^m(x,y,z) &= r^{-2n-1}\, \widetilde{H}_n^m(x,y,z) = r^{-n-1}\, \widetilde{Y}_n^m(\varphi,\theta) = r^{-n-1} p_n^m(\varphi)\sin m\theta, \end{aligned} \tag{12.61}$$

for solving such exterior problems. For the capacitor problem, we need only those that are independent of θ, whereby $m = 0$. We write the resulting solution as a series

$$u(x,y,z) = \frac{1}{2}\sum_{n=0}^{\infty} c_n K_n^0(x,y,z) = \frac{1}{2}\sum_{n=0}^{\infty} c_n\, r^{-n-1}\, P_n(\cos\varphi). \tag{12.62}$$

The boundary conditions

$$u|_{r=1} = \frac{1}{2}\sum_{n=0}^{\infty} c_n P_n(\cos\varphi) = h(\varphi) \equiv \begin{cases} 1, & 0 \le \varphi < \frac{1}{2}\pi, \\ 0, & \frac{1}{2}\pi < \varphi \le \pi, \end{cases}$$

are identical to those in the previous example. Therefore, the coefficients are given by (12.59), leading to the series expansion

$$\begin{aligned} u(x,y,z) &= \frac{1}{2\,r} + \frac{3\cos\varphi}{4\,r^2} - \frac{21\cos\varphi + 35\cos 3\varphi}{128\,r^4} + \cdots \\ &= \frac{1}{2\sqrt{x^2+y^2+z^2}} + \frac{3\,z}{4\,(x^2+y^2+z^2)^{3/2}} + \frac{21\,(x^2+y^2)\,z - 14\,z^3}{32\,(x^2+y^2+z^2)^{7/2}} + \cdots . \end{aligned} \tag{12.63}$$

Observe that the higher-order terms become negligible at large distances, and hence the potential is asymptotic to that associated with a point charge concentrated at the origin of magnitude $\frac{1}{2}$, which is the average of the boundary potential over the sphere. This is indicative of a general fact, to be explored in Exercise 12.2.32.

Exercises

12.2.1. A solid ball of radius R has its upper hemispherical surface held at temperature T_1 and its lower hemispherical surface held at temperature T_0. Find the resulting equilibrium temperature.

12.2.2. A solid ball has its top hemispherical surface insulated and its bottom hemispherical surface held at a fixed temperature of 10°. Find its equilibrium temperature.

12.2.3. Find the potential inside a spherical capacitor of radius R when the upper hemisphere is at potential α and the lower is at β.

12.2.4. Find the potential $u(x,y,z)$ inside a unit spherical capacitor that has the indicated boundary values on the unit sphere $x^2+y^2+z^2 = 1$: (*a*) x, (*b*) x^2+y^2, (*c*) x^3. *Hint*: The potential is a polynomial.

12.2.5. Each point on the spherical boundary of a solid ball of radius 1 has temperature equal to its zenith angle φ. (*a*) Find the value of the equilibrium temperature at the center of the ball. (*b*) Find the Taylor polynomial of degree 3, based at the origin, for the equilibrium temperature distribution.

12.2.6. Solve Exercise 12.2.5 when the boundary temperature equals (*a*) $\cos\varphi$, (*b*) $\cos\theta$, (*c*) θ.

12.2.7. A solid spherical container of radius 3 cm contains a hollow spherical cavity of radius 1 cm in its center. The inner cavity is filled with boiling water at 100°, while the entire container is immersed in an ice water bath at 0°. Assume that the container is in thermal equilibrium. *True or false*: The temperature at a point half-way between the container's inner and outer boundaries is 50°. If true, explain. If false, what is the temperature at such a point?

12.2.8. Find the electrostatic potential between two concentric spherical metal shells of respective radii 1 and 1.2, given that the inner shell is grounded, while the outer shell has potential equal to 1.

♢ 12.2.9. Use the chain rule to establish the formula (12.16) for the Laplacian in spherical coordinates.

♢ 12.2.10. (*a*) Prove that $t = \pm 1$ are both regular singular points for the order 0 Legendre differential equation (12.28). (*b*) Prove that the Legendre eigenvalue problem (12.27–28) is defined by a self-adjoint operator with respect to the L^2 inner product on the cut locus $[-1, 1]$. (*c*) Discuss the orthogonality of the Legendre polynomials.

♢ 12.2.11. Solve Exercise 12.2.10 for the Legendre eigenvalue problem (12.26–27) of order m along with the relevant Ferrers eigenfunctions.

♢ 12.2.12. Suppose $m > 0$. (*a*) Find the Green's function for the boundary value problem

$$(1-t^2)\,\frac{d^2P}{dt^2} - 2t\,\frac{dP}{dt} - \frac{m^2}{1-t^2}\,P = f(t), \qquad |\,P(-1)\,|,\ |\,P(1)\,| < \infty.$$

Hint: The homogeneous differential equation has solutions $\left(\dfrac{1+t}{1-t}\right)^{\frac{m}{2}}$ and $\left(\dfrac{1-t}{1+t}\right)^{\frac{m}{2}}$.
(*b*) Use part (*a*) to prove completeness of the Ferrers functions of order $m > 0$ on $[-1, 1]$.
(*c*) Explain why there is no Green's function in the order $m = 0$ case.
Remark: When $m = 0$, one can use the trick of Example 9.49 to prove completeness. Although the Green's function for the modified operator does not have an explicit elementary formula, one can prove that it has logarithmic singularities at the endpoints, and hence finite double L^2 norm. See [**120**; §43] for details.

12.2.13. What happens when $n < m$ in formula (12.31)?

♢ 12.2.14. Prove that the Legendre polynomial (12.29) has the explicit formula

$$P_n(t) = \sum_{0 \le 2m \le n} (-1)^m \,\frac{(2n-2m)!}{2^n\,(n-m)!\,m!\,(n-2m)!}\,t^{n-2m}. \tag{12.64}$$

♢ 12.2.15. Prove the following recurrence relation for the Ferrers functions:

$$P_n^{m+1}(t) = \sqrt{1-t^2}\;\frac{dP_n^m}{dt} + \frac{m\,t}{\sqrt{1-t^2}}\,P_n^m(t). \tag{12.65}$$

♡ 12.2.16. In this exercise, we determine the L^2 norms of the Ferrers functions. (*a*) First, prove that $\displaystyle\int_{-1}^{1} (1-t^2)^n\,dt = \frac{2^{2n+1}\,(n!)^2}{(2n+1)!}$. *Hint*: Set $t = \cos\theta$ and then integrate by parts repeatedly. (*b*) Prove that $\| P_n \|^2 = \dfrac{2}{2n+1}$. *Hint*: Integrate by parts repeatedly and then use part (*a*). (*c*) Prove that $\| P_n^{m+1} \|^2 = (n-m)(n+m+1)\,\| P_n^m \|^2$. *Hint*: Use (12.65) and an integration by parts. (*d*) Finally, prove that $\| P_n^m \|^2 = \dfrac{2}{2n+1}\,\dfrac{(n+m)!}{(n-m)!}$.

12.2.17. (*a*) Prove that $P_n^m(t)$ is an even or odd function according to whether $m + n$ is an even or odd integer. (*b*) Prove that its Fourier form, $p_n^m(\varphi)$, depends only on $\cos n\,\varphi$, $\cos(n-2)\varphi$, $\cos(n-4)\varphi$, $\dots$ if m is even, and only on $\sin n\,\varphi$, $\sin(n-2)\varphi$, $\sin(n-4)\varphi$, $\dots$ if m is odd.

12.2.18. Let m be fixed. Are the functions $p_n^m(\varphi)$ for $n = 0, 1, 2, \dots$ mutually orthogonal with respect to the standard L^2 inner product on $[0, \pi]$? If not, is there an inner product that makes them orthogonal functions?

12.2.19. Prove that the surfaces defined by the first three spherical harmonics Y_0^0, Y_1^0, and Y_1^1, as in Figure 12.5, are all spheres. Find their centers and radii.

♢ 12.2.20. Explain why the surface defined by $r = \widetilde{Y}_n^m(\varphi, \theta)$ is obtained by rotating that defined by $r = Y_n^m(\varphi, \theta)$ around the z–axis by 90°.

♢ 12.2.21. Prove directly that the spherical Laplacian Δ_S is a self-adjoint linear operator with respect to the inner product (12.40).

♢ 12.2.22. (*a*) In view of Exercise 12.2.21, which orthogonality relations in (12.41) follow from their status as eigenfunctions of the spherical Laplacian?
(*b*) Prove the general orthogonality formulae by direct computation.

♢ 12.2.23. State and prove the orthogonality of the complex spherical harmonics (12.46). Then establish the following formula for their norms:

$$\| \mathcal{Y}_n^m \|^2 = \iint_{S_1} | \mathcal{Y}_n^m |^2 \, dS = \frac{4\pi(n+m)!}{(2n+1)(n-m)!} \qquad \begin{array}{l} n = 0, 1, 2, \dots, \\ m = -n, -n+1, \dots, n. \end{array} \tag{12.66}$$

♢ 12.2.24. Prove the formulae (12.42) for the norms of the spherical harmonics. *Hint*: Use Exercise 12.2.16.

♢ 12.2.25. Justify the formulas in (12.50) for (*a*) H_1^0, (*b*) H_2^0, (*c*) $\widetilde{H}_2^1$.

12.2.26. Find formulas for the following harmonic polynomials (*i*) in spherical coordinates; (*ii*) in rectangular coordinates: (*a*) H_4^0, (*b*) H_4^4, (*c*) $\widetilde{H}_4^4$.

12.2.27. Explain why every polynomial solution of the Laplace equation is a linear combination of the harmonic polynomials (12.49). *Hint*: Look at its Taylor series.

12.2.28. (*a*) Prove that if $u(x, y, z)$ is any harmonic polynomial, then so are $u(y, x, z)$, $u(z, x, y)$, and all other functions obtained by permuting the variables x, y, z. (*b*) Discuss the effect of such permutations on the basis harmonic polynomials $H_n^m(x, y, z)$ appearing in (12.50).

12.2.29. Find the formulas in rectangular coordinates for the following complementary harmonic functions: (*a*) K_0^0, (*b*) K_1^1, (*c*) K_2^0, (*d*) $\widetilde{K}_2^1$.

♢ 12.2.30. Let $u(x, y, z)$ be a harmonic function defined on the unit ball $r \leq 1$. Prove that its gradient at the center, $\nabla u(\mathbf{0})$, equals the average of the vector field $\mathbf{v}(\mathbf{x}) = \mathbf{x}\, u(\mathbf{x})$ over the unit sphere $r = 1$.

♢ 12.2.31. (*a*) Suppose $u(x, y, z)$ is a solution to the Laplace equation. Prove that the function $U(x, y, z) = r^{-1}\, u(x/r^2, y/r^2, z/r^2)$ obtained by *inversion* is also a solution. (*b*) Explain how inversion can be used to solve boundary value problems on the exterior of a sphere. (*c*) Use inversion to relate the solutions to Examples 12.6 and 12.7.

♢ 12.2.32. Suppose $u(r, \varphi, \theta)$ is the potential exterior to a spherical capacitor of unit radius. (*a*) Prove that $\lim_{r \to \infty} r\, u(r, \varphi, \theta)$ equals the average value of u on the sphere. (*b*) Use Exercise 12.2.31 to deduce this result as a consequence of Theorem 12.4.

12.2.33. (*a*) Write out, using spherical coordinates, formulas for the L^2 inner product and norm for scalar fields $f(r, \varphi, \theta)$ and $g(r, \varphi, \theta)$ on a solid ball of unit radius centered at the origin. (*b*) Let $f(x, y, z) = z$ and $g(x, y, z) = x^2 + y^2$. Find $\| f \|, \| g \|$ and $\langle f, g \rangle$. (*c*) Verify the Cauchy–Schwarz and triangle inequalities for these two functions.

♢ 12.2.34. Use separation of variables to construct a Fourier series solution to the Laplace equation on a rectangular box, $B = \{ 0 < x < a,\ 0 < y < b,\ 0 < z < c \}$, subject to the Dirichlet boundary conditions $u(x, y, z) = \begin{cases} h(x, y), & z = 0, \quad 0 < x < a, \quad 0 < y < b, \\ 0, & \text{at all other points in } \partial B. \end{cases}$

12.2.35. Find the equilibrium temperature distribution inside a unit cube that has 100° temperature on its top face, 0° on its bottom face, while all four side faces are insulated.

12.2.36. Solve Exercise 12.2.35 when the top face of the cube has temperature $u(x, y, 1) = \cos \pi x \, \cos \pi y$.

♣ 12.2.37. A solid unit cube is in thermal equilibrium when subject to 100° temperature on its top face and 0° on all other faces. *True or false*: The temperature at the center equals the average temperature over the surface of the cube.

12.2.38. Solve the boundary value problem

$$\begin{gathered} -\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} - \frac{\partial^2 u}{\partial z^2} + u = \cos x \cos y, \qquad 0 < x, y, z < \pi, \\ u(x, y, 0) = 1, \quad \frac{\partial u}{\partial z}(x, y, \pi) = \frac{\partial u}{\partial y}(x, 0, z) = \frac{\partial u}{\partial y}(x, \pi, z) = \frac{\partial u}{\partial z}(0, y, z) = \frac{\partial u}{\partial x}(\pi, y, z) = 0. \end{gathered}$$

12.2.39. Let C be the cylinder of height 1 and diameter 1 that sits on the (x, y)–plane centered on the z–axis. (*a*) Write out, in cylindrical coordinates, the explicit formula for the L^2 inner product and norm on C.
(*b*) Let $f(x, y, z) = z$ and $g(x, y, z) = x^2 + y^2$. Find $\| f \|, \| g \|$ and $\langle f, g \rangle$.
(*c*) Verify the Cauchy–Schwarz and triangle inequalities for these two functions.

♢ 12.2.40. (*a*) Write out the Laplace equation in cylindrical coordinates.
(*b*) Use separation of variables to construct a series solution to the Laplace equation on the cylinder $C = \{ x^2 + y^2 < 1,\ 0 < z < 1 \}$, subject to the Dirichlet boundary conditions

$$u(x, y, z) = \begin{cases} h(x, y), & z = 0, \quad x^2 + y^2 < 1, \\ 0, & \text{at all other points in } \partial C. \end{cases}$$

12.2.41. A cylinder of radius 1 and height 2 has 100° temperature on its top face, 0° on its bottom face, while its curved side is fully insulated. Find its equilibrium temperature distribution.

12.2.42. Solve Exercise 12.2.41 if the curved sides are kept at 0° instead.

12.3 Green's Functions for the Poisson Equation

We now turn to the inhomogeneous form of the three-dimensional Laplace equation: the *Poisson equation*

$$-\Delta u = f, \tag{12.67}$$

on a solid domain $\Omega \subset \mathbb{R}^3$. In order to uniquely specify the solution, we must impose appropriate boundary conditions: Dirichlet or mixed. (As in the planar version, Neumann boundary value problems have either infinitely many solutions or no solutions, depending upon whether the Fredholm conditions are satisfied or not.) We only need to discuss the case of homogeneous boundary conditions, since, by linear superposition, an inhomogeneous boundary value problem can be split into a homogeneous boundary value problem for the inhomogeneous Poisson equation along with an inhomogeneous boundary value problem for the homogeneous Laplace equation.

As in Chapter 6, we begin by analyzing the case of a delta function inhomogeneity that is concentrated at a single point in the domain. Thus, for each $\boldsymbol{\xi} = (\xi, \eta, \zeta) \in \Omega$, the *Green's function* $G(\mathbf{x}; \boldsymbol{\xi}) = G(x, y, z; \xi, \eta, \zeta)$ is the unique solution to the Poisson equation

$$-\Delta u = \delta(\mathbf{x} - \boldsymbol{\xi}) = \delta(x - \xi)\, \delta(y - \eta)\, \delta(z - \zeta) \qquad \text{for all} \qquad \mathbf{x} \in \Omega, \tag{12.68}$$

subject to the chosen homogeneous boundary conditions. The solution to the general Poisson equation (12.67) is then obtained by superposition: We write the forcing function

$$f(x, y, z) = \iiint_\Omega f(\xi, \eta, \zeta)\, \delta(x - \xi)\, \delta(y - \eta)\, \delta(z - \zeta)\, d\xi\, d\eta\, d\zeta \tag{12.69}$$

as a linear superposition of delta functions. By linearity, the solution

$$u(x, y, z) = \iiint_\Omega f(\xi, \eta, \zeta)\, G(x, y, z; \xi, \eta, \zeta)\, d\xi\, d\eta\, d\zeta \tag{12.70}$$

to the homogeneous boundary value problem for the Poisson equation (12.67) is then given as the corresponding superposition of the Green's function solutions.

The Green's function can also be used to solve the inhomogeneous Dirichlet boundary value problem

$$-\Delta u = 0, \qquad \mathbf{x} \in \Omega, \qquad\qquad u = h, \qquad \mathbf{x} \in \partial\Omega. \tag{12.71}$$

The same argument that was used in the two-dimensional situation produces the solution

$$u(\mathbf{x}) = -\iint_{\partial\Omega} \frac{\partial G}{\partial \mathbf{n}}\,(\mathbf{x};\boldsymbol{\xi})\, h(\boldsymbol{\xi})\, dS, \tag{12.72}$$

where the normal derivative is taken with respect to the variable $\boldsymbol{\xi} \in \partial\Omega$. In the case that Ω is a solid ball, this integral formula effectively sums the spherical harmonic series (12.51); see Theorem 12.12 below.

The Free–Space Green's Function

Only in a few specific instances is an explicit formula for the Green's function known. Nevertheless, certain general guiding features can be readily established. The starting point is to investigate the Poisson equation (12.68) when the domain $\Omega = \mathbb{R}^3$ is all of three-dimensional space. We impose boundary constraints by seeking a solution that goes to zero, $u(\mathbf{x}) \to 0$, at large distances, $\| \mathbf{x} \| \to \infty$. Since the Laplacian operator is invariant under translations, we can, without loss of generality, place our delta impulse at the origin, and concentrate on solving the particular case

$$-\Delta u = \delta(\mathbf{x})\,, \qquad\qquad \mathbf{x} \in \mathbb{R}^3.$$

Since $\delta(\mathbf{x}) = 0$ for all $\mathbf{x} \neq \mathbf{0}$, the desired solution will, in fact, be a solution to the homogeneous Laplace equation

$$\Delta u = 0, \qquad\qquad \mathbf{x} \neq \mathbf{0},$$

save, possibly, for a singularity at the origin.

The Laplace equation models the equilibria of a uniform isotropic medium, and so, as noted in Exercise 12.1.7, is also invariant under three-dimensional rotations. This suggests that, in any radially symmetric configuration, the solution should depend only on the distance $r = \| \mathbf{x} \|$ from the origin. Referring to the spherical coordinate form (12.16) of the Laplacian operator, if u is a function of r only, then its derivatives with respect to the angular coordinates φ, θ are zero, and so $u(r)$ solves the ordinary differential equation

$$\frac{d^2u}{dr^2} + \frac{2}{r}\,\frac{du}{dr} = 0. \tag{12.73}$$

This equation is, in effect, a first-order linear ordinary differential equation for $v = du/dr$ and hence is particularly easy to solve:

$$\frac{du}{dr} = v(r) = -\,\frac{b}{r^2}\,, \qquad \text{and hence} \qquad u(r) = a + \frac{b}{r}\,,$$

where a, b are arbitrary constants. The constant solution $u(r) = a$ does not die away at large distances, nor does it have a singularity at the origin. Therefore, if our intuition is valid, the desired solution should be of the form

$$u = \frac{b}{r} = \frac{b}{\| \mathbf{x} \|} = \frac{b}{\sqrt{x^2 + y^2 + z^2}}\,. \tag{12.74}$$

Indeed, this function is harmonic — solves Laplace's equation — everywhere away from the origin and has a singularity at $\mathbf{x} = \mathbf{0}$.

The solution (12.74) is, up to a constant multiple, the three-dimensional Newtonian gravitational potential due to a point mass at the origin. Its gradient,

$$\mathbf{f}(\mathbf{x}) \;=\; \nabla \left(\frac{b}{\| \mathbf{x} \|} \right) \;=\; - \frac{b\,\mathbf{x}}{\| \mathbf{x} \|^3}\,, \tag{12.75}$$

defines the gravitational force vector at the point $\mathbf{x}$. When $b > 0$, the force $\mathbf{f}(\mathbf{x})$ points toward the mass at the origin. Its magnitude

$$\| \mathbf{f} \| \;=\; \frac{b}{\| \mathbf{x} \|^2} \;=\; \frac{b}{r^2}$$

is proportional to the reciprocal of the squared distance, which is the well-known inverse square law of three-dimensional Newtonian gravity. Formula (12.75) can also be interpreted as the electrostatic force due to a concentrated electric charge at the origin, with (12.74) giving the corresponding Coulomb potential. The constant b is positive when the charges are of opposite signs, leading to an attractive force, and negative in the repulsive case of like charges.

Returning to our problem, the remaining task is to fix the multiple b such that the Laplacian of our candidate solution (12.74) has a delta function singularity at the origin; equivalently, we must determine $a = 1/b$ such that

$$-\,\Delta(r^{-1}) = a\,\delta(\mathbf{x}). \tag{12.76}$$

This equation is certainly valid away from the origin, since $\delta(\mathbf{x}) = 0$ when $\mathbf{x} \neq \mathbf{0}$. To investigate near the singularity, we integrate both sides of (12.76) over a small solid ball $B_\varepsilon = \{\, \| \mathbf{x} \| \leq \varepsilon \,\}$ of radius ε:

$$-\iiint_{B_\varepsilon} \Delta(r^{-1})\; dx\,dy\,dz = \iiint_{B_\varepsilon} a\,\delta(\mathbf{x})\,dx\,dy\,dz = a, \tag{12.77}$$

where we used the definition of the delta function to evaluate the right-hand side. On the other hand, since $\Delta\, r^{-1} = \nabla \cdot \nabla\, r^{-1}$, we can use the divergence theorem (12.8) to evaluate the left-hand integral, whence

$$\iiint_{B_\varepsilon} \Delta(r^{-1})\,dx\,dy\,dz = \iiint_{B_\varepsilon} \nabla \cdot \nabla(r^{-1})\,dx\,dy\,dz = \iint_{S_\varepsilon} \frac{\partial}{\partial \mathbf{n}} \left(\frac{1}{r} \right) dS,$$

where the surface integral is over the bounding sphere $S_\varepsilon = \partial B_\varepsilon = \{\, \| \mathbf{x} \| = \varepsilon \,\}$. The sphere's unit normal $\mathbf{n}$ points in the radial direction, and hence the normal derivative coincides with differentiation with respect to r; in particular,

$$\frac{\partial}{\partial \mathbf{n}} \left(\frac{1}{r} \right) = \frac{\partial}{\partial r} \left(\frac{1}{r} \right) = - \frac{1}{r^2}\,.$$

The surface integral can now be explicitly evaluated:

$$\iint_{S_\varepsilon} \frac{\partial}{\partial \mathbf{n}} \left(\frac{1}{r} \right) dS = - \iint_{S_\varepsilon} \frac{1}{r^2}\; dS = - \iint_{S_\varepsilon} \frac{1}{\varepsilon^2}\; dS = -\,4\,\pi,$$

since S_ε has surface area $4\,\pi\,\varepsilon^2$. Substituting this result back into (12.77), we conclude that

$$a = 4\,\pi, \qquad \text{and hence} \qquad -\,\Delta\, r^{-1} = 4\,\pi\,\delta(\mathbf{x}). \tag{12.78}$$

This is our desired formula! We conclude that a solution to the Poisson equation with a delta function impulse at the origin is

$$G(x,y,z) = \frac{1}{4\pi r} = \frac{1}{4\pi \,\|\, \mathbf{x} \,\|} = \frac{1}{4\pi \sqrt{x^2+y^2+z^2}}\,, \tag{12.79}$$

which is the three-dimensional Newtonian potential due to a unit point mass situated at the origin.

If the singularity is concentrated at some other point $\boldsymbol{\xi} = (\xi, \eta, \zeta)$, then we merely translate the preceding solution. This leads immediately to the *free-space Green's function*

$$G(\mathbf{x}; \boldsymbol{\xi}) = G(\mathbf{x} - \boldsymbol{\xi}) = \frac{1}{4\pi \,\|\, \mathbf{x} - \boldsymbol{\xi} \,\|} = \frac{1}{4\pi \sqrt{(x-\xi)^2 + (y-\eta)^2 + (z-\zeta)^2}}\,. \tag{12.80}$$

The superposition principle (12.70) implies the following integral formula for the solutions to the Poisson equation on all of three-dimensional space.

Theorem 12.8. *Assuming that $f(\mathbf{x}) \to 0$ sufficiently rapidly as $\| \mathbf{x} \| \to \infty$, a particular solution to the Poisson equation*

$$-\Delta u = f, \qquad \text{for} \qquad \mathbf{x} \in \mathbb{R}^3, \tag{12.81}$$

is given by

$$u_\star(\mathbf{x}) = \frac{1}{4\pi} \iiint_{\mathbb{R}^3} \frac{f(\boldsymbol{\xi})}{\| \mathbf{x} - \boldsymbol{\xi} \|}\, d\boldsymbol{\xi} = \frac{1}{4\pi} \iiint_{\mathbb{R}^3} \frac{f(\xi,\eta,\zeta)\; d\xi\, d\eta\, d\zeta}{\sqrt{(x-\xi)^2 + (y-\eta)^2 + (z-\zeta)^2}}\,. \tag{12.82}$$

The general solution is $u(x,y,z) = u_\star(x,y,z) + w(x,y,z)$, where $w(x,y,z)$ is an arbitrary harmonic function.

Example 12.9. In this example, we compute the gravitational (or electrostatic) potential in three-dimensional space due to a uniform solid ball, e.g., a spherical planet such as the Earth. By rescaling, it suffices to consider the case in which the forcing function is equal to 1 inside a ball of radius 1 and zero outside:

$$f(\mathbf{x}) = \begin{cases} 1, & \| \mathbf{x} \| < 1, \\ 0, & \| \mathbf{x} \| > 1. \end{cases}$$

The particular solution to the resulting Poisson equation (12.81) is given by the integral

$$u(\mathbf{x}) = \frac{1}{4\pi} \iiint_{\| \boldsymbol{\xi} \| < 1} \frac{1}{\| \mathbf{x} - \boldsymbol{\xi} \|}\, d\xi\, d\eta\, d\zeta. \tag{12.83}$$

Clearly, since the forcing function is radially symmetric, the solution $u = u(r)$ is also radially symmetric. To evaluate the integral, then, we can take $\mathbf{x} = (0, 0, z)$ to lie on the z–axis, so that $r = \| \mathbf{x} \| = | z |$. We use cylindrical coordinates $\boldsymbol{\xi} = (\rho \cos\theta, \rho \sin\theta, \zeta)$, so that

$$\| \mathbf{x} - \boldsymbol{\xi} \| = \sqrt{\rho^2 + (z-\zeta)^2}\,.$$

The integral in (12.83) can then be explicitly computed:

$$\begin{aligned} \frac{1}{4\pi} \int_{-1}^{1} \int_{0}^{\sqrt{1-\zeta^2}} \int_{0}^{2\pi} \frac{\rho\, d\theta\, d\rho\, d\zeta}{\sqrt{\rho^2 + (z-\zeta)^2}} \\ = \frac{1}{2} \int_{-1}^{1} \Big(\sqrt{1 + z^2 - 2z\zeta} \;-\; | z - \zeta | \Big)\, d\zeta = \begin{cases} \dfrac{1}{3\,|z|}\,, & |z| \geq 1, \\[2ex] \dfrac{1}{2} - \dfrac{z^2}{6}\,, & |z| \leq 1. \end{cases} \end{aligned}$$

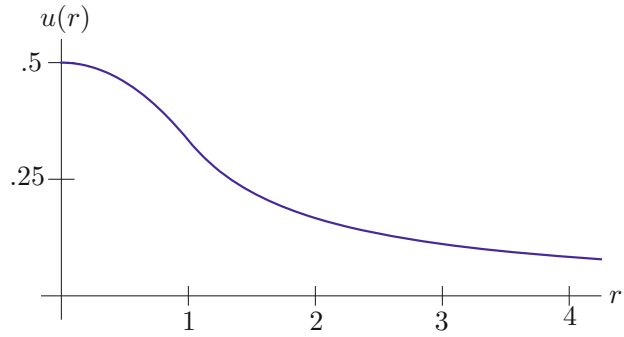


Figure 12.6. Solution to Poisson's equation in a solid ball.

Therefore, by radial symmetry, the solution is

$$u(\mathbf{x}) = \begin{cases} \dfrac{1}{3\,r}\,, & r = \| \mathbf{x} \| \geq 1, \\ \dfrac{1}{2} - \dfrac{r^2}{6}\,, & r = \| \mathbf{x} \| \leq 1, \end{cases} \tag{12.84}$$

plotted, as a function of $r = \| \mathbf{x} \|$, in Figure 12.6. Note that, outside the solid ball, the solution is a Newtonian potential corresponding to a concentrated point mass of magnitude $\frac{4}{3}\pi$ — the total mass of the planet. We have thus demonstrated a well-known result in gravitation and electrostatics: the exterior potential due to a spherically symmetric mass (or electrically charged body) is the same as if all the mass (charge) were concentrated at its center. In the darkness of outer space, if you cannot see a spherical planet, you can determine only its mass, not its size, by measuring its external gravitational force.

Bounded Domains and the Method of Images

Suppose we now wish to solve the inhomogeneous Poisson equation (12.67) on a bounded domain $\Omega \subset \mathbb{R}^3$. To construct the desired Green's function, we proceed as follows. The Newtonian potential (12.80) is a particular solution to the underlying inhomogeneous equation

$$-\Delta u = \delta(\mathbf{x} - \boldsymbol{\xi}), \qquad \mathbf{x} \in \Omega, \tag{12.85}$$

but it almost surely does not have the proper boundary values on $\partial\Omega$. By linearity, the general solution to such an inhomogeneous linear equation must take the form

$$u(\mathbf{x}) = \frac{1}{4\,\pi \,\| \mathbf{x} - \boldsymbol{\xi} \|} \; - \; v(\mathbf{x}), \tag{12.86}$$

where the first term is a particular solution, while $v(\mathbf{x})$ is an arbitrary solution to the homogeneous equation $\Delta v = 0$, i.e., an arbitrary harmonic function. The solution (12.86) satisfies the homogeneous boundary conditions, provided the boundary values of $v(\mathbf{x})$ match those of the Green's function. Let us explicitly state the result in the Dirichlet case.

Theorem 12.10. *The Green's function for the homogeneous Dirichlet boundary value problem*

$$-\Delta u = f \quad \text{for} \quad \mathbf{x} \in \Omega, \qquad u = 0 \quad \text{for} \quad \| \mathbf{x} \| \in \partial\Omega,$$

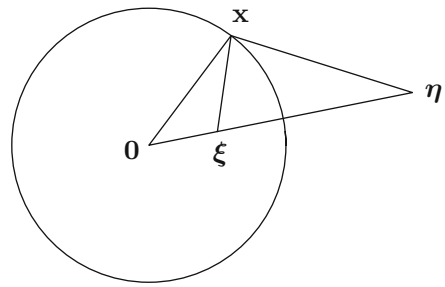


Figure 12.7. Method of Images for the unit ball.

for the Poisson equation in a domain $\Omega \subset \mathbb{R}^3$ *has the form*

$$G(\mathbf{x};\boldsymbol{\xi}) = \frac{1}{4\pi \,\|\, \mathbf{x} - \boldsymbol{\xi} \,\|} \;-\; v(\mathbf{x};\boldsymbol{\xi}), \qquad \mathbf{x},\boldsymbol{\xi} \in \Omega, \tag{12.87}$$

where $v(\mathbf{x};\boldsymbol{\xi})$ *is the harmonic function of* $\mathbf{x} \in \Omega$ *that satisfies*

$$v(\mathbf{x};\boldsymbol{\xi}) = \frac{1}{4\pi \,\|\, \mathbf{x} - \boldsymbol{\xi} \,\|} \qquad \textit{for all} \qquad \mathbf{x} \in \partial\Omega. \tag{12.88}$$

In this manner, we have reduced the determination of the Green's function to the solution to a particular family of Laplace boundary value problems, which are parametrized by the point $\boldsymbol{\xi} \in \Omega$. In certain domains with simple geometry, the Method of Images can be used to produce an explicit formula for the Green's function. As in Section 6.3, the idea is to match the boundary values of the free-space Green's function due to a delta impulse at a point inside the domain with one or more additional Green's functions corresponding to impulses at points outside the domain — the "image points".

The case of a solid ball of radius 1 with Dirichlet boundary conditions is the easiest to handle. Indeed, the *same* geometric construction that we used for a planar disk, redrawn in Figure 12.7, applies here. Although identical to Figure 6.13, we are re-interpreting it as a three-dimensional diagram, with the circle representing the unit sphere, while the lines remain lines. The required image point is given by *inversion*:

$$\boldsymbol{\eta} = \frac{\boldsymbol{\xi}}{\|\, \boldsymbol{\xi} \,\|^2}, \qquad \text{whereby} \qquad \|\, \boldsymbol{\xi} \,\| = \frac{1}{\|\, \boldsymbol{\eta} \,\|}.$$

By the similar triangles argument used before, we have

$$\frac{\|\, \boldsymbol{\xi} \,\|}{\|\, \mathbf{x} \,\|} = \frac{\|\, \mathbf{x} \,\|}{\|\, \boldsymbol{\eta} \,\|} = \frac{\|\, \mathbf{x} - \boldsymbol{\xi} \,\|}{\|\, \mathbf{x} - \boldsymbol{\eta} \,\|}, \qquad \text{and therefore} \qquad \|\, \mathbf{x} \,\| = 1.$$

As a result, the function

$$v(\mathbf{x},\boldsymbol{\xi}) = \frac{1}{4\pi} \, \frac{\|\, \boldsymbol{\eta} \,\|}{\|\, \mathbf{x} - \boldsymbol{\eta} \,\|} = \frac{1}{4\pi} \, \frac{\|\, \boldsymbol{\xi} \,\|}{\|\, \boldsymbol{\xi} - \|\, \boldsymbol{\xi} \,\|^2 \, \mathbf{x} \,\|}$$

has the same boundary values on the unit sphere as the Newtonian potential:

$$\frac{1}{4\pi} \, \frac{\|\, \boldsymbol{\eta} \,\|}{\|\, \mathbf{x} - \boldsymbol{\eta} \,\|} \;=\; \frac{1}{4\pi \,\|\, \mathbf{x} - \boldsymbol{\xi} \,\|} \qquad \text{whenever} \qquad \|\, \mathbf{x} \,\| = 1.$$

We conclude that their difference

$$G(\mathbf{x};\boldsymbol{\xi}) = \frac{1}{4\pi}\left(\frac{1}{\| \mathbf{x} - \boldsymbol{\xi} \|} \;-\; \frac{\| \boldsymbol{\xi} \|}{\| \boldsymbol{\xi} - \| \boldsymbol{\xi} \|^2 \mathbf{x} \|} \right) \tag{12.89}$$

has the required properties of the Green's function: it satisfies the Laplace equation inside the unit ball except at the delta function singularity $\mathbf{x} = \boldsymbol{\xi}$, and, moreover, $G(\mathbf{x};\boldsymbol{\xi}) = 0$ has homogeneous Dirichlet conditions on the spherical boundary $\| \mathbf{x} \| = 1$.

With the Green's function in hand, we can apply the general superposition formula (12.70) to arrive at a solution to the Dirichlet boundary value problem for the Poisson equation in the unit ball.

Theorem 12.11. *The solution to the homogeneous Dirichlet boundary value problem*

$$-\Delta u = f \qquad \text{for} \qquad \| \mathbf{x} \| < 1, \qquad\qquad u = 0 \qquad \text{for} \qquad \| \mathbf{x} \| = 1,$$

on the unit ball is given by the integral

$$u(\mathbf{x}) = \frac{1}{4\pi} \iiint_{\| \boldsymbol{\xi} \| \le 1} \left(\frac{1}{\| \mathbf{x} - \boldsymbol{\xi} \|} \;-\; \frac{\| \boldsymbol{\xi} \|}{\| \boldsymbol{\xi} - \| \boldsymbol{\xi} \|^2 \mathbf{x} \|} \right) f(\boldsymbol{\xi})\, d\xi\, d\eta\, d\zeta. \tag{12.90}$$

By the same token, formula (12.72) provides a solution to the inhomogeneous Dirichlet boundary value problem for the Laplace equation on a ball.

Theorem 12.12. *The solution to the homogeneous Dirichlet boundary value problem*

$$-\Delta u = 0 \qquad \text{for} \qquad \| \mathbf{x} \| < 1, \qquad\qquad u = h \qquad \text{for} \qquad \| \mathbf{x} \| = 1,$$

is given by the following surface integral:

$$u(\mathbf{x}) = \frac{1}{4\pi} \iint_{\| \boldsymbol{\xi} \| = 1} \frac{1 - \| \mathbf{x} \|^2}{\| \boldsymbol{\xi} - \mathbf{x} \|^3}\, h(\boldsymbol{\xi})\, dS. \tag{12.91}$$

Proof: We start with the explicit formula (12.89) for the Green's function on the unit ball. Since the normal derivative on the unit sphere $\| \boldsymbol{\xi} \| = 1$ can be written as $\partial/\partial \mathbf{n} = \boldsymbol{\xi} \cdot \nabla_{\boldsymbol{\xi}}$, a short computation demonstrates that

$$\frac{\partial G}{\partial \mathbf{n}}(\mathbf{x};\boldsymbol{\xi}) = \frac{1}{4\pi}\left(\frac{\mathbf{x}\cdot\boldsymbol{\xi} - \| \boldsymbol{\xi} \|^2}{\| \mathbf{x} - \boldsymbol{\xi} \|^3} \;-\; \frac{\| \boldsymbol{\xi} \|^3 \big(\mathbf{x}\cdot\boldsymbol{\xi} - \| \boldsymbol{\xi} \|^2 \| \mathbf{x} \|^2 \big)}{\| \boldsymbol{\xi} - \| \boldsymbol{\xi} \|^2 \mathbf{x} \|^3} \right) = \frac{1}{4\pi} \frac{\| \mathbf{x} \|^2 - 1}{\| \boldsymbol{\xi} - \mathbf{x} \|^3}.$$

The solution formula (12.91) thus immediately follows from (12.72). *Q.E.D.*

For example, the series solution (12.60) to the spherical capacitor problem of Example 12.6 can thus be re-expressed as a surface integral:

$$\begin{aligned} u(x,y,z) &= \frac{1}{4\pi} \iint_{\{\, \xi^2+\eta^2+\zeta^2 = 1,\ \zeta > 0 \,\}} \frac{(1 - x^2 - y^2 - z^2)\, dS}{\big[\, (\xi - x)^2 + (\eta - y)^2 + (\zeta - z)^2 \,\big]^{3/2}} \\ &= \int_{-\pi}^{\pi} \int_{0}^{\pi/2} \frac{(1 - x^2 - y^2 - z^2) \sin\varphi\, d\varphi\, d\theta}{\big[\, (\cos\theta \sin\varphi - x)^2 + (\sin\theta \sin\varphi - y)^2 + (\cos\varphi - z)^2 \,\big]^{3/2}}\,. \end{aligned}$$

Exercises

12.3.1. Find the equilibrium temperature of a sphere of radius 1 whose boundary is held at 0° while a concentrated unit heat source is applied at (*a*) the center; (*b*) a point half-way between the center and the boundary.

12.3.2. A hot soldering iron is continually applied to the north pole of a solid spherical ball of radius 1. Find the equilibrium temperature.

12.3.3. Write down the gravitional potential — both external and internal — due to a spherical planet of radius R composed out of a uniform material with density ρ.

12.3.4. (*a*) Find the gravitational potential due to a spherical shell of unit density obtained by carving out a spherical cavity of radius a from a solid ball of radius $b > a$. *Hint*: Use the solution to Exercise 12.3.3. (*b*) What is the gravitational force inside the cavity? (*c*) Show that outside the shell, the gravitational potential is as if the entire mass were concentrated at the origin.

♣ 12.3.5. (*a*) Write down an integral formula for the gravitational potential and gravitational force field due to a mass of unit density in the shape of a solid unit cube that is centered at the origin. (*b*) Use numerical integration to determine the gravitational force vector at the points $(3, 0, 0)$ and $\left(\sqrt{3}, \sqrt{3}, \sqrt{3}\,\right)$. Before doing the calculation, see whether you can predict which experiences a stronger force, and then check your prediction numerically. (*c*) Suppose the mass is re-formed into a sphere. How does this affect the gravitational force at the two points? First predict whether it will increase, decrease, or stay the same. Then test your prediction by computing the values and comparing with those you computed in part (*b*).

12.3.6. A thin hollow metal sphere of unit radius is grounded. Find the electrostatic potential inside the sphere due to a small solid metal ball of radius $\rho < 1$ placed at its center, assuming unit charge density throughout the ball.

12.3.7. A thin straight rod of unit density and length 2ℓ is fixed on the z–axis centered at the origin. Find the induced (*a*) gravitational potential and (*b*) gravitational force experienced by a point (x, y, z) not on the rod.

♡ 12.3.8. (*a*) Find the gravitational force due to a thin, uniform straight rod of unit density and infinite length by letting $\ell \to \infty$ in your solution to Exercise 12.3.7(*b*). (*b*) Show that the force field of part (*a*) has a potential function that can be identified with the two-dimensional logarithmic gravitational potential due to a point mass at the origin. Thus, two-dimensional gravitation can be regarded as a cross-section of three-dimensional gravitation due to infinitely long vertical line masses. (*c*) Is your potential function the limit, as $\ell \to \infty$, of the potential function you found in Exercise 12.3.7(a)? Discuss.

12.3.9. Which well-known solutions to the Laplace equation comes from setting $m = n = 0$ in (12.61)?

12.3.10. Use the Fredholm Alternative to analyze the existence and uniqueness of solutions to the homogeneous Neumann boundary value problem for the Poisson equation on a bounded domain $\Omega \subset \mathbb{R}^3$.

♢ 12.3.11. Mimic the proof of Theorem 6.19 to establish the solution formula (12.72).

12.3.12. Use the Method of Images to find the Green's function for a solid hemisphere of unit radius subject to homogeneous Dirichlet boundary conditions.

12.4 The Heat Equation for Three–Dimensional Media

Thermal diffusion in a uniform isotropic solid body $\Omega \subset \mathbb{R}^3$ is modeled by the three-dimensional *heat equation*

$$\frac{\partial u}{\partial t} = \gamma\, \Delta u = \gamma \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \right), \qquad (x, y, z) \in \Omega. \tag{12.92}$$

The positivity of the body's thermal diffusivity, $\gamma > 0$, is required on both physical and mathematical grounds. The physical derivation is exactly the same as that for the two-dimensional version (11.1), and does not need to be repeated in detail. Briefly, Fourier's law expresses the heat flux vector as a multiple of the temperature gradient, $\mathbf{w} = -\kappa \nabla u$, while energy conservation implies that its divergence is proportional to the rate of change of temperature: $\nabla \cdot \mathbf{w} = -\sigma u_t$. Combining these two physical laws and assuming uniformity, whereby κ and σ are constant, produces (12.92) with $\gamma = \kappa/\sigma$.

As always, we must impose suitable boundary conditions: either Dirichlet conditions $u = h$ that specify the boundary temperature; (homogeneous) Neumann conditions $\partial u/\partial \mathbf{n} = 0$ corresponding to an insulated boundary; or a mixture of the two. Given the body's temperature

$$u(t_0, x, y, z) = f(x, y, z) \tag{12.93}$$

at an initial time t_0, it can be proved, [**38**, **61**, **99**], that the resulting initial-boundary value problem is well-posed, which means that there is a unique classical solution $u(t, x, y, z)$, defined at all subsequent times $t > t_0$, that depends continuously on the initial data.

As in the one- and two-dimensional versions, we begin by restricting our attention to homogeneous boundary conditions. Separation of variables works as usual, and we quickly review the basic ideas. One begins by imposing an exponential solution ansatz

$$u(t, \mathbf{x}) = e^{-\lambda t}\, v(\mathbf{x}).$$

Substituting into the differential equation and canceling the exponentials, it follows that v satisfies the Helmholtz eigenvalue problem

$$\gamma\, \Delta v + \lambda\, v = 0,$$

subject to the relevant boundary conditions. For Dirichlet and mixed boundary conditions, the Laplacian is a positive definite operator, and hence the eigenvalues are all strictly positive,

$$0 < \lambda_1 \le \lambda_2 \le \cdots, \qquad \text{with} \qquad \lambda_n \longrightarrow \infty, \qquad \text{as} \qquad n \to \infty.$$

Moreover, on a bounded domain, the Helmholtz eigenfunctions are complete, and so linear superposition implies that the solution can be written as an eigenfunction series

$$u(t, \mathbf{x}) = \sum_{n=1}^{\infty} c_n\, e^{-\lambda_n t}\, v_n(\mathbf{x}). \tag{12.94}$$

The coefficients c_n are uniquely prescribed by the initial condition (12.93):

$$u(t_0, \mathbf{x}) = \sum_{n=1}^{\infty} c_n\, e^{-\lambda_n t_0}\, v_n(\mathbf{x}) = f(\mathbf{x}). \tag{12.95}$$

Self-adjointness of the boundary value problem implies orthogonality of the eigenfunctions, and hence the coefficients are obtained via the usual inner product formulae:

$$c_n = e^{\lambda_n t_0} \frac{\langle f, v_n \rangle}{\| v_n \|^2} = e^{\lambda_n t_0} \frac{\iiint_\Omega f(\mathbf{x})\, v_n(\mathbf{x})\, dx\, dy\, dz}{\iiint_\Omega v_n(\mathbf{x})^2\, dx\, dy\, dz}. \tag{12.96}$$

The resulting solution decays exponentially fast to thermal equilibrium, $u(t, \mathbf{x}) \to 0$ as $t \to \infty$, typically at a rate equal to the smallest positive eigenvalue $\lambda_1 > 0$, although special solutions, whose initial series coefficients vanish, will decay at a faster rate governed by a higher eigenvalue. Since the higher modes — the terms with $n \gg 0$ — go to zero extremely rapidly with increasing t, the solution can be well approximated by the first few terms in its eigenfunction expansion. As a consequence, the heat equation rapidly smooths out discontinuities and eliminates high-frequency noise in the initial data.

Unfortunately, explicit formulas for the eigenfunctions and eigenvalues are rare. Most explicit eigensolutions of the Helmholtz boundary value problem require a further separation of variables. In a rectangular box, one separates the solution into a product of functions depending upon the individual Cartesian coordinates, and the eigenfunctions are written as products of trigonometric functions; see Exercise 12.4.1 for details. In a cylindrical domain, the separation is effected in cylindrical coordinates, which leads to eigensolutions involving trigonometric and Bessel functions, as outlined in Exercise 12.4.5. The most interesting and enlightening case is a spherical domain, and we treat this particular problem in complete detail in the ensuing subsection.

Exercises

♢ 12.4.1. Let $B = \{ 0 < x < a,\ 0 < y < b,\ 0 < z < c \}$ be a solid box of size $a \times b \times c$.
(*a*) Write down an initial-boundary value problem for the thermodynamics of the box when all its sides are all held at 0° and its initial temperature is $f(x, y, z)$. (*b*) Use separation of variables to construct the normal mode solutions. (*c*) Write down a series representing the general solution to the initial-boundary value problem. What are the formulas for the coefficients in your series? (*d*) What is the equilibrium temperature? How fast does the temperature in the box decay to equilibrium?

12.4.2. *True or false*: In the context of Exercise 12.4.1, among all boxes of a given volume V, a cube decays slowest to thermal equilibrium. What is the cube's decay rate?

12.4.3. Answer Exercises 12.4.1 and 12.4.2 when the top of the box, where $z = c$, is insulated.

12.4.4. A rectangular brick of size 1 cm × 2 cm × 3 cm made out of material with diffusion coefficient $\gamma = 6$ is insulated on five sides, while one of its small ends is held at temperature $u(x, y, 0) = \cos \pi x \, \cos 2\pi y$. (*a*) Find the eventual equilibrium temperature distribution. (*b*) If the brick is initially heated in an oven, how fast does it return to equilibrium?

♢ 12.4.5. Let $C = \left\{ 0 \le \sqrt{x^2 + y^2} < a,\ 0 < z < h \right\}$ be a solid cylinder of radius a and height h.
(*a*) Write down an initial-boundary value problem in cylindrical coordinates for the thermodynamics of the cylinder when its sides, top, and bottom are all held at 0°.
(*b*) Use separation of variables to write down a series representing the general solution to the initial-boundary value problem. What are the formulas for the coefficients in your series?

(*c*) What is the eventual equilibrium temperature?
(*d*) How fast does the temperature in the cylinder go to equilibrium?

12.4.6. Find the solution to the initial-boundary value problem in Exercise 12.4.5 when the initial temperature of the cylinder is uniformly 30°. *Hint*: Use (11.112) to evaluate the coefficients.

♡ 12.4.7. A cylindrical can that contains 355 ml of soda is removed from the refrigerator. Find the optimal cylindrical shape for such a can in order to keep the soda cold the longest. Is this the manufactured shape of a standard soda can?

♡ 12.4.8. *True or false*: Among all solid cylinders of a given volume, the one that reaches thermal equilibrium the slowest, when subject to homogeneous Dirichlet boundary conditions, is the one that has the least surface area. Justify your answer.

♡ 12.4.9. Among all fully insulated solid cylinders of unit volume, which cools down
(*i*) the slowest? (*ii*) the fastest?

♢ 12.4.10. Write down a series for the solution to the homogeneous Neumann boundary value problem for the heat equation on a bounded domain $\Omega \subset \mathbb{R}^3$, corresponding to the thermodynamics of a completely insulated solid body. What is the equilibrium temperature of the body? Does the solution decay to equilibrium? If so, how fast?

♢ 12.4.11. Suppose $u(t, x, y, z)$ is a solution to the heat equation on a fully insulated bounded domain $\Omega \subset \mathbb{R}^3$. Use the identities in Exercise 12.1.11 to prove the following:
(*a*) The total heat $H(t) = \iiint_\Omega u(t, x, y, z)\, dx\, dy\, dz$ is conserved, i.e., is constant. Explain how this can be used to determine the equilibrium temperature of the body.
(*b*) If u is a non-equilibrium solution, its squared L^2 norm $E(t) = \iiint_\Omega u(t, x, y, z)^2\, dx\, dy\, dz$ is a strictly decreasing function of t.
(*c*) Use part (*b*) to prove uniqueness of solutions to the initial value problem.

♢ 12.4.12. State and prove a Maximum Principle for the three-dimensional heat equation.

Heating of a Ball

Our goal is to study heat propagation in a solid spherical body, e.g., the Earth.[†] For simplicity, we take the diffusivity $\gamma = 1$, and consider the heat equation on a solid spherical ball of unit radius, $B_1 = \{\, \| \mathbf{x} \| < 1 \,\}$, that is subject to homogeneous Dirichlet boundary conditions. Once we know how to solve this particular case, an easy scaling argument, as outlined in Exercise 12.4.16, will allow us to find the solution for a ball of arbitrary radius and general diffusivity.

As usual, when dealing with a spherical geometry, we adopt spherical coordinates r, φ, θ, as in (12.15), in terms of which the heat equation takes the form

$$\frac{\partial u}{\partial t} = \Delta u = \frac{\partial^2 u}{\partial r^2} + \frac{2}{r}\,\frac{\partial u}{\partial r} + \frac{1}{r^2}\,\frac{\partial^2 u}{\partial \varphi^2} + \frac{\cos \varphi}{r^2 \sin \varphi}\,\frac{\partial u}{\partial \varphi} + \frac{1}{r^2 \sin^2 \varphi}\,\frac{\partial^2 u}{\partial \theta^2}\,, \tag{12.97}$$

where we have used our handy spherical coordinate formula (12.16) for the Laplacian. The

[†] In this admittedly simplistic model, we are assuming that the Earth is composed of a completely uniform and isotropic solid material.

standard diffusive separation of variables ansatz

$$u(t, r, \varphi, \theta) = e^{-\lambda t}\, v(r, \varphi, \theta)$$

requires us to analyze the spherical coordinate form of the Helmholtz equation

$$\Delta v + \lambda\, v = \frac{\partial^2 v}{\partial r^2} + \frac{2}{r}\,\frac{\partial v}{\partial r} + \frac{1}{r^2}\,\frac{\partial^2 v}{\partial \varphi^2} + \frac{\cos\varphi}{r^2 \sin\varphi}\,\frac{\partial v}{\partial \varphi} + \frac{1}{r^2 \sin^2\varphi}\,\frac{\partial^2 v}{\partial \theta^2} + \lambda\, v = 0 \tag{12.98}$$

on the unit ball $\Omega = \{\, r < 1 \,\}$ under homogeneous Dirichlet boundary conditions. To make further progress, we invoke a second variable separation, splitting off the radial coordinate by setting

$$v(r, \varphi, \theta) = p(r)\, w(\varphi, \theta).$$

The function w must be 2π–periodic in θ and well defined on the z–axis, i.e., when $\varphi = 0, \pi$. Substituting this ansatz into (12.98), and separating all the r-dependent terms from those terms depending on the angular variables φ, θ leads to a pair of differential equations involving a separation constant, denoted by μ. The first is an ordinary differential equation

$$r^2\,\frac{d^2 p}{dr^2} + 2\,r\,\frac{dp}{dr} + (\lambda\, r^2 - \mu) p = 0, \tag{12.99}$$

for the radial component $p(r)$, while the second is a familiar partial differential equation

$$\Delta_S\, w + \mu\, w = \frac{\partial^2 w}{\partial \varphi^2} + \frac{\cos\varphi}{\sin\varphi}\,\frac{\partial w}{\partial \varphi} + \frac{1}{\sin^2\varphi}\,\frac{\partial^2 w}{\partial \theta^2} + \mu\, w = 0, \tag{12.100}$$

for its angular counterpart $w(\varphi, \theta)$. The operator Δ_S is the *spherical Laplacian* from (12.19). In Section 12.2, we showed that its eigenvalues are

$$\mu_m = m(m+1) \qquad \text{for} \qquad m = 0, 1, 2, 3, \ldots.$$

The m^{th} eigenvalue admits $2m+1$ linearly independent eigenfunctions: the spherical harmonics $Y_m^0, \ldots, Y_m^m, \widetilde{Y}_m^1, \ldots, \widetilde{Y}_m^m$ defined in (12.38).

Spherical Bessel Functions

The radial ordinary differential equation (12.99) can be solved by setting

$$q(r) = \sqrt{r}\, p(r). \tag{12.101}$$

We use the product rule to relate the derivatives of q and p, whereby

$$p = \frac{q}{r^{1/2}}\,, \qquad \frac{dp}{dr} = \frac{1}{r^{1/2}}\,\frac{dq}{dr} - \frac{q}{2\,r^{3/2}}\,, \qquad \frac{d^2 p}{dr^2} = \frac{1}{r^{1/2}}\,\frac{d^2 q}{dr^2} - \frac{1}{r^{3/2}}\,\frac{dq}{dr} + \frac{3\,q}{4\,r^{5/2}}\,.$$

Substituting these expressions back into (12.99) with $\mu = \mu_m = m(m+1)$ and multiplying the resulting equation by $\sqrt{r}$, we discover that $q(r)$ must solve the differential equation

$$r^2\,\frac{d^2 q}{dr^2} + r\,\frac{dq}{dr} + \big[\, \lambda\, r^2 - \big(m + \tfrac{1}{2}\big)^2 \,\big] q = 0, \tag{12.102}$$

which we recognize as the rescaled Bessel equation (11.56) of half-integer order $m + \frac{1}{2}$. Consequently, the solution to (12.102) that remains bounded at $r = 0$ is (up to a scalar multiple) the rescaled Bessel function

$$q(r) = J_{m+1/2}\big(\sqrt{\lambda}\, r\big).$$

The corresponding solution

$$p(r) = r^{-1/2}\, J_{m+1/2}\big(\sqrt{\lambda}\, r\big) \tag{12.103}$$

to (12.99) is important enough to warrant a special name.

Definition 12.13. The *spherical Bessel function* of order $m \geq 0$ is defined by the formula

$$S_m(x) = \sqrt{\frac{\pi}{2\,x}}\; J_{m+1/2}(x). \tag{12.104}$$

Remark: The multiplicative factor $\sqrt{\pi/2}$ is included in the definition so as to avoid annoying factors of $\sqrt{\pi}$ and $\sqrt{2}$ in the subsequent formulas.

Surprisingly, unlike the Bessel functions of integer order, the spherical Bessel functions are all elementary functions! Comparing (12.104) with (11.105), we see that the spherical Bessel function of order 0 is

$$S_0(x) = \frac{\sin x}{x}\,. \tag{12.105}$$

The corresponding explicit formulas for the higher-order spherical Bessel functions can be obtained through the general recurrence relation

$$S_{m+1}(x) = -\,\frac{dS_m}{dx} + \frac{m}{x}\, S_m(x), \tag{12.106}$$

which is a consequence of the Bessel function recurrence formula (11.111). Indeed,

$$\begin{aligned}
\frac{dS_m}{dx} &= \sqrt{\frac{\pi}{2\,x}}\;\frac{dJ_{m+1/2}}{dx} - \frac{1}{2}\sqrt{\frac{\pi}{2}}\;\frac{1}{x^{3/2}}\; J_{m+1/2}(x) \\
&= -\sqrt{\frac{\pi}{2\,x}}\;\left[\, J_{m+3/2}(x) + \frac{m+\frac{1}{2}}{x}\; J_{m+1/2}(x)\,\right] - \frac{1}{2}\sqrt{\frac{\pi}{2}}\;\frac{1}{x^{3/2}}\; J_{m+1/2}(x) \\
&= -\sqrt{\frac{\pi}{2\,x}}\; J_{m+3/2}(x) + \frac{m}{x}\,\sqrt{\frac{\pi}{2\,x}}\; J_{m+1/2}(x) = -\,S_{m+1}(x) + \frac{m}{x}\, S_m(x).
\end{aligned}$$

The next few spherical Bessel functions are, therefore,

$$\begin{aligned}
S_1(x) &= -\,\frac{dS_0}{dx} && = -\,\frac{\cos x}{x} + \frac{\sin x}{x^2}\,, \\
S_2(x) &= -\,\frac{dS_1}{dx} + \frac{S_1}{x} && = -\,\frac{\sin x}{x} - \frac{3\cos x}{x^2} + \frac{3\sin x}{x^3}\,, \\
S_3(x) &= -\,\frac{dS_2}{dx} + \frac{2\,S_2}{x} = \frac{\cos x}{x} - \frac{6\sin x}{x^2} - \frac{15\cos x}{x^3} + \frac{15\sin x}{x^4}\,,
\end{aligned} \tag{12.107}$$

and so on. Figure 11.4 provides graphs of the first four spherical Bessel functions on the interval $0 \leq x \leq 20$; the vertical axes range from $-.5$ to 1.0. We note that

$$S_0(0) = 1, \qquad \text{whereas} \qquad S_m(0) = 0 \qquad \text{for} \qquad m > 0, \tag{12.108}$$

whose proof is the task of Exercise 12.4.26. Thus, our radial solution (12.103) is, apart from an inessential constant multiple, a rescaled spherical Bessel function of order m:

$$p(r) = S_m\big(\sqrt{\lambda}\, r\big)\,.$$

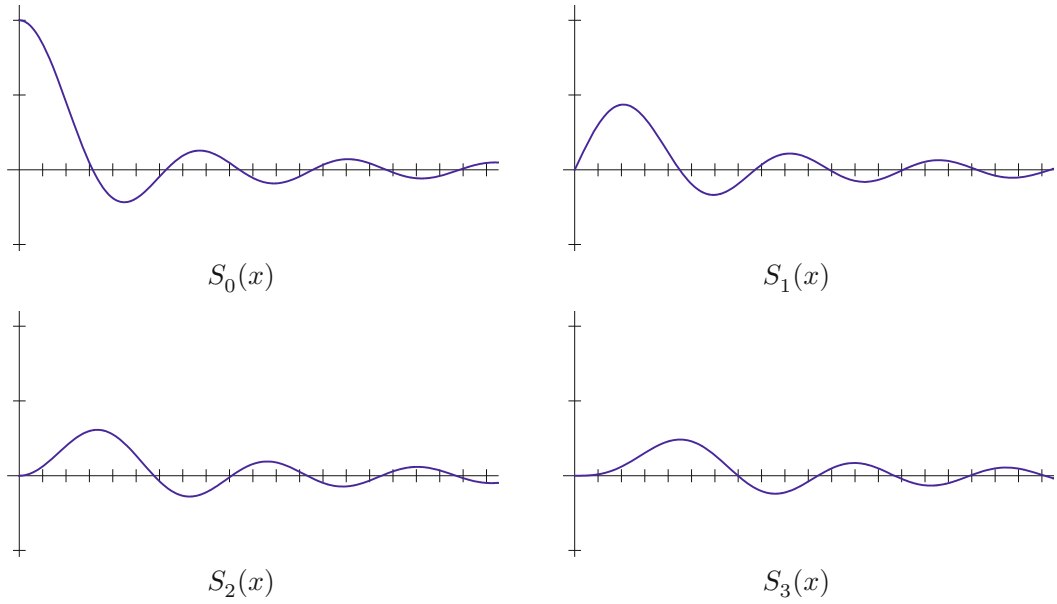


Figure 12.8. Spherical Bessel functions.

So far, we have not taken into account the (homogeneous) Dirichlet boundary condition at $r = 1$. This requires

$$p(1) = 0, \qquad \text{and hence} \qquad S_m\big(\sqrt{\lambda}\,\big) = 0.$$

Therefore, $\sqrt{\lambda}$ must be a root of the m^{th} order spherical Bessel function. We introduce the notation

$$0 < \sigma_{m,1} < \sigma_{m,2} < \sigma_{m,3} < \ \cdots$$

to denote the successive (positive) *spherical Bessel roots*, satisfying

$$S_m(\sigma_{m,n}) = 0 \qquad \text{for} \qquad n = 1, 2, \ \ldots \ . \tag{12.109}$$

In particular the roots of the zero$^{\text{th}}$ order spherical Bessel function $S_0(x) = x^{-1} \sin x$ are just the integer multiples of π:

$$\sigma_{0,n} = n\,\pi \qquad \text{for} \qquad n = 1, 2, \ \ldots \ .$$

The higher-order roots are not expressible in terms of known constants. A table of all spherical Bessel roots that are < 13 appears below. The columns of the table are indexed by m, the order, while the rows are indexed by n, the root number.

Re-assembling the individual constituents, we have now demonstrated that the separable eigenfunctions of the Helmholtz equation on a solid ball of radius 1, when subject to homogeneous Dirichlet boundary conditions, are products of spherical Bessel functions and spherical harmonics,

$$\begin{aligned} v_{k,m,n}(r,\varphi,\theta) &= S_m(\sigma_{m,n}\,r)\,Y_m^k(\varphi,\theta), \\ \widetilde{v}_{k,m,n}(r,\varphi,\theta) &= S_m(\sigma_{m,n}\,r)\,\widetilde{Y}_m^k(\varphi,\theta), \end{aligned} \qquad \begin{aligned} &m = 0, 1, 2, \ldots \ , \\ &k = 0, \ldots, m, \\ &n = 1, 2, 3, \ldots \ . \end{aligned} \tag{12.110}$$

Spherical Bessel Roots $\sigma_{m,n}$

n \ m	0	1	2	3	4	5	6	7	
1	3.1416	4.4934	5.7635	8.1826	9.3558	10.5128	11.6570	12.7908	...
2	6.2832	7.7253	9.0950	11.7049	12.9665	⋮	⋮	⋮	
3	9.4248	10.9041	12.3229	⋮	⋮				
4	12.5664	⋮	⋮						
⋮	⋮								

The corresponding eigenvalues

$$\lambda_{m,n} = \sigma_{m,n}^2, \qquad m = 0, 1, 2, \dots, \qquad n = 1, 2, 3, \dots, \tag{12.111}$$

are the squared spherical Bessel roots. Since there are $2m + 1$ independent spherical harmonics of order m, the eigenvalue $\lambda_{m,n}$ admits $2m + 1$ linearly independent eigenfunctions, namely $v_{0,m,n}, \dots, v_{m,m,n}, \widetilde{v}_{1,m,n}, \dots, \widetilde{v}_{m,m,n}$. In particular, the radially symmetric solutions are the eigenfunctions with $k = m = 0$:

$$v_n(r) = v_{0,0,n}(r) = S_0(\sigma_{0,n}\, r) = \frac{\sin n\pi\, r}{n\pi\, r}, \qquad n = 1, 2, \dots. \tag{12.112}$$

Further analysis, cf. [**34**], demonstrates that the separable solutions (12.110) form a complete system of eigenfunctions for the Helmholtz equation on the unit ball with homogeneous Dirichlet boundary conditions.

We have thus completely determined the basic separable solutions to the heat equation on a solid unit ball subject to homogeneous Dirichlet boundary conditions. They are products of exponential functions of time, spherical Bessel functions of the radius, and spherical harmonics:

$$\begin{aligned} u_{k,m,n}(t, r, \varphi, \theta) &= e^{-\sigma_{m,n}^2 t}\, S_m(\sigma_{m,n}\, r)\, Y_m^k(\varphi, \theta), \\ \widetilde{u}_{k,m,n}(t, r, \varphi, \theta) &= e^{-\sigma_{m,n}^2 t}\, S_m(\sigma_{m,n}\, r)\, \widetilde{Y}_m^k(\varphi, \theta). \end{aligned} \tag{12.113}$$

The general solution can be written as an infinite "Fourier–Bessel–spherical harmonic" series in these fundamental modes:

$$\begin{aligned} u(t, r, \varphi, \theta) = \sum_{m=0}^{\infty} \sum_{n=1}^{\infty} e^{-\sigma_{m,n}^2 t}\, S_m(\sigma_{m,n}\, r) \Bigg(& \frac{c_{0,m,n}}{2}\, Y_m^0(\varphi, \theta) \\ & + \sum_{k=1}^{m} \Big[c_{k,m,n} Y_m^k(\varphi, \theta) + \widetilde{c}_{k,m,n} \widetilde{Y}_m^k(\varphi, \theta) \Big] \Bigg). \end{aligned} \tag{12.114}$$

The series' coefficients are uniquely prescribed by the initial data $u(0, r, \varphi, \theta) = f(r, \varphi, \theta)$,

and their explicit formulae[†]

$$c_{k,m,n} = \frac{(2m+1)(m-k)!}{\pi\,(m+k)!\,S_{m+1}(\sigma_{m,n})^2} \int_{-\pi}^{\pi}\int_{0}^{\pi}\int_{0}^{1} f(r,\varphi,\theta)\, v_{k,m,n}(r,\varphi,\theta)\, r^2 \sin\varphi\, dr\, d\varphi\, d\theta,$$
$$\widetilde{c}_{k,m,n} = \frac{(2m+1)(m-k)!}{\pi\,(m+k)!\,S_{m+1}(\sigma_{m,n})^2} \int_{-\pi}^{\pi}\int_{0}^{\pi}\int_{0}^{1} f(r,\varphi,\theta)\, \widetilde{v}_{k,m,n}(r,\varphi,\theta)\, r^2 \sin\varphi\, dr\, d\varphi\, d\theta, \tag{12.115}$$

follow from the usual orthogonality relations among the eigenfunctions, combined with the formulas

$$\| v_{0,m,n} \| = \sqrt{\frac{2\pi}{2m+1}}\; S_{m+1}(\sigma_{m,n}),$$
$$\| v_{k,m,n} \| = \| \widetilde{v}_{k,m,n} \| = \sqrt{\frac{\pi(m+k)!}{(2m+1)(m-k)!}}\; S_{m+1}(\sigma_{m,n}), \qquad k > 0, \tag{12.116}$$

for their norms, to be established in Exercise 12.4.29. In particular, the slowest-decaying mode is the spherically symmetric function

$$u_{0,0,1}(t,r) = \frac{e^{-\pi^2 t}\sin \pi r}{\pi r}\,, \tag{12.117}$$

corresponding to the smallest eigenvalue $\lambda_{0,1} = \sigma_{0,1}^2 = \pi^2$. Therefore, typically, the decay to thermal equilibrium of a unit sphere is at an exponential rate of $\pi^2 \approx 9.8696$, or, to a very rough approximation, 10.

Exercises

12.4.13. It takes a solid ball of radius 1 cm ten minutes to return to (approximate) thermal equilibrium. How long does it take a similar ball of radius 2?

12.4.14. If a 200-gram potato served hot from the oven takes 15 minutes until its maximum temperature is less than 40° C, how long does it take a 300-gram potato of the same shape to cool off?

♡ 12.4.15. A uniform solid metal ball of radius 1 meter, with diffusion coefficient $\gamma = 2$, is taken from a 300° oven and immersed in a bucket of ice water. (*a*) Write down an initial-boundary value problem that describes the temperature of the ball. (*b*) Find a series solution for the temperature. (*c*) At what time is the temperature $\le 50°$ throughout the ball?

♢ 12.4.16. Find the decay rate to thermal equilibrium of a solid spherical ball of radius R and diffusion coefficient γ when subject to homogeneous Dirichlet boundary conditions.

12.4.17. *True or false*: A heated solid hemisphere placed in a 0° environment cools down twice as fast as a solid sphere of the same radius made out of the same material.

12.4.18. A fully insulated solid spherical ball of radius 1 has initial temperature distribution $f(r,\varphi,\theta)$. (*a*) Write down a formula for the equilibrium temperature of the ball. (*b*) What is the rate of decay of the ball to thermal equilibrium?

[†] We use the spherical coordinate form of the L^2 inner product on the ball.

12.4.19. Which cools down to equilibrium faster: a fully insulated solid ball or one whose boundary is held fixed at 0°? How much faster?

12.4.20. A solid sphere and solid cube are made out of the same material and have the same volume. Both are heated in an oven and then submerged in a large vat of water. Which will cool down faster? Explain and justify your answer.

12.4.21. Answer Exercise 12.4.20 when the two solids have the same surface area.

12.4.22. Suppose the solid spherical shell in Exercise 12.2.7 starts off at room temperature. Assuming that the water in the center remains at 100°, find the rate at which the shell tends to thermal equilibrium.

♡ 12.4.23. The thermodynamics of a thin, uniform, spherical shell of unit radius is governed by the *spherical heat equation* $u_t = \gamma\, \Delta_S u$, $u(0, \varphi, \theta) = f(\varphi, \theta)$, in which Δ_S is the spherical Laplacian (12.19). The solution $u(t, \varphi, \theta)$ represents the temperature of the point on the unit sphere with angular coordinates φ, θ, while $f(\varphi, \theta)$ is the initial temperature distribution. (*a*) Find the eigensolutions. (*b*) Write down the solution to the initial value problem as a series in eigensolutions. (*c*) What is the final equilibrium temperature of the spherical shell? (*d*) What is its rate of decay to equilibrium? (*e*) Find the solution and the final equilibrium temperature when $f(\varphi, \theta) = (i)\ \sin\varphi \cos\theta$; $(ii)\ \cos 2\varphi$.

12.4.24. A spherical potato, of radius $R = 7.5\,\text{cm}$ and thermal diffusivity $\gamma = .3\,\text{cm}^2/\text{sec}$, is initially at room temperature, 25° C, and is placed in a pot of boiling water at 100° C. The potato is cooked when it has reached the temperature of at least 90° C throughout. How long do you have to wait until the potato is done?

12.4.25. (*a*) Explain why the spherical Bessel function $S_1(x)$ is bounded at $x = 0$. What is $S_1(0)$? (*b*) Answer the same question for $S_2(x)$.

♢ 12.4.26. Prove the formulae (12.108).

♢ 12.4.27. (*a*) Find a recurrence relation expressing the spherical Bessel function $S_{m-1}(x)$ in terms of $S_m(x)$. (*b*) Prove that

$$\frac{d}{dx}\Big[\, x^3 \Big(S_m(x)^2 - S_{m-1}(x)\, S_{m+1}(x) \Big) \,\Big] = 2\, x^2\, S_m(x)^2.$$

♢ 12.4.28. Let $m \geq 0$ be a fixed integer. (*a*) Prove that the rescaled spherical Bessel functions $v_n(r) = S_m(\sigma_{m,n}\, r)$, $n = 1, 2, \dots$, are mutually orthogonal under the inner product $\langle\, f\,, g\,\rangle = \displaystyle\int_0^1 f(r)\, g(r)\, r^2\, dr$. (*b*) Prove that $\| v_n \| = \frac{1}{\sqrt{2}}\, |\, S_{m+1}(\sigma_{m,n})\,|$. *Hint*: Mimic the method outlined in Exercise 11.4.22, using the identity in Exercise 12.4.27(*b*).

♢ 12.4.29. (*a*) Use the result of Exercise 12.4.28 to prove the formulae (12.116) for the L^2 norms of the eigenfunctions (12.110). (*b*) Justify the formulae (12.115).

The Fundamental Solution of the Heat Equation

For the heat equation (as well as more general diffusion equations), the fundamental solution measures the response of the body to an instantaneously applied concentrated unit heat source. Thus, given a point $\boldsymbol{\xi} = (\xi, \eta, \zeta) \in \Omega$ within the body, the *fundamental solution*

$$u(t, \mathbf{x}) = F(t, \mathbf{x}; \boldsymbol{\xi}) = F(t, x, y, z; \xi, \eta, \zeta)$$

solves the initial-boundary value problem

$$u_t = \Delta u, \qquad u(0, \mathbf{x}) = \delta(\mathbf{x} - \boldsymbol{\xi}), \qquad \text{for} \qquad \mathbf{x} \in \Omega, \quad t > 0, \tag{12.118}$$

subject to the selected homogeneous boundary conditions — Dirichlet, Neumann, or mixed.

Explicit formulas for the fundamental solution are rare, although in bounded domains it is possible to construct it as an eigenfunction series, as described in Section 9.5. The one case amenable to a complete analysis is that in which the heat is distributed over all of three-dimensional space, so $\Omega = \mathbb{R}^3$. We recall that Lemma 11.11 showed how to construct solutions of the two-dimensional heat equation as products of one-dimensional solutions. In a similar manner, if $p(t,x)$, $q(t,x)$, and $r(t,x)$ are any three solutions to the one-dimensional heat equation $u_t = \gamma\, u_{xx}$, then their product

$$u(t,x,y,z) = p(t,x)\, q(t,y)\, r(t,z) \tag{12.119}$$

is a solution to the three-dimensional heat equation

$$u_t = \gamma\, (u_{xx} + u_{yy} + u_{zz}).$$

In particular, choosing

$$p(t,x) = \frac{e^{-(x-\xi)^2/4\gamma t}}{2\sqrt{\pi\gamma t}}, \qquad q(t,y) = \frac{e^{-(y-\eta)^2/4\gamma t}}{2\sqrt{\pi\gamma t}}, \qquad r(t,z) = \frac{e^{-(z-\zeta)^2/4\gamma t}}{2\sqrt{\pi\gamma t}},$$

to all be one-dimensional fundamental solutions, we are immediately led to the fundamental solution in the form of a three-dimensional *Gaussian filter*.

Theorem 12.14. *The fundamental solution*

$$F(t,\mathbf{x};\boldsymbol{\xi}) = F(t,\mathbf{x}-\boldsymbol{\xi}) = \frac{e^{-\|\,\mathbf{x}-\boldsymbol{\xi}\,\|^2/(4\gamma t)}}{8\,(\pi\gamma t)^{3/2}} \tag{12.120}$$

solves the three-dimensional heat equation $u_t = \gamma\,\Delta u$ on $\mathbb{R}^3$ for $t > 0$, with an initial temperature equal to a delta function concentrated at the point $\mathbf{x} = \boldsymbol{\xi}$.

Thus, the initially concentrated heat energy immediately begins to spread out in a spherically symmetric manner, with a minuscule, but nonzero effect that is felt immediately arbitrarily far away from the initial concentration. At each individual point $\mathbf{x} \in \mathbb{R}^3$, after an initial warm-up, the temperature decays back to zero at a rate proportional to $t^{-3/2}$ — more rapidly than in two dimensions, because, intuitively, there are more directions in which the heat energy can disperse.

To solve a more general initial value problem with the initial temperature distributed over all of space, we first write

$$u(0,\mathbf{x}) = f(\mathbf{x}) = \iiint f(\boldsymbol{\xi})\,\delta(\mathbf{x}-\boldsymbol{\xi})\,d\xi\,d\eta\,d\zeta$$

as a linear superposition of delta functions. By linearity, the solution to the initial value problem is given by the corresponding superposition

$$u(t,\mathbf{x}) = \frac{1}{8\,(\pi\gamma t)^{3/2}} \iiint f(\boldsymbol{\xi})\, e^{-\|\,\mathbf{x}-\boldsymbol{\xi}\,\|^2/(4\gamma t)}\,d\xi\,d\eta\,d\zeta \tag{12.121}$$

of the fundamental solutions. Since the fundamental solution has exponential decay as $\|\,\mathbf{x}\,\| \to \infty$, the superposition formula is valid even for initial temperature distributions that are moderately increasing at large distances. We remark that the integral (12.121) has the form of a three-dimensional convolution

$$u(t,\mathbf{x}) = F(t,\mathbf{x}) * f(\mathbf{x}) = \iiint f(\boldsymbol{\xi})\,F(t,\mathbf{x}-\boldsymbol{\xi})\,d\xi\,d\eta\,d\zeta \tag{12.122}$$

of the initial data with a one-parameter family of increasingly spread-out Gaussian filters. Thus, as before, convolution with a Gaussian filter has a smoothing effect on the initial temperature distribution.

Exercises

12.4.30. *True or false*: In a three-dimensional medium, heat energy propagates at infinite speed.

12.4.31. A solid spherical ball of radius 1 is heated to 100° and inserted into a three-dimensional medium filling the rest of $\mathbb{R}^3$ with uniform temperature 0°.
(*a*) Write down an integral formula for the subsequent temperature distribution over $\mathbb{R}^3$ at time $t > 0$, assuming a common diffusion coefficient $\gamma = 1$.
(*b*) Evaluate the resulting integral using spherical coordinates.

12.4.32. (a) Prove that $u(t, r)$ is a spherically symmetric solution to the three-dimensional heat equation if and only if $w(t, r) = r\, u(t, r)$ solves the one-dimensional heat equation: $w_t = w_{rr}$. (*b*) *True or false*: If $w(t, r)$ is the fundamental solution for the one-dimensional heat equation based at $r = 0$, then $u(t, r) = w(t, r)/r$ is the fundamental solution for the three-dimensional heat equation based at the origin.

12.4.33. Construct the solution to the initial value problem in Exercise 12.4.31 using radial symmetry and Exercise 12.4.32.

♡ 12.4.34. Suppose that, as Earth orbits the sun, its surface is subject to yearly periodic temperature variations $a \cos \omega\, t$, where the frequency ω is given by (4.56). (*a*) Assuming, for simplicity, that the Earth is a homogeneous solid ball, of radius R, formulate an initial-boundary value problem that governs the temperature fluctuations within the Earth due to its orbiting the sun. (*b*) At what depth does the temperature vary out of phase with the surface, i.e., is the warmest in winter and coldest in summer? Compare your answer with the root cellar computation at the end of Section 4.1. *Hint*: Use Exercise 12.4.32.

12.4.35. (a) Prove that if $u(t, x)$ is any (sufficiently smooth) solution to the heat equation, so is its time derivative $v = \partial u/\partial t$. (*b*) Write out the time derivative of the fundamental solution, and the initial value problem it satisfies.

12.4.36. Write down an explicit eigenfunction series for the fundamental solution $F(t, \mathbf{x}; \boldsymbol{\xi})$ to the heat equation in a unit cube with thermal diffusivity $\gamma = 1$ that is subject to homogeneous Dirichlet boundary conditions.

12.4.37. Write down an explicit eigenfunction series for the fundamental solution $F(t, \mathbf{x}; \boldsymbol{\xi})$ to the heat equation in a ball of radius 1 that has thermal diffusivity $\gamma = 1$ and is subject to homogeneous Dirichlet boundary conditions.

♢ 12.4.38. Justify the statement that formula (12.119) provides a solution to the three-dimensional heat equation.

12.4.39. Fill in the details of the proof of Theorem 12.14.

12.5 The Wave Equation for Three–Dimensional Media

The *three-dimensional wave equation*

$$u_{tt} = c^2 \Delta u = c^2 (u_{xx} + u_{yy} + u_{zz}), \tag{12.123}$$

in which $c > 0$ denotes the speed of light, governs the propagation of waves in a homogeneous isotropic three-dimensional medium, e.g., electromagnetic waves (light, X-rays, radio waves, etc.) in empty space. In this context, while the electric and magnetic vector fields $\mathbf{E}, \mathbf{B}$ are intrinsically coupled by the more complicated system of Maxwell's equations, each individual component satisfies the wave equation; see Exercise 12.5.14 for details.

The wave equation also models certain restricted classes of vibrations of a uniform solid body. The solution $u(t, \mathbf{x}) = u(t, x, y, z)$ represents a scalar-valued displacement of the body at time t and position $\mathbf{x} = (x, y, z) \in \Omega \subset \mathbb{R}^3$. For example, $u(t, \mathbf{x})$ might represent the radial displacement of the body. One imposes suitable boundary conditions, e.g., Dirichlet, Neumann, or mixed, on $\partial\Omega$, along with a pair of initial conditions

$$u(0, \mathbf{x}) = f(\mathbf{x}), \qquad \frac{\partial u}{\partial t}(0, \mathbf{x}) = g(\mathbf{x}), \qquad \mathbf{x} \in \Omega, \tag{12.124}$$

that specify the body's initial displacement and initial velocity. As long as the initial and boundary data are reasonably nice, there exists a unique classical solution to the initial-boundary value problem for all $-\infty < t < \infty$, cf. [**38**, **61**, **99**]. Thus, in contrast to the heat equation, one can follow solutions to the wave equation both forwards and backwards in time.

Let us focus our attention on the homogeneous boundary value problem. The fundamental vibrational modes are found by imposing our usual trigonometric ansatz

$$u(t, x, y, z) = \cos(\omega t)\, v(x, y, z) \qquad \text{or} \qquad \sin(\omega t)\, v(x, y, z).$$

Substituting into the wave equation (12.123), we discover (yet again) that $v(x, y, z)$ must be an eigenfunction for the associated Helmholtz eigenvalue problem

$$\Delta v + \lambda\, v = 0, \qquad \text{where} \qquad \lambda = \frac{\omega^2}{c^2}\,, \tag{12.125}$$

coupled to the relevant boundary conditions. In the positive definite cases, i.e., Dirichlet and mixed boundary conditions, the eigenvalues $\lambda_k = \omega_k^2/c^2 > 0$ are all positive. Each eigenfunction $v_k(x, y, z)$ yields two normal vibrational modes

$$u_k(t, x, y, z) = \cos(\omega_k t)\, v_k(x, y, z), \qquad \widetilde{u}_k(t, x, y, z) = \sin(\omega_k t)\, v_k(x, y, z),$$

of frequency $\omega_k = c\sqrt{\lambda_k}$ equal to the square root of the corresponding eigenvalue multiplied by the wave speed. The general solution is a quasiperiodic linear combination,

$$u(t, x, y, z) = \sum_{k=1}^{\infty} \big[\, a_k \cos(\omega_k t) + b_k \sin(\omega_k t) \,\big]\, v_k(x, y, z), \tag{12.126}$$

of the eigenmodes. The coefficients a_k, b_k are uniquely prescribed by the initial conditions (12.124). Thus,

$$u(0, x, y, z) = \sum_{k=1}^{\infty} a_k\, v_k(x, y, z) = f(x, y, z),$$

$$\frac{\partial u}{\partial t}(0, x, y, z) = \sum_{k=1}^{\infty} \omega_k\, b_k\, v_k(x, y, z) = g(x, y, z).$$

The explicit formulas follow immediately from the orthogonality of the eigenfunctions:

$$a_k = \frac{\langle f, v_k \rangle}{\| v_k \|^2} = \frac{\iiint_\Omega f\, v_k \, dx\, dy\, dz}{\iiint_\Omega v_k^2 \, dx\, dy\, dz}, \qquad b_k = \frac{1}{\omega_k} \frac{\langle g, v_k \rangle}{\| v_k \|^2} = \frac{\iiint_\Omega g\, v_k \, dx\, dy\, dz}{\omega_k \iiint_\Omega v_k^2 \, dx\, dy\, dz}. \tag{12.127}$$

In the positive semi-definite Neumann case, there is an additional zero eigenvalue $\lambda_0 = 0$ corresponding to the constant null eigenfunction $v_0(x, y, z) \equiv 1$. This results in two additional terms in the eigenfunction expansion — a constant term

$$a_0 = \frac{1}{\operatorname{vol} \Omega} \iiint_\Omega f(x, y, z)\, dx\, dy\, dz$$

that equals the average initial displacement, and an unstable mode $b_0\, t$ that grows linearly in time, whose speed

$$b_0 = \frac{1}{\operatorname{vol} \Omega} \iiint_\Omega g(x, y, z)\, dx\, dy\, dz$$

is the average initial velocity over the entire body. Thus, the unstable mode will be excited if and only if there is a nonzero net initial velocity: $b_0 \neq 0$.

Most of the basic solution techniques we learned in the two-dimensional case apply here, and we will not dwell on the details. The case of a rectangular box is a particularly straightforward application of the method of separation of variables, and is outlined in the exercises. A similar analysis, now in cylindrical coordinates, can be applied to the case of a vibrating cylinder. The most interesting case is that of a solid spherical ball, which is the subject of the next subsection.

Vibration of Balls and Spheres

Let us focus on the radial vibrations of a solid ball, as modeled by the three-dimensional wave equation (12.123). The solution $u(t, x, y, z)$ represents the radial displacement of the "atom" that is situated at position (x, y, z) when the ball is at rest.

For simplicity, we look at the Dirichlet boundary value problem on the unit ball $B_1 = \{ \| \mathbf{x} \| < 1 \}$. The normal modes of vibration are governed by the Helmholtz equation (12.125) subject to homogeneous Dirichlet boundary conditions. According to (12.110), the eigenfunctions are

$$\begin{aligned} v_{0,m,n}(r, \varphi, \theta) &= S_m(\sigma_{m,n}\, r)\, Y_m^0(\varphi, \theta), & & n = 1, 2, 3, \dots, \\ v_{k,m,n}(r, \varphi, \theta) &= S_n(\sigma_{n,m}\, r)\, Y_m^k(\varphi, \theta), & \text{for} \quad & m = 0, 1, 2, \dots, \\ \widetilde{v}_{k,m,n}(r, \varphi, \theta) &= S_m(\sigma_{m,n}\, r)\, \widetilde{Y}_m^k(\varphi, \theta), & & k = 1, 2, \dots, m. \end{aligned} \tag{12.128}$$

Here S_m denotes the m^{th} order spherical Bessel function (12.104), $\sigma_{m,n}$ is its n^{th} root, as in (12.109), while $Y_n^m, \widetilde{Y}_n^m$ are the spherical harmonics (12.38). Each eigenvalue

$$\lambda_{m,n} = \sigma_{m,n}^2, \qquad m = 0, 1, 2, \dots, \qquad n = 1, 2, 3, \dots,$$

corresponds to $2\,m + 1$ independent eigenfunctions, namely

$$v_{k,m,0}(r, \varphi, \theta),\ v_{k,m,1}(r, \varphi, \theta),\ \dots,\ v_{k,m,m}(r, \varphi, \theta),\ \widetilde{v}_{k,m,1}(r, \varphi, \theta),\ \dots,\ \widetilde{v}_{k,m,m}(r, \varphi, \theta).$$

Consequently, the fundamental vibrational frequencies of a solid ball

$$\omega_{m,n} = c\sqrt{\lambda_{m,n}} = c\,\sigma_{m,n}, \qquad m = 0, 1, 2, \ldots, \qquad n = 1, 2, 3, \ldots, \tag{12.129}$$

are equal to the spherical Bessel roots $\sigma_{m,n}$ multiplied by the wave speed. There is a total of $2(2m+1)$ independent vibrational modes associated with each distinct frequency (12.129), namely

$$\begin{aligned}
u_{0,m,n}(t,r,\varphi,\theta) &= \cos(c\,\sigma_{m,n}t)\; S_m(\sigma_{m,n}\,r)\; Y_m^0(\varphi,\theta),\\
\widehat{u}_{0,m,n}(t,r,\varphi,\theta) &= \sin(c\,\sigma_{m,n}t)\; S_m(\sigma_{m,n}\,r)\; Y_m^0(\varphi,\theta),\\
u_{k,m,n}(t,r,\varphi,\theta) &= \cos(c\,\sigma_{m,n}t)\; S_m(\sigma_{m,n}\,r)\; Y_m^k(\varphi,\theta),\\
\widehat{u}_{k,m,n}(t,r,\varphi,\theta) &= \sin(c\,\sigma_{m,n}t)\; S_m(\sigma_{m,n}\,r)\; Y_m^k(\varphi,\theta),\\
\widetilde{u}_{k,m,n}(t,r,\varphi,\theta) &= \cos(c\,\sigma_{m,n}t)\; S_m(\sigma_{m,n}\,r)\; \widetilde{Y}_m^k(\varphi,\theta),\\
\widehat{\widetilde{u}}_{k,m,n}(t,r,\varphi,\theta) &= \sin(c\,\sigma_{m,n}t)\; S_m(\sigma_{m,n}\,r)\; \widetilde{Y}_m^k(\varphi,\theta),
\end{aligned} \qquad \begin{array}{l} n = 1, 2, 3, \ldots\,,\\ m = 0, 1, 2, \ldots\,,\\ k = 1, 2, \ldots, m. \end{array} \tag{12.130}$$

In particular, the radially symmetric modes of vibration have, according to (12.105), the elementary form

$$\begin{aligned}
u_{0,0,n}(r,\varphi,\theta) &= \cos(c\,n\,\pi\,t)\; S_0(n\,\pi\,r) = \frac{\cos c\,n\,\pi\,t\;\sin n\,\pi\,r}{r},\\
\widehat{u}_{0,0,n}(r,\varphi,\theta) &= \sin(c\,n\,\pi\,t)\; S_0(n\,\pi\,r) = \frac{\sin c\,n\,\pi\,t\;\sin n\,\pi\,r}{r},
\end{aligned} \qquad n = 1, 2, 3, \ldots\,. \tag{12.131}$$

Their vibrational frequencies, $\omega_{0,n} = c\,n\,\pi$, are integral multiples of the lowest frequency $\omega_{0,1} = c\,\pi$. Thus, intriguingly, if you excite only the radially symmetric modes, the resulting motion of the ball is periodic. However, more general vibrations are only quasiperiodic.

Adopting the same scaling argument as in (11.166), we conclude that the fundamental frequencies for a solid ball of radius R and wave speed c are given by $\omega_{m,n} = c\,\sigma_{m,n}/R$. The relative vibrational frequencies

$$\frac{\omega_{m,n}}{\omega_{0,1}} = \frac{\sigma_{m,n}}{\sigma_{0,1}} = \frac{\sigma_{m,n}}{\pi} \tag{12.132}$$

are independent of the size of the ball R or the wave speed c. In the accompanying table, we display all relative vibrational frequencies that are less than 4 in magnitude.

Relative Spherical Bessel Roots $\sigma_{m,n}/\sigma_{0,1}$

n \ m	0	1	2	3	4	6	7	8	…
1	1.0000	1.4303	1.8346	2.2243	2.6046	2.9780	3.3463	3.7105	…
2	2.0000	2.4590	2.8950	3.3159	3.7258	⋮	⋮	⋮	
3	3.0000	3.4709	3.9225	⋮	⋮				
4	4.0000	⋮	⋮						
⋮	⋮								

The purely radial modes of vibration (12.131) have individual frequencies

$$\omega_{0,n} = \frac{n\,\pi\,c}{R}\,, \qquad \text{so} \qquad \frac{\omega_{0,n}}{\omega_{0,1}} = n,$$

which appear in the first column of the table. The lowest frequency is $\omega_{0,1} = \pi\,c/R$, corresponding to a vibration with period $2\,\pi/\omega_{0,1} = 2\,R/c$. In particular, for the Earth, the radius $R \approx 6000$ km, and the wave speed in rock is, on average, $c \approx 5$ km/sec, so that the fundamental mode of vibration has period $2\,R/c \approx 2400$ seconds, or 40 minutes. Of course, we have suppressed almost all interesting terrestrial geology in this very crude approximation, which has been based on the assumption that the Earth is a uniform spherical body, globally vibrating only in its radial direction. A more realistic modeling of the vibrations of the Earth requires an understanding of the basic partial differential equations of linear and nonlinear elastodynamics, [**7**, **49**]. Nonuniformities in the Earth lead to scattering of the vibrational waves, which are then used to locate subterranean geological structures, e.g., oil and gas deposits. Localized vibrations of the Earth are also known as *seismic waves*, and, of course, earthquakes are their most severe manifestation. We refer the interested reader to [**5**] for an introduction to mathematical seismology. Understanding terrestrial vibrations is an issue of critical importance in geophysics and civil engineering, including the design of structures, buildings, and bridges, requiring the avoidance of potentially catastrophic resonant frequencies.

Example 12.15. The radial vibrations of a hollow thin spherical shell (e.g., an elastic balloon) are governed by the differential equation

$$\frac{\partial^2 u}{\partial t^2} = c^2\,\Delta_S[u] = c^2\left(\frac{\partial^2 u}{\partial \varphi^2} + \frac{\cos\varphi}{\sin\varphi}\,\frac{\partial u}{\partial \varphi} + \frac{1}{\sin^2\varphi}\,\frac{\partial^2 u}{\partial \theta^2}\right), \tag{12.133}$$

where Δ_S denotes the spherical Laplacian (12.19). The radial displacement $u(t,\varphi,\theta)$ of a point on the sphere depends only on time t and the angular coordinates φ, θ. The solution $u(t,\varphi,\theta)$ is required to be $2\,\pi$–periodic in the azimuthal angle θ and bounded at the poles, where $\varphi = 0$ and π.

According to (12.38), the n^{th} eigenvalue of the spherical Laplacian, $\lambda_n = n(n+1)$, possesses $2\,n+1$ linearly independent eigenfunctions, namely, the spherical harmonics

$$Y_n^0(\varphi,\theta), \quad Y_n^1(\varphi,\theta), \quad \dots, \quad Y_n^n(\varphi,\theta), \quad \widetilde{Y}_n^1(\varphi,\theta), \quad \dots, \quad \widetilde{Y}_n^n(\varphi,\theta).$$

As a consequence, the fundamental frequencies of vibration for a spherical shell are

$$\omega_n = c\,\sqrt{\lambda_n} = c\,\sqrt{n(n+1)}\,, \qquad n = 1, 2, \dots\,. \tag{12.134}$$

The vibrational solutions are quasiperiodic combinations of the fundamental spherical harmonic modes

$$\begin{aligned} &\cos\bigl(\sqrt{n(n+1)}\,t\,\bigr)\; Y_n^m(\varphi,\theta), \qquad &&\sin\bigl(\sqrt{n(n+1)}\,t\,\bigr)\; Y_n^m(\varphi,\theta),\\ &\cos\bigl(\sqrt{n(n+1)}\,t\,\bigr)\; \widetilde{Y}_n^m(\varphi,\theta), \qquad &&\sin\bigl(\sqrt{n(n+1)}\,t\,\bigr)\; \widetilde{Y}_n^m(\varphi,\theta). \end{aligned} \tag{12.135}$$

Representative graphs can be seen in Figure 12.5. The smallest positive eigenvalue is $\lambda_1 = 2$, yielding a lowest tone of frequency $\omega_1 = c\,\sqrt{2}$. The higher-order frequencies are irrational multiples of the fundamental frequency, implying that a vibrating spherical bell sounds dissonant to our ears.

One further remark is in order. The spherical Laplacian operator is only positive semi-definite, since the lowest mode has eigenvalue $\lambda_0 = 0$, which corresponds to the constant

null eigenfunction $v_0(\varphi, \theta) = Y_0^0(\varphi, \theta) \equiv 1$. Therefore, the wave equation (12.133) admits an unstable mode $b_{0,0}\, t$, corresponding to a uniform radial inflation; its coefficient

$$b_{0,0} = \frac{3}{4\pi} \iint_{S_1} \frac{\partial u}{\partial t}\,(0, \varphi, \theta)\, dS$$

represents the shell's average initial velocity. The existence of such an unstable mode is an artifact of the simplified linear model we are using, which fails to account for nonlinearly elastic effects that serve to constrain the inflation of a spherical balloon.

Exercises

12.5.1. Find the eigenfunction series solution to the initial-boundary value problem for the wave equation $u_{tt} = \Delta u$ on a unit cube $C = \{0 < x, y, z < 1\}$, subject to homogeneous Dirichlet boundary conditions and one of the following sets of initial conditions:
(*a*) $u(0, x, y, z) = 1$, $u_t(0, x, y, z) = 0$; (*b*) $u(0, x, y, z) = 0$, $u_t(0, x, y, z) = 1$;
(*c*) $u(0, x, y, z) = \sin \pi x \, \sin \pi y \, \sin \pi z$, $u_t(0, x, y, z) = 0$; (*d*) $u(0, x, y, z) = \sin 3\pi x$, $u_t(0, x, y, z) = \sin 2\pi y$; (*e*) $u(0, x, y, z) = 0$, $u_t(0, x, y, z) = x\, y\, z\, (1 - x)(1 - y)(1 - z)$.

12.5.2. Suppose the cube in Exercise 12.5.1 is subject to homogeneous Neumann boundary conditions. Which of the preceding initial value problems leads to an unstable motion of the cube?

12.5.3. (*a*) Find the separable periodic vibrations of a unit cube subject to homogeneous Dirichlet boundary conditions. (*b*) Can you find a periodic mode that is not separable?

12.5.4. Answer Exercise 12.5.3 when one face of the cube is left free, while the other five faces are fixed.

12.5.5. Given a material with wave speed $c = 1.5$ cm/sec, find the natural vibrational frequencies of a solid rectangular box of size 1 cm $\times$ 2 cm $\times$ 3 cm whose sides are held fixed. List the lowest five such frequencies in order. Does the box vibrate periodically?

12.5.6. Find the natural vibrational frequencies of a solid cylinder of height 2, radius 1, and wave speed $c = 1$, when (*a*) all sides are fixed; (*b*) the top and bottom of the cylinder are free, while the curved side is fixed; (*c*) the curved side of the cylinder is free, while the top and bottom are fixed.

12.5.7. Among all solid cylinders of unit volume with fixed boundary, find the one that vibrates the slowest.

12.5.8. Does a solid spherical ball that is subject to homogeneous Neumann boundary conditions vibrate (*i*) faster, (*ii*) slower, or (*iii*) at the same rate as the same ball subject to homogeneous Dirichlet conditions. If your answer is (*i*) or (*ii*), estimate how much faster or slower.

12.5.9. A solid cube and solid sphere are made of the same material and have the same volume. Which vibrates faster when subject to homogeneous Dirichlet boundary conditions?

12.5.10. Assuming that they both have the same wave speed and fixed boundaries, which vibrates faster: a solid sphere or a circular membrane of the same radius?

12.5.11. A uniform, solid spherical planet is floating freely in outer space. Find its three slowest resonant frequencies.

12.5.12. *True or false*: Suppose we have two uniform solid bodies composed of the same material. If the first body cools down to thermal equilibrium the fastest, then it also vibrates the fastest. Explain your answer.

12.5.13. (*a*) Define what is meant by a nodal curve and a nodal region on a vibrating thin spherical shell. (*b*) *True or false*: All the nodal curves are arcs of circles.

♡ 12.5.14. The propagation of electromagnetic waves (including light) is governed by the electric field $\mathbf{E}(t, \mathbf{x})$ and magnetic field $\mathbf{B}(t, \mathbf{x})$, which are both time-dependent vector fields defined for $\mathbf{x} = (x, y, z)$ in a domain $\Omega \subset \mathbb{R}^3$. In empty space, *Maxwell's equations* (as formulated by Heaviside) are

$$\nabla \cdot \mathbf{E} = 0, \qquad \nabla \cdot \mathbf{B} = 0, \qquad \frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \mathbf{E}, \qquad \frac{\partial \mathbf{E}}{\partial t} = \frac{1}{\mu_0\, \epsilon_0} \nabla \times \mathbf{B}, \tag{12.136}$$

where μ_0, ϵ_0 are, respectively, the *permeability* and *permittivity* constants. Prove that all individual components of $\mathbf{E}$ and $\mathbf{B}$ satisfy the scalar wave equation. What is the wave speed, i.e., the speed of light in empty space?

12.6 Spherical Waves and Huygens' Principle

For any dynamical partial differential equation, the fundamental solution measures the effect of applying an instantaneous concentrated unit impulse at a single point. Two representative physical effects to keep in mind are the light waves emanating from a sudden concentrated blast, e.g., a lightning bolt or a stellar supernova, and the sound waves due to an explosion or thunderclap, propagating in air at a much slower speed. Linear superposition utilizes the fundamental solution to build up more general solutions to initial value problems. For the wave and other second-order vibrational equations, the impulse can be applied either to the initial displacement or to the initial velocity, resulting in two distinct types of fundamental solution. The general solution to the initial value problem will be obtained by a double superposition. In this section, we derive explicit formulas for the two fundamental solutions for the three-dimensional wave equation on all of space, leading to Kirchhoff's formula for the solution to the general initial value problem. An important consequence is Huygens' Principle, which states that, in three-dimensional space, localized initial disturbances remain localized as they propagate. In the final subsection, we apply the method of descent to our three-dimensional solution formulas in order to solve the two-dimensional wave equation, for which, surprisingly, Huygens' Principle is no longer valid.

Spherical Waves

In a uniform isotropic medium, an initial concentrated blast results in a spherically expanding wave, moving away at the speed of light (or sound) in all directions. Invoking translation invariance, we will assume, without loss of generality, that the source of the disturbance is at the origin, and so the solution $u(t, \mathbf{x})$ should depend only on the distance $r = \| \mathbf{x} \|$ from the source. We adopt spherical coordinates and look for a solution $u = u(t, r)$ to the three-dimensional wave equation (12.123) with no angular dependence. Substituting the formula (12.16) for the spherical Laplacian and setting both angular derivatives to 0, we are led to the partial differential equation

$$\frac{\partial^2 u}{\partial t^2} = c^2 \left(\frac{\partial^2 u}{\partial r^2} + \frac{2}{r} \frac{\partial u}{\partial r} \right), \tag{12.137}$$

which governs the propagation of spherically symmetric waves in three-dimensional space. Surprisingly, we can explicitly solve (12.137). The secret is to multiply both sides of the equation by r:

$$\frac{\partial^2 (r\,u)}{\partial t^2} = r\,\frac{\partial^2 u}{\partial t^2} = c^2 \left(r\,\frac{\partial^2 u}{\partial r^2} + 2\,\frac{\partial u}{\partial r} \right) = c^2\,\frac{\partial^2}{\partial r^2}\,(r\,u).$$

Thus, the function

$$w(t,r) = r\,u(t,r)$$

satisfies the one-dimensional wave equation

$$\frac{\partial^2 w}{\partial t^2} = c^2\,\frac{\partial^2 w}{\partial r^2}\,. \tag{12.138}$$

According to Theorem 2.14, the general solution to the one-dimensional wave equation (12.138) can be written in d'Alembert form

$$w(t,r) = p(r - ct) + q(r + ct),$$

where $p(\xi)$ and $q(\eta)$ are arbitrary functions of a single characteristic variable. Therefore, spherically symmetric solutions to the three-dimensional wave equation assume the form

$$u(t,r) \;=\; \frac{p(r-ct)}{r} \;+\; \frac{q(r+ct)}{r}\,. \tag{12.139}$$

The first summand,

$$u(t,r) = \frac{p(r-ct)}{r}\,, \tag{12.140}$$

represents a wave moving at speed c in the direction of increasing r, and so describes the illumination from a variable light source that is concentrated at the origin, e.g., a pulsating quasar in interstellar space. To highlight this interpretation, let us concentrate on the case that $p(\xi) = \delta(\xi - a)$ is a delta function, keeping in mind that more general solutions can then be assembled by linear superposition. The induced solution

$$u(t,r) = \frac{\delta(r - ct - a)}{r} = \frac{\delta\big(r - c\,(t - t_0)\big)}{r}\,, \qquad \text{where} \qquad t_0 = -\,\frac{a}{c}\,, \tag{12.141}$$

represents a spherical wave propagating through space. At the instant $t = t_0$, the light is entirely concentrated at the origin, $r = 0$. The signal then moves away from the origin in all directions at speed c. At each later time $t > t_0$, the wave remains concentrated on the surface of a sphere of radius $r = c\,(t - t_0)$. Its intensity at each point on the sphere, however, has decreased by a factor $1/r$, and so, the farther the light travels away from the source, the dimmer it becomes. A stationary observer sitting at a fixed point in space will see only an instantaneous flash of light of intensity $1/r$ as the spherical wave passes by at time $t = t_0 + r/c$, where r is the observer's distance from the light source. A similar statement holds for sound waves — to an observer, the sound of a distant explosion will last momentarily. Thunder and lightning are the most familiar examples of this everyday phenomenon.

On the other hand, for $t < t_0$, the impulse is concentrated at a negative radius $r = c\,(t - t_0) < 0$. To interpret this, note that, for spherical coordinates (12.15), replacing r by $-r$ has the same effect as changing $\mathbf{x}$ to the antipodal point $-\mathbf{x}$. Thus, the solution (12.141) represents a concentrated spherically symmetric light wave arriving from the edges

of the universe at speed c that strengthens in intensity as it collapses into the origin at $t = t_0$. After collapse, it immediately reappears and expands back out into the universe.

The second solution in the d'Alembert formula (12.139) has, in fact, exactly the same physical form under the antipodal identification. Indeed, if we set

$$\widetilde{r} = -r, \qquad \widetilde{p}(\xi) = -q(-\xi), \qquad \text{then} \qquad \frac{q(r+ct)}{r} = \frac{\widetilde{p}(\widetilde{r}-ct)}{\widetilde{r}}\,.$$

Thus, the second d'Alembert solution is redundant, and we only need to consider solutions of the form (12.140) from now on.

To effectively utilize such spherical wave solutions, we need to understand the nature of their originating singularity. For simplicity, we set $t_0 = 0$ in (12.141) and concentrate on the particular solution

$$u(t,r) = \frac{\delta(r-ct)}{r}\,, \tag{12.142}$$

which apparently has a bad singularity at the origin, $r = 0$, at the initial time $t = 0$. We need to pin down precisely which sort of distribution (generalized function) it represents. Invoking the limiting definition is tricky, and it will be easier to work with the dual characterization of a distribution as a linear functional. Thus, at a fixed time $t \geq 0$, we must evaluate the inner product[†]

$$\langle\, u(t,\cdot)\,, f\,\rangle = \iiint u(t,x,y,z)\, f(x,y,z)\, dx\, dy\, dz$$

of the solution with a smooth test function $f(\mathbf{x}) = f(x,y,z)$. We rewrite the triple integral in spherical coordinates, whereby

$$\langle\, u(t,\cdot)\,, f\,\rangle = \int_{-\pi}^{\pi}\int_{0}^{\pi}\int_{0}^{\infty} \frac{\delta(r-ct)}{r}\, f(r,\varphi,\theta)\, r^2 \sin\varphi\; dr\, d\varphi\, d\theta.$$

When $t \neq 0$, the r integration can be immediately evaluated, and so

$$\langle\, u(t,\cdot)\,, f\,\rangle = c\,t \int_{-\pi}^{\pi}\int_{0}^{\pi} f(c\,t,\varphi,\theta)\, \sin\varphi\; d\varphi\, d\theta = 4\,\pi\, c\,t\, \mathrm{M}_{c\,t}\,[\,f\,]\,, \tag{12.143}$$

where

$$\mathrm{M}_{c\,t}\,[\,f\,] = \frac{1}{4\,\pi}\int_{-\pi}^{\pi}\int_{0}^{\pi} f(c\,t,\varphi,\theta)\, \sin\varphi\; d\varphi\, d\theta \;=\; \frac{1}{4\,\pi\, c^2\, t^2}\iint_{S_{c\,t}} f\; dS \tag{12.144}$$

is the *mean* or *average value* of the function f on the sphere $S_{c\,t} = \big\{\, \|\,\mathbf{x}\,\| = c\,t\,\big\}$ of radius $r = c\,t$ and, hence, surface area $4\,\pi\, c^2\, t^2$. In particular, in the limit as the sphere's radius $c\,t \to 0$, by continuity, the mean reduces to just the value of the function at the origin:

$$\lim_{t\,\to\,\infty} \mathrm{M}_{c\,t}\,[\,f\,] = \mathrm{M}_0\,[\,f\,] = f(\mathbf{0}). \tag{12.145}$$

Thus, (12.143) implies that

$$\lim_{t\,\to\,\infty} \langle\, u(t,\cdot)\,, f\,\rangle = \langle\, u(0,\cdot)\,, f\,\rangle = 0 \qquad \text{for } \textit{all} \text{ functions} \quad f,$$

[†] For fixed t, we use $u(t,\cdot)$ to indicate the real-valued function $(x,y,z) \mapsto u(t,x,y,z)$ on $\mathbb{R}^3$.

and hence $u(0, x, y, z) \equiv 0$ represents a zero initial displacement. In other words, there is, in fact, *no singularity* in the solution at $t = 0$!

In the absence of any initial displacement, how, then, can the solution (12.142) be nonzero? Clearly, this must be the result of a nonzero initial velocity. To evaluate $\partial u/\partial t$, we differentiate (12.143), whereby

$$\begin{aligned}\left\langle \frac{\partial u}{\partial t}\,,\, f \right\rangle &= \frac{\partial}{\partial t}\,\langle\, u(t,\cdot)\,, f\,\rangle = \frac{\partial}{\partial t}\left(c\,t \int_{-\pi}^{\pi}\int_{0}^{\pi} f(c\,t,\varphi,\theta)\,\sin\varphi\;d\varphi\,d\theta\right)\\ &= c\int_{-\pi}^{\pi}\int_{0}^{\pi} f(c\,t,\varphi,\theta)\,\sin\varphi\;d\varphi\,d\theta + c^2\,t\int_{-\pi}^{\pi}\int_{0}^{\pi}\frac{\partial f}{\partial r}\,(c\,t,\varphi,\theta)\,\sin\varphi\;d\varphi\,d\theta\\ &= 4\,\pi\,c\,\mathrm{M}_{c\,t}\,[\,f\,] + 4\,\pi\,c^2\,t\,\mathrm{M}_{c\,t}\left[\,\frac{\partial f}{\partial r}\,\right]. \end{aligned} \tag{12.146}$$

The result is a linear combination of the means of f and its radial derivative f_r over the sphere of radius $c\,t$. In the limit, the second term goes to 0, and so, by (12.145),

$$\lim_{t\,\to\,0}\,\langle\, u_t\,, f\,\rangle = 4\,\pi\,c\,\mathrm{M}_0\,[\,f\,] = 4\,\pi\,c\,f(\mathbf{0}).$$

Since this holds for all test functions f, we conclude that the initial velocity of our solution is a multiple of a delta function at the origin:

$$u_t(0, r) = 4\,\pi\,c\,\delta(\mathbf{x}).$$

Dividing through by $4\,\pi\,c$, we find that the spherical expanding wave

$$u(t, r) = \frac{\delta(r - c\,t)}{4\,\pi\,c\,r} \tag{12.147}$$

solves the initial value problem

$$u(0, \mathbf{x}) \equiv 0, \qquad \frac{\partial u}{\partial t}\,(0, \mathbf{x}) = \delta(\mathbf{x}),$$

corresponding to an initial unit-velocity impulse concentrated at the origin. This solution can be viewed as the three-dimensional version of the hammer-blow solution to the one-dimensional wave equation discussed in Exercise 6.3.28.

More generally, we use the translational symmetry of the wave equation to conclude that the function

$$G(t, \mathbf{x}; \boldsymbol{\xi}) = \frac{\delta\big(\,\|\,\mathbf{x} - \boldsymbol{\xi}\,\| - c\,t\,\big)}{4\,\pi\,c\,\|\,\mathbf{x} - \boldsymbol{\xi}\,\|}\,, \qquad t \geq 0, \tag{12.148}$$

is the *fundamental solution* to the wave equation resulting from a unit-velocity impulse concentrated at the point $\boldsymbol{\xi}$ at the initial time $t = 0$:

$$G(0, \mathbf{x}; \boldsymbol{\xi}) = 0, \qquad \frac{\partial G}{\partial t}\,(0, \mathbf{x}; \boldsymbol{\xi}) = \delta(\mathbf{x} - \boldsymbol{\xi}). \tag{12.149}$$

With this in hand, we can apply linear superposition to solve the zero initial displacement initial value problem

$$u(0, x, y, z) = 0, \qquad \frac{\partial u}{\partial t}\,(0, x, y, z) = g(x, y, z). \tag{12.150}$$

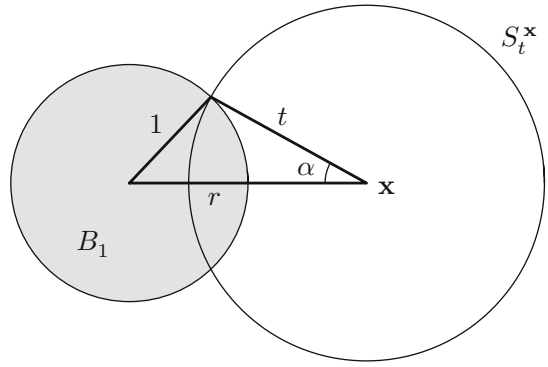


Figure 12.9. Cross-section of a sphere intersecting a ball.

Namely, we write the initial velocity

$$g(\mathbf{x}) = \iiint g(\boldsymbol{\xi})\, \delta(\mathbf{x} - \boldsymbol{\xi})\, d\xi\, d\eta\, d\zeta$$

as a superposition of impulses, and immediately conclude that the relevant solution is the selfsame superposition of spherical waves:

$$\begin{aligned} u(t, \mathbf{x}) &= \frac{1}{4\pi c} \iiint g(\boldsymbol{\xi})\, \frac{\delta\big(\,\|\,\mathbf{x} - \boldsymbol{\xi}\,\| - c\,t\,\big)}{\|\,\mathbf{x} - \boldsymbol{\xi}\,\|}\; d\xi\, d\eta\, d\zeta \\ &= \frac{1}{4\pi c^2\, t} \iint_{\|\,\boldsymbol{\xi} - \mathbf{x}\,\| = c\,t} g(\boldsymbol{\xi})\, dS = t\, \mathrm{M}_{c\,t}^{\mathbf{x}}[\,g\,]\,. \end{aligned} \tag{12.151}$$

Thus, the value of our solution at position $\mathbf{x}$ and time $t > 0$ is equal to t times the *mean* of the initial velocity function g over the sphere of radius $r = c\,t$ centered at the point $\mathbf{x}$.

Example 12.16. Let us set the wave speed $c = 1$. Suppose that the initial velocity

$$g(\mathbf{x}) = \begin{cases} 1, & \|\,\mathbf{x}\,\| < 1, \\ 0, & \|\,\mathbf{x}\,\| > 1, \end{cases}$$

is 1 inside the unit ball B_1 centered at the origin and 0 outside. To solve the corresponding initial velocity problem, we must compute the average value of g over a sphere

$$S_t^{\mathbf{x}} = \{\, \boldsymbol{\xi} \mid \|\,\boldsymbol{\xi} - \mathbf{x}\,\| = t\,\}$$

of radius $t > 0$ centered at a point $\mathbf{x} \in \mathbb{R}^3$. Since $g = 0$ outside the unit ball, its average will be equal to the surface area of that part of the sphere that is contained inside the unit ball, namely $S_t^{\mathbf{x}} \cap B_1$, divided by the total surface area of $S_t^{\mathbf{x}}$, namely $4\pi t^2$.

To compute this quantity, let $r = \|\,\mathbf{x}\,\|$. If $t > r + 1$ or $0 < t < r - 1$, then the sphere of radius t lies entirely outside the unit ball, and so the average is 0; if $0 < t < 1 - r$, which requires $r < 1$ and so $\mathbf{x} \in B_1$, then the sphere lies entirely within the unit ball, and so the average is 1. Otherwise, referring to Figure 12.9 and Exercise 12.6.7, we see that the area of the spherical cap $S_t^{\mathbf{x}} \cap B_1$ is given by

$$2\pi t^2(1 - \cos\alpha) = 2\pi t^2 \left(1 - \frac{r^2 + t^2 - 1}{2\,r\,t}\right) = \frac{\pi\, t}{r}\,[\,1 - (t - r)^2\,]\,, \tag{12.152}$$

where α denotes the angle between the line joining the centers of the two spheres and the circle formed by their intersection, whose value is prescribed by the Law of Cosines. Assembling the different subcases, we conclude that

$$\mathrm{M}^{\mathbf{x}}_{ct}[\,g\,] = \begin{cases} 1, & 0 \le t \le 1 - r, \\ \dfrac{1 - (t-r)^2}{4\,r\,t}, & |\,r - 1\,| \le t \le r + 1, \\ 0, & 0 \le t \le r - 1 \quad \text{or} \quad t \ge r + 1. \end{cases} \tag{12.153}$$

The solution (12.151) is obtained by multiplying by t, and hence for $t \ge 0$,

$$u(t,\mathbf{x}) = \begin{cases} t, & 0 \le t \le 1 - \|\,\mathbf{x}\,\|, \\ \dfrac{1 - \big(t - \|\,\mathbf{x}\,\|\big)^2}{4\,\|\,\mathbf{x}\,\|}, & \big|\,\|\,\mathbf{x}\,\| - 1\,\big| \le t \le \|\,\mathbf{x}\,\| + 1, \\ 0, & 0 \le t \le \|\,\mathbf{x}\,\| - 1 \quad \text{or} \quad t \ge \|\,\mathbf{x}\,\| + 1. \end{cases} \tag{12.154}$$

The resulting function is not smooth at the interfaces $t = \big|\,\|\,\mathbf{x}\,\| - 1\,\big|$ and $\|\,\mathbf{x}\,\| + 1$, and hence does not qualify as a classical solution. Nevertheless, it can be shown that (12.154) is a bona fide weak solution to the initial value problem.

The first two rows of Figure 12.10 plot the solution as a function of time for several fixed values of $r = \|\,\mathbf{x}\,\|$. An observer sitting at the origin will see a linearly increasing light intensity followed by a sudden blackout. At other points inside the sphere, there is a similar linear increase, while the subsequent decrease follows a parabolic arc; if the observer is closer to the edge of the ball than its center, the parabolic portion will continue to increase for a while before eventually tapering off. On the other hand, an observer sitting outside the sphere will experience, after an initially dark period, a symmetric parabolic increase to a maximal intensity and then a decrease back to dark after a total time lapse of 2. The second two rows plot the solution as a function of r for various fixed times. Note that, up until time $t = 1$, the light spreads out while increasing in intensity near the origin, after which the solution is of gradually decreasing magnitude, supported within the domain lying between two concentric spheres of respective radii $t - 1$ and $t + 1$.

Returning to the general situation, we note that the solution formula (12.151) handles only nonzero initial velocities. What about solutions resulting from a nonzero initial displacement? Surprisingly, the answer comes from differentiation! The key observation is that if $u(t,\mathbf{x})$ is any (sufficiently smooth) solution to the wave equation, then so is its time derivative

$$v(t,\mathbf{x}) = \frac{\partial u}{\partial t}\,(t,\mathbf{x}). \tag{12.155}$$

This follows at once from differentiating both sides of the wave equation with respect to t and using the equality of mixed partial derivatives. Physically, this implies that the velocity of a wave obeys the same evolutionary principle as the wave itself, which is a manifestation of the linearity and time-independence (autonomy) of the equation.

Now, suppose u has initial conditions

$$u(0,\mathbf{x}) = f(\mathbf{x}), \qquad u_t(0,\mathbf{x}) = g(\mathbf{x}). \tag{12.156}$$

What are the initial conditions for its derivative $v = u_t$? Clearly, its initial displacement

$$v(0,\mathbf{x}) = u_t(0,\mathbf{x}) = g(\mathbf{x}) \tag{12.157}$$

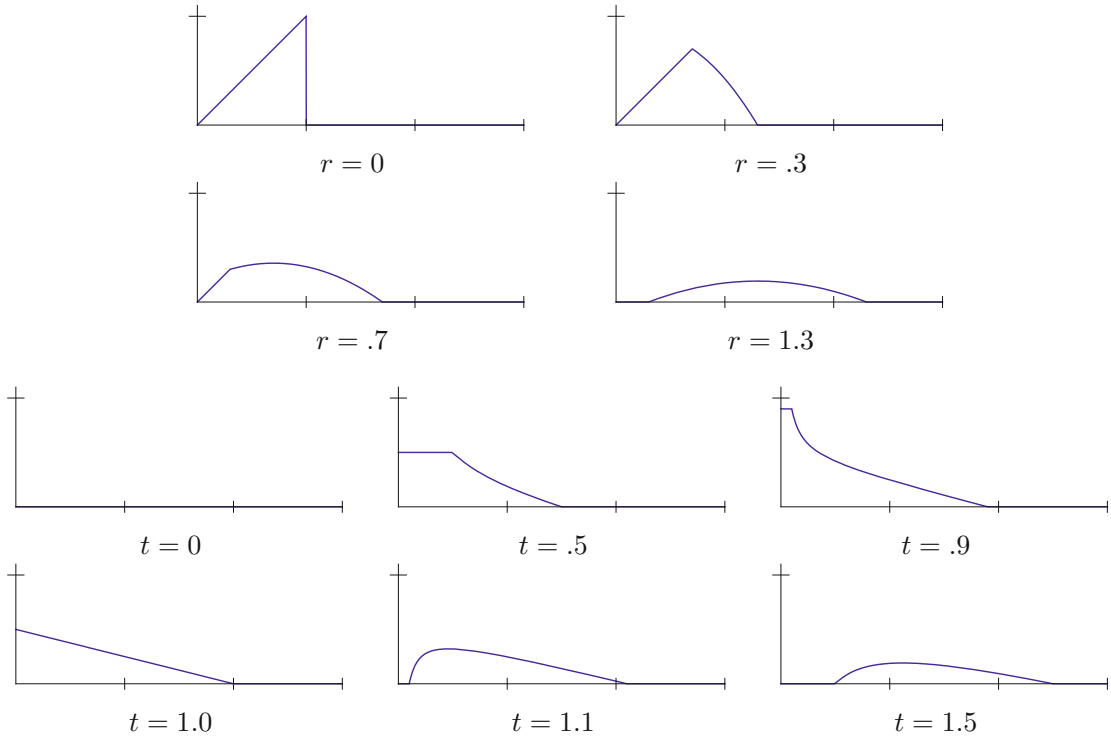


Figure 12.10. Wave equation solution $u(t, r)$ due to an initial velocity of the unit ball.

equals the initial velocity of u. As for its initial velocity, we have

$$\frac{\partial v}{\partial t} = \frac{\partial^2 u}{\partial t^2} = c^2 \, \Delta u,$$

because we are assuming that u solves the wave equation. Thus, at the initial time, the velocity,

$$\frac{\partial v}{\partial t}\,(0, \mathbf{x}) = c^2 \, \Delta u(0, \mathbf{x}) = c^2 \, \Delta f(\mathbf{x}), \tag{12.158}$$

equals c^2 times the Laplacian of the initial displacement f. In particular, if u satisfies the initial conditions

$$u(0, \mathbf{x}) = 0, \qquad u_t(0, \mathbf{x}) = g(\mathbf{x}), \tag{12.159}$$

then $v = u_t$ satisfies the initial conditions

$$v(0, \mathbf{x}) = g(\mathbf{x}), \qquad v_t(0, \mathbf{x}) = 0. \tag{12.160}$$

Thus, paradoxically, to solve the initial displacement problem we differentiate the initial velocity solution (12.151) with respect to t, and hence

$$v(t, \mathbf{x}) = \frac{\partial u}{\partial t}\,(t, \mathbf{x}) = \frac{\partial}{\partial t}\left(t \, \mathrm{M}^{\mathbf{x}}_{ct}\,[\,g\,] \right) = \mathrm{M}^{\mathbf{x}}_{ct}\,[\,g\,] + c\,t\,\mathrm{M}^{\mathbf{x}}_{ct}\left[\, \frac{\partial g}{\partial \mathbf{n}} \,\right], \tag{12.161}$$

where we have made use of our computation in (12.146). Therefore, $v(t, \mathbf{x})$ is a linear combination of the mean of the function g and the mean of its normal or radial derivative $\partial g/\partial \mathbf{n} = \partial g/\partial r$, taken over a sphere of radius $c\,t$ centered at the point $\mathbf{x}$. In particular, to obtain the solution corresponding to a concentrated initial displacement,

$$F(0, \mathbf{x}; \boldsymbol{\xi}) = \delta(\mathbf{x} - \boldsymbol{\xi}), \qquad \frac{\partial F}{\partial t}\,(0, \mathbf{x}; \boldsymbol{\xi}) = 0, \tag{12.162}$$

we differentiate the solution (12.148), resulting in

$$F(t, \mathbf{x}; \boldsymbol{\xi}) = \frac{\partial G}{\partial t}\,(t, \mathbf{x}; \boldsymbol{\xi}) = -\,\frac{\delta'\big(\,\|\,\mathbf{x} - \boldsymbol{\xi}\,\| - c\,t\,\big)}{4\,\pi\,\|\,\boldsymbol{\xi} - \mathbf{x}\,\|}\,, \tag{12.163}$$

which is the fundamental solution for the initial displacement problem. Thus, interestingly, a concentrated initial displacement spawns a spherically expanding doublet, cf. Figure 6.6, whereas a concentrated initial velocity spawns an expanding spherical singlet or delta wave.

Example 12.17. Let $c = 1$. Consider the initial conditions

$$u(0, \mathbf{x}) = f(\mathbf{x}) = \begin{cases} 1, & \|\,\mathbf{x}\,\| < 1, \\ 0, & \|\,\mathbf{x}\,\| > 1, \end{cases} \qquad \frac{\partial u}{\partial t}\,(0, \mathbf{x}) = 0, \tag{12.164}$$

modeling the effect of an instantaneously illuminated solid ball. To obtain the resulting solution, we differentiate (12.154) with respect to t, leading to

$$u(t, \mathbf{x}) = \begin{cases} 1, & 0 \le t < 1 - \|\,\mathbf{x}\,\|, \\ \dfrac{\|\,\mathbf{x}\,\| - t}{2\,\|\,\mathbf{x}\,\|}, & \big|\,\|\,\mathbf{x}\,\| - 1\,\big| \le t \le \|\,\mathbf{x}\,\| + 1, \\ 0, & 0 \le t < \|\,\mathbf{x}\,\| - 1 \quad \text{or} \quad t > 1 + \|\,\mathbf{x}\,\|. \end{cases} \tag{12.165}$$

As illustrated in the first two rows of Figure 12.11, an observer sitting at the center of the ball will see a constant light intensity until $t = 1$, at which time the solution suddenly goes dark. At other points inside the ball, $0 < r < 1$, the downwards jump in intensity arrives sooner, and even goes below 0, followed by a further linear decrease, and finally a jump back to quiescence. An observer placed outside the ball, at radius $r = \|\,\mathbf{x}\,\| > 1$, will experience, after an initially dark period, a sudden increase in the light intensity at time $t = r - 1$, followed by a linear decrease to negative, followed by a jump back up to darkness at time $t = r + 1$. The farther away from the source, the fainter the light. In the second two rows, we plot the same solution as a function of r for different values of t. Note the sudden appearance of a $1/r$ singularity at the origin at time $t = 1$, which is the result of a focusing of the initial discontinuities of $u(0, \mathbf{x}) = f(\mathbf{x})$ on the surface of the unit sphere. Afterwards, the residual radially symmetric disturbance moves off to ∞ while gradually decreasing in intensity. Again, the discontinuities imply that (12.165) is not a classical solution, but it does qualify as a weak solution to the initial value problem.

Kirchhoff's Formula and Huygens' Principle

Linearly combining the two solution formulas (12.151) and (12.161) establishes *Kirchhoff's formula* (first discovered by Poisson), which is the three-dimensional counterpart to d'Alembert's solution formula for the wave equation.

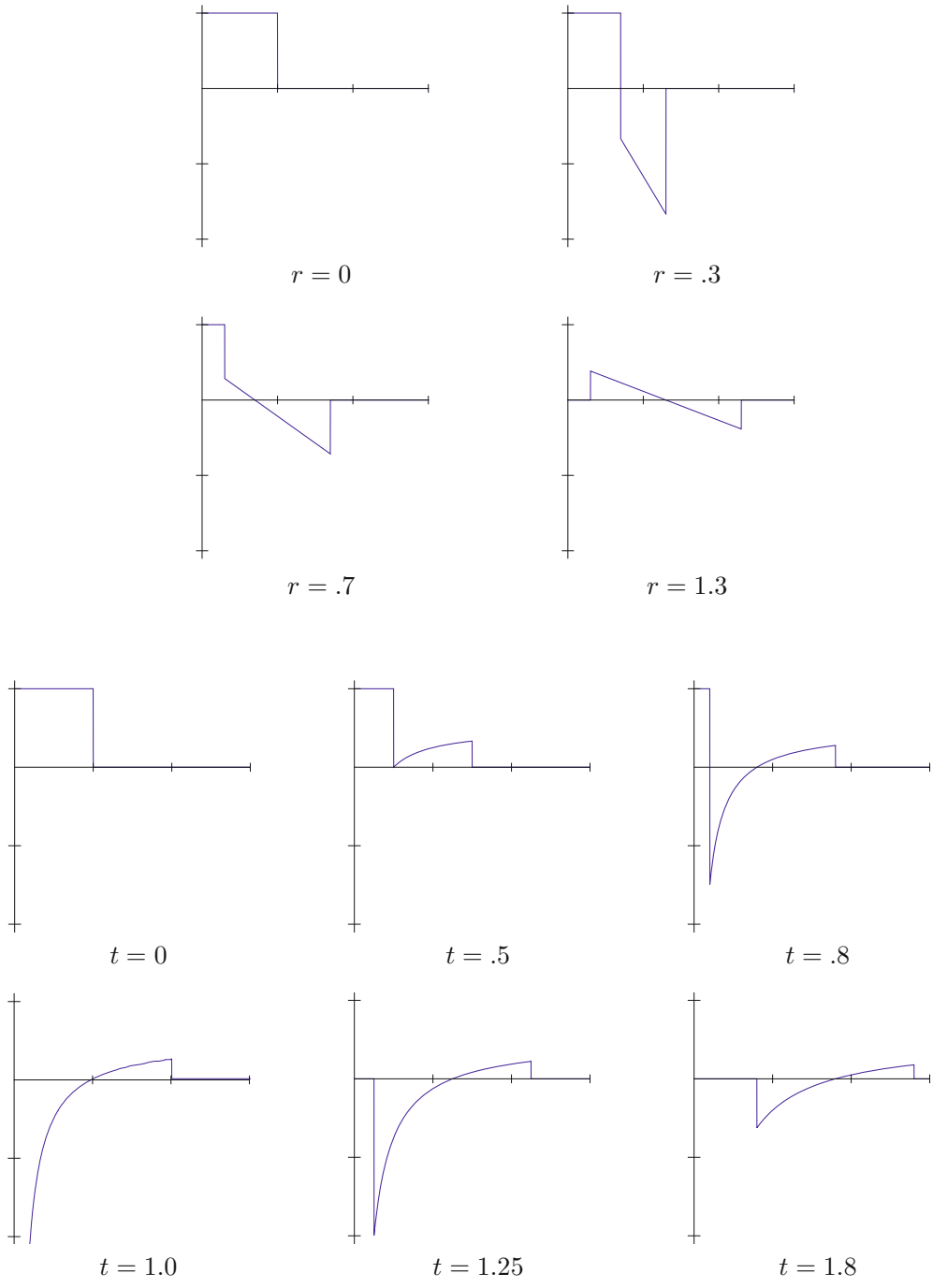


Figure 12.11. Wave equation solution $u(t, r)$ due to an initial displacement of the unit ball .

Theorem 12.18. *The solution to the initial value problem*

$$u_{tt} = c^2\,\Delta u, \qquad u(0,\mathbf{x}) = f(\mathbf{x}), \qquad \frac{\partial u}{\partial t}\,(0,\mathbf{x}) = g(\mathbf{x}), \qquad \mathbf{x} \in \mathbb{R}^3, \tag{12.166}$$

for the wave equation in three-dimensional space is given by

$$u(t,\mathbf{x}) = \frac{\partial}{\partial t}\left(t\,\mathrm{M}^{\mathbf{x}}_{ct}[\,f\,]\right) \;+\; t\,\mathrm{M}^{\mathbf{x}}_{ct}[\,g\,] = \mathrm{M}^{\mathbf{x}}_{ct}[\,f\,] + c\,t\,\mathrm{M}^{\mathbf{x}}_{ct}\left[\,\frac{\partial f}{\partial \mathbf{n}}\,\right] \;+\; t\,\mathrm{M}^{\mathbf{x}}_{ct}[\,g\,], \tag{12.167}$$

where $\mathrm{M}^{\mathbf{x}}_{ct}[\,f\,]$ *denotes the average of the function* f *over the sphere* $S^{\mathbf{x}}_{ct} = \{\,\|\,\boldsymbol{\xi} - \mathbf{x}\,\| = c\,t\,\}$ *of radius* $c\,t$ *centered at the point* $\mathbf{x}$.

A crucially important consequence of the Kirchhoff solution formula is a celebrated physical principle first set out by the pioneering seventeenth century Dutch scientist Christiaan Huygens.† Roughly, *Huygens' Principle* states that, in three-dimensional space, localized solutions to the wave equation remain localized. More concretely, (12.167) implies that the value of the solution at a point $\mathbf{x}$ and time t depends only on the values of the initial displacements and velocities at a distance $c\,t$ away. Thus, all signals propagate along the relativistic light cone

$$c^2\,t^2 = x^2 + y^2 + z^2$$

in four-dimensional Minkowski space-time. Physically, Huygens' Principle assures us that any light that we witness at time t arrived from points that lie a distance exactly $d = c(t - t_0)$ away at an earlier time $t_0 < t$. In particular, a localized initial signal, whether initial displacement or initial velocity, that is concentrated near a point produces a response that remains concentrated on an ever expanding sphere surrounding the point. In our three-dimensional universe, we witness the light from a sudden explosion or lightning bolt for only a brief moment, after which the view returns to darkness. Similarly, a sharp sound, e.g., a thunderclap, remains sharply concentrated with diminishing magnitude as it propagates through space. Huygens' Principle is responsible for the important astronomical fact that the light we now observe from a distant star was generated at a single past time that is directly proportional to the star's distance from the Earth. Remarkably, as we will show in the following subsection, Huygens' Principle *does not hold* in a two-dimensional universe! There, initially concentrated light and sound impulses will spread out as time progresses, and their effect will be experienced over an extended time range; see below for details.

Exercises

12.6.1. Solve the wave equation in three-dimensional space for the following initial conditions:
(*a*) $u(0,x,y,z) = x + z$, $u_t(0,x,y,z) = 0$; (*b*) $u(0,x,y,z) = 0$, $u_t(0,x,y,z) = y$;
(*c*) $u(0,x,y,z) = 1/(1 + x^2 + y^2 + z^2)$, $u_t(0,x,y,z) = 0$,
(*d*) $u(0,x,y,z) = 0$, $u_t(0,x,y,z) = 1/(1 + x^2 + y^2 + z^2)$.

12.6.2. At what points in space-time does a three-dimensional wave vanish if it vanishes outside a sphere of radius R at the initial time $t = 0$?

† Don't even bother trying to pronounce his name correctly unless you are Dutch!

12.6.3. Consider the initial value problem

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}, \qquad u(0,x,y,z) = 0, \qquad \frac{\partial u}{\partial t}(0,x,y,z) = \begin{cases} 1, & 0 < x,y,z < 1, \\ 0, & \text{otherwise,} \end{cases}$$

i.e., the initial velocity is 1 inside a unit cube and 0 outside the cube. We interpret the solution $u(t,x,y,z)$ as the intensity of light at a given point in space-time, measured in units that make the speed of light $c = 1$. (*a*) Write down an integral formula for $u(t,x,y,z)$. (*b*) Suppose a light sensor is placed at the point $(2,2,1)$. For which values of $t > 0$ will the sensor register a nonzero signal? Sketch a rough graph of what the sensor measures. (You do not need to find the precise formula, but explain how you obtained your graph.) (*c*) *True or false*: The solution $u(t,x,y,z) \geq 0$ at all points in space-time.

12.6.4. Is (12.151) a solution to the wave equation for $t < 0$? If not, write down a solution formula that is valid for negative t.

12.6.5. *True or false*: The function $u(t,x,y,z)$ defined by (12.154) is everywhere continuous.

12.6.6. A thermonuclear explosion occurs at the center of the Earth. Would you feel the effect first through a motion at the surface or a change in temperature at the surface? Discuss.

♢ 12.6.7. Prove that the area of the spherical cap $S_t^{\mathbf{x}} \cap B_1$ is given by formula (12.152).

Descent to Two Dimensions

So far, we have found explicit formulas for the solution to the wave equation on the one-dimensional line, and in three-dimensional space. The two-dimensional case

$$u_{tt} = c^2 \Delta u = c^2 (u_{xx} + u_{yy}) \tag{12.168}$$

is, counterintuitively, more complicated! For instance, seeking a radially symmetric solution $u(t,r)$ requires solving the partial differential equation

$$\frac{\partial^2 u}{\partial t^2} = c^2 \left(\frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r} \right), \tag{12.169}$$

which, unlike its three-dimensional counterpart (12.137), is not so easily integrated.

However, our solution to the three-dimensional problem can be adapted to construct a solution to the two-dimensional problem using the so-called *Method of Descent*. Observe that any solution $u(t,x,y)$ to the two-dimensional wave equation (12.168) can be viewed as a solution to the three-dimensional wave equation (12.123) that does not depend upon the vertical z coordinate, whence $\partial u/\partial z = 0$. Clearly, if the three-dimensional initial data does not depend on z, then the resulting solution $u(t,x,y)$ will also be independent of z.

Consider first the zero initial displacement condition

$$u(0,x,y) = 0, \qquad \frac{\partial u}{\partial t}(0,x,y) = g(x,y). \tag{12.170}$$

In the three-dimensional solution formula (12.151), if $g(x,y)$ does not depend on the z–coordinate, then the integrals over the upper and lower hemispheres

$$S_{ct}^{+} = \big\{ \| \boldsymbol{\xi} - \mathbf{x} \| = c\,t, \;\; \zeta \geq z \big\}, \qquad S_{ct}^{-} = \big\{ \| \boldsymbol{\xi} - \mathbf{x} \| = c\,t, \;\; \zeta \leq z \big\},$$

are identical. To evaluate these integrals, we parametrize the upper hemisphere as the graph of

$$\zeta = z + \sqrt{c^2 t^2 - (\xi - x)^2 - (\eta - y)^2} \quad \text{over the disk} \quad D_{ct}^{\mathbf{x}} = \big\{ (\xi - x)^2 + (\eta - y)^2 \le c^2 t^2 \big\},$$

concluding that

$$\begin{aligned} u(t, x, y) &= \frac{1}{4\pi c^2 t} \iint_{S_{ct}} g(\xi, \eta)\, dS = \frac{1}{2\pi c^2 t} \iint_{S_{ct}^+} g(\xi, \eta)\, dS \\ &= \frac{1}{2\pi c} \iint_{D_{ct}^{\mathbf{x}}} \frac{g(\xi, \eta)}{\sqrt{c^2 t^2 - (\xi - x)^2 - (\eta - y)^2}}\, d\xi\, d\eta \end{aligned} \tag{12.171}$$

solves the initial value problem (12.170). In particular, if we take the initial velocity

$$\frac{\partial u}{\partial t}(0, x, y) = g(x, y) = \delta(x)\, \delta(y)$$

to be a unit impulse concentrated at the origin, then the resulting solution is

$$u(t, x, y) = \begin{cases} \dfrac{1}{2\pi c \sqrt{c^2 t^2 - x^2 - y^2}}, & x^2 + y^2 < c^2 t^2, \\ 0, & x^2 + y^2 > c^2 t^2. \end{cases} \tag{12.172}$$

An observer sitting at distance $r = \| \mathbf{x} \| = \sqrt{x^2 + y^2}$ from the origin will first witness a concentrated displacement singularity at time $t = r/c$. However, in contrast to the three-dimensional solution, even after the impulse passes by, there will continue to be a decreasing, but nonzero, signal of magnitude roughly proportional to $1/t$. In Figure 12.12, we plot the solution (12.172) for unit wave speed $c = 1$. The first row plots intensity as a function of t at three different radii; note that the initial singularity, indicated by a spike in the graph, is followed by a progressively smaller residual displacement, which never entirely disappears. The second row shows the displacement at three different times as a function of $r = \| \mathbf{x} \|$.

As in the three-dimensional case, the solution to the initial displacement conditions

$$u(0, x, y) = f(x, y), \qquad \frac{\partial u}{\partial t}(0, x, y) = 0, \tag{12.173}$$

can then be obtained by differentiation of (12.171) with respect to t, and so

$$u(t, x, y) = \frac{1}{2\pi c} \frac{\partial}{\partial t} \iint_{D_{ct}^{\mathbf{x}}} \frac{f(\xi, \eta)}{\sqrt{c^2 t^2 - (\xi - x)^2 - (\eta - y)^2}}\, d\xi\, d\eta. \tag{12.174}$$

As before, starting with a concentrated impulse, an observer will witness, after a time lapse $t = r/c$, an abrupt impulse passing by, followed by a progressively decaying residual wave. The general solution to the two-dimensional wave equation on all of $\mathbb{R}^2$ is a linear combination of these two types of solutions (12.171, 174).

As a consequence of these considerations, we discover that Huygens' Principle is *not* valid in a two-dimensional universe. The solution to the two-dimensional wave equation at a point $\mathbf{x}$ at time t depends on the initial displacement and velocity on the entire disk of radius ct centered at the point, and not just on the points lying a distance ct away. So a two-dimensional creature would experience not just an initial effect of a concentrated sound or light wave, but also an "afterglow" of slowly diminishing magnitude. It would be

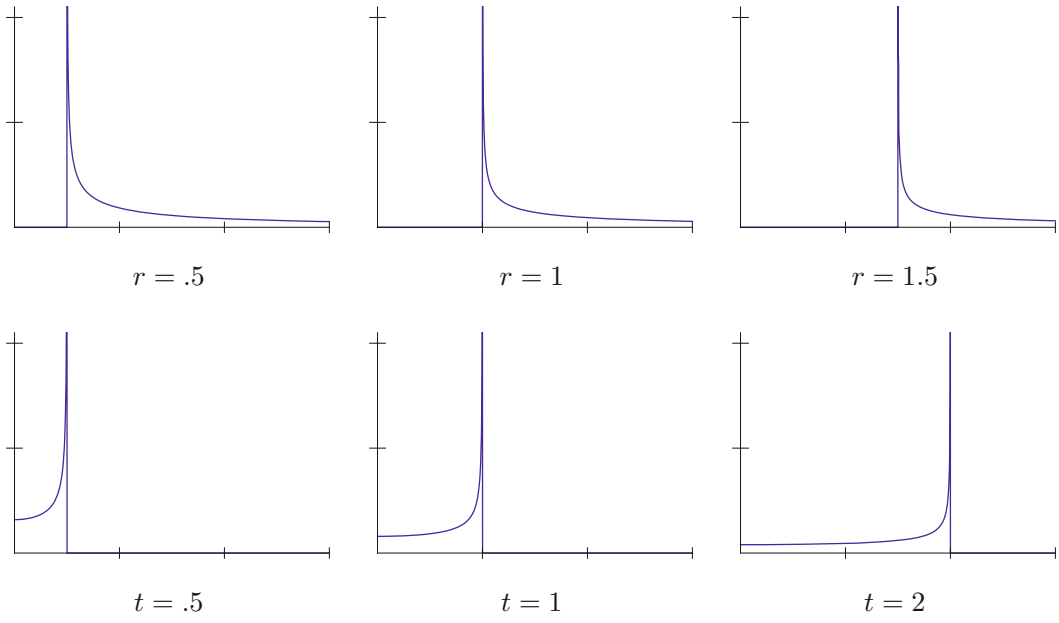


Figure 12.12. Solution to the two-dimensional wave equation for a concentrated impulse.

like living in a permanent echo chamber, and so understanding and acting upon sensory phenomena would be considerably more challenging. In general, it can be proved that Huygens' Principle for the wave equation is valid only in spaces of odd dimension $n = 2k + 1 \geq 3$; see also [**15**] for recent advances in the classification of partial differential equations that admit a Huygens' Principle.

Remark: Since the solutions to the two-dimensional wave equation can be interpreted as three-dimensional solutions with no z dependence, a concentrated delta impulse for the two-dimensional wave equation would correspond to a three-dimensional initial impulse that is concentrated along an entire vertical line, e.g., an instantaneous lightning bolt in the form of an infinite straight line. An observer fixed in space will first encounter the light flash arriving from the closest point on the line, but will subsequently experience the gradually decreasing effect of the light emitted by points that lie progressively farther away along the line. This accounts for the two-dimensional afterglow in formula (12.172).

Exercises

12.6.8. Solve initial value problem for the two-dimensional wave equation with the following initial data (*a*) $u(0, x, y) = x - y$, $u_t(0, x, y) = 0$; (*b*) $u(0, x, y) = 0$, $u_t(0, x, y) = y$.

12.6.9. (a) Prove that $u(t,x,y) = 1/\sqrt{x^2+y^2-c^2t^2}$ is a solution to the two-dimensional wave equation on the domain $\Omega = \{\, x^2+y^2 > c^2t^2 \,\}$ exterior to the light cone passing through the origin. What is the corresponding initial data at $t = 0$? (b) Use part (a) to solve the initial value problem $u(0,x,y) = 0$, $u_t(0,x,y) = 1/\sqrt{x^2+y^2}$, on Ω.

12.6.10. Consider the two-dimensional wave equation on $\mathbb{R}^2$ with wave speed $c = 1$. Write down an integral formula for the solution to the following initial value problems. You need not evaluate the integrals. (a) $u(0,x,y) = x^3 - y^3$, $u_t(0,x,y) = 0$; (b) $u(0,x,y) = 0$, $u_t(0,x,y) = y^2$; (c) $u(0,x,y) = x^2 + y^2$, $u_t(0,x,y) = -x^2 - y^2$.

12.6.11. (a) Find the solution to the two-dimensional wave equation whose initial displacement is a concentrated delta impulse at the origin and whose initial velocity is zero.
(b) Is your expression a classical solution when $t > 0$?
(c) *True or false*: The solution tends to 0 uniformly as $t \to \infty$.

12.6.12. Use separation of variables to write down an eigenfunction series solution to the partial differential equation (12.169) when subject to homogeneous Dirichlet boundary conditions at $r = 1$ and bounded at $r = 0$.

♢ 12.6.13. Write down the fundamental solution for the one-dimensional wave equation with (a) a concentrated initial displacement at the origin; (b) a concentrated initial velocity at the origin. (c) Discuss the validity of Huygens' Principle in a one-dimensional universe.

12.6.14. Discuss how you can construct solutions to the one-dimensional wave equation by descent from the three-dimensional wave equation.

12.7 The Hydrogen Atom

A *hydrogen atom* consists of a single electron circling an atomic nucleus that contains a single proton, which, owing to it relatively tiny size, is assumed to be entirely concentrated at the origin. As a result of quantization of the corresponding classical Coulomb problem, the Schrödinger equation[†] governing the dynamical behavior of the electron moving around the nucleus takes the explicit form

$$\mathrm{i}\,\hbar\,\frac{\partial \psi}{\partial t} = -\,\frac{\hbar^2}{2\,M}\,\Delta\psi - \frac{\alpha^2}{r}\,\psi = -\,\frac{\hbar^2}{2\,M}\left(\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial y^2}\right) - \frac{\alpha^2\,\psi}{\sqrt{x^2+y^2+z^2}}\,. \tag{12.175}$$

Here $\psi(t,x,y,z)$ denotes the electron's time-dependent wave function, which, at each time t, prescribes its quantum probability density as it circles the nucleus. In the quantized Hamiltonian operator $K = -\frac{1}{2}\,(\hbar^2/M)\,\Delta - \alpha^2/r$, the coefficient of the Laplacian depends on Planck's constant $\hbar$ and the electron's mass M. The final term represents the three-dimensional electromagnetic (Coulomb) potential function $V(\mathbf{x}) = \alpha^2/r$ attracting the electron to the nucleus, with α representing the electron's (and proton's) charge, while $r = \|\,\mathbf{x}\,\|$ is its distance from the nucleus. Incidentally, the quantum-mechanical Schrödinger equation for multi-electron atoms or even molecules is not so difficult to write down, but its solution, even for, say, the helium atom, is *much* more difficult, and is still a major

[†] The reader is referred to (9.151) and the subsequent discussion for generalities regarding the Schrödinger equation and quantum mechanics.

challenge for numerical analysts, even on today's supercomputers, [**116**]. Thus, to keep matters as simple as possible, we will consider only the case of a single electron hydrogen atom here.

Bound States

According to the analysis in Section 9.5, the normal mode solutions to the Schrödinger equation are of the form

$$\psi(t,x,y,z) = e^{\,\mathrm{i}\,\lambda\, t/\hbar}\, v(x,y,z),$$

where v is an eigenfunction of the Hamiltonian operator with eigenvalue λ, and hence satisfies

$$\frac{\hbar^2}{2\,M}\,\Delta v + \left(\lambda + \frac{\alpha^2}{r}\right) v = 0. \tag{12.176}$$

The *bound states* of the atom, in which the electron remains trapped by the nucleus, are represented by the nonzero solutions to the eigenvalue problem (12.176) with unit L^2 norm:

$$\| v \|^2 = \iiint | v(x,y,z) |^2 \, dx\, dy\, dz = 1.$$

The eigenvalue λ specifies the bound state's energy, and is necessarily negative: $\lambda < 0$. Since we are working on an unbounded domain, the bound states do *not* form a complete system of eigenfunctions, and so not every wave function $\varphi \in \mathrm{L}^2(\mathbb{R}^3)$ can be approximated by an eigenfunction series. The missing data are the so-called *scattering states* arising from the *continuous spectrum* of the Schrödinger operator; these represent electrons that scatter off the nucleus, and so do not remain trapped in an orbit. (For the classical Kepler problem of a planet circling a sun, the bound states would correspond to planets following bounded elliptic orbits, while the scattering states correspond to interstellar comets and the like moving along unbounded hyperbolic or parabolic trajectories.) We will leave the discussion of the quantum-mechanical scattering states and the associated continuous spectrum to a more advanced treatment of the subject, [**72**, **95**].

To understand the bound states, we will apply the method of separation of variables. We begin by rewriting the eigenvalue problem (12.176) in spherical coordinates:

$$\frac{\hbar^2}{2\,M}\left(\frac{\partial^2 v}{\partial r^2} + \frac{2}{r}\,\frac{\partial v}{\partial r} + \frac{1}{r^2}\,\frac{\partial^2 v}{\partial \varphi^2} + \frac{\cos\varphi}{r^2\sin\varphi}\,\frac{\partial v}{\partial \varphi} + \frac{1}{r^2\sin^2\varphi}\,\frac{\partial^2 v}{\partial \theta^2}\right) + \left(\lambda + \frac{\alpha^2}{r}\right) v = 0. \tag{12.177}$$

We then separate off the radial coordinate, setting

$$v(r,\varphi,\theta) = p(r)\, w(\varphi,\theta).$$

The angular component satisfies the spherical Helmholtz equation

$$\Delta_S\, w + \mu\, w = \frac{\partial^2 w}{\partial \varphi^2} + \frac{\cos\varphi}{\sin\varphi}\,\frac{\partial w}{\partial \varphi} + \frac{1}{\sin^2\varphi}\,\frac{\partial^2 w}{\partial \theta^2} + \mu\, w = 0,$$

which we have already solved; see (12.21) and the ensuing discussion. The eigensolutions are spherical harmonics, which, because the quantum-mechanical solutions are intrinsically complex-valued, we take in their complex form (12.46). The associated eigenvalue

$$\mu = l\,(l+1), \qquad \text{where the integer} \qquad l = 0, 1, 2, \ldots \tag{12.178}$$

is known as the *angular quantum number*, admits a total of $2l + 1$ linearly independent eigenfunctions

$$\mathcal{Y}_l^m(\varphi, \theta) = P_l^m(\cos\varphi)\, e^{\,\mathrm{i}\, m\, \theta}, \qquad m = -l, -l+1, \ldots, l-1, l. \tag{12.179}$$

The radial equation with the separation constant (12.178) is

$$\frac{\hbar^2}{2\,M}\left(\frac{d^2p}{dr^2} + \frac{2}{r}\,\frac{dp}{dr}\right) + \left(\lambda + \frac{\alpha^2}{r} - \frac{l\,(l+1)}{r^2}\right) p = 0. \tag{12.180}$$

To eliminate the physical parameters, let us rescale the radial coordinate by setting

$$s = \sigma\, r, \qquad \text{where} \qquad \sigma = \frac{2\,\sqrt{-2\,M\,\lambda}}{\hbar}\,, \tag{12.181}$$

given that $\lambda < 0$. The resulting ordinary differential equation for the rescaled function

$$P(s) = p\left(\frac{s}{\sigma}\right)$$

is

$$\frac{d^2P}{ds^2} + \frac{2}{s}\,\frac{dP}{ds} - \left(\frac{1}{4} - \frac{n}{s} + \frac{l\,(l+1)}{s^2}\right) P = 0, \tag{12.182}$$

where

$$n = \frac{2\,M\,\alpha^2}{\sigma\,\hbar^2} = \frac{\alpha^2}{\hbar}\,\sqrt{-\frac{M}{2\,\lambda}}\,. \tag{12.183}$$

Equation (12.182) is a version of the *generalized Laguerre differential equation* — see Exercise 12.7.4 below — named after the nineteenth-century French mathematician Edmond Laguerre, who studied its solutions well before the appearance of quantum mechanics. Since we are searching for bound states, the relevant solutions should be defined on $0 \le s < \infty$, remain bounded at $s = 0$, and go to zero as $s \to \infty$:

$$\lim_{s\,\to\,0^+} P(s) < \infty, \qquad \lim_{s\,\to\,\infty} P(s) = 0. \tag{12.184}$$

The proof of the following key result is outlined in Exercises 12.7.4–5.

Theorem 12.19. *For each pair of nonnegative integers $0 \le l < n$, the boundary value problem* (12.182, 184) *has the eigensolution*

$$P_l^n(s) = s^l\, e^{-\,s/2}\, L_{n-l-1}^{2\,l+1}(s), \tag{12.185}$$

where

$$L_k^j(s) = \frac{s^{-\,j} e^s}{k!}\,\frac{d^k}{ds^k}\left[\, s^{j+k} e^{-\,s}\,\right] = \sum_{i\,=\,0}^{k} \frac{(-1)^i}{i!}\binom{j+k}{j+i}\, s^i, \qquad j, k = 0, 1, 2, \ldots\,, \tag{12.186}$$

are known as generalized† Laguerre polynomials.

† The ordinary *Laguerre polynomials* are $L_k(s) = L_k^0(s)$.

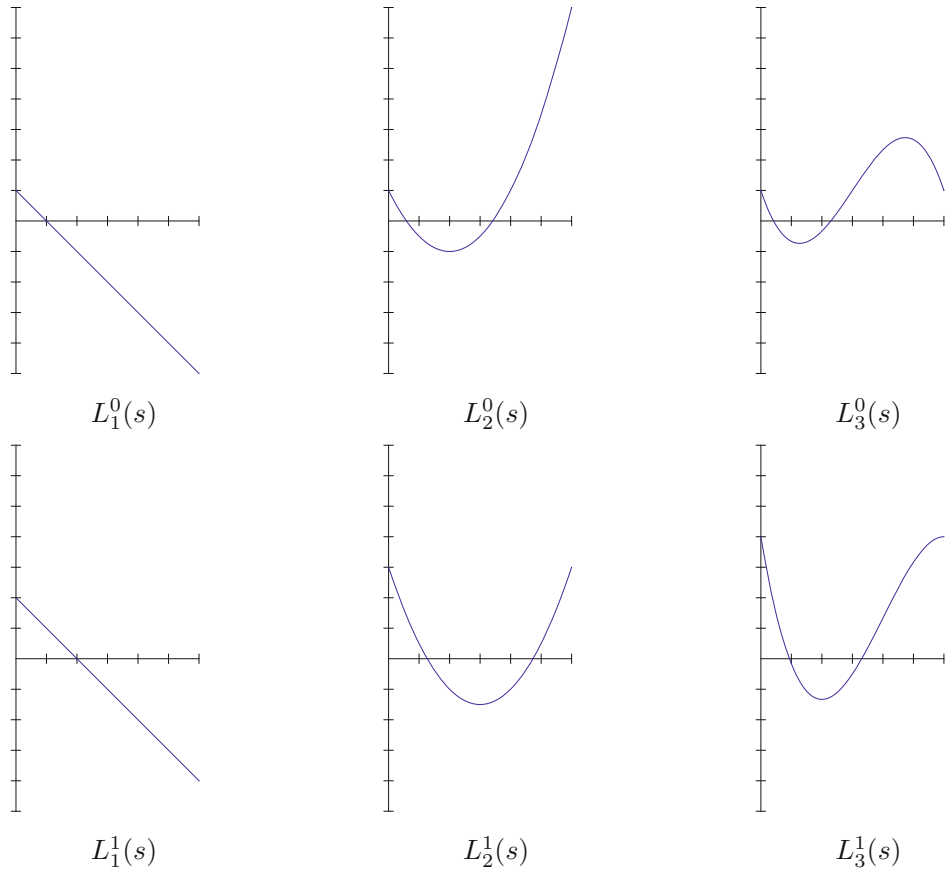


Figure 12.13. Generalized Laguerre polynomials.

The first few generalized Laguerre polynomials are

$$\begin{aligned}
&L_0^0(s) = 1, \quad L_1^0(s) = 1 - s, \quad L_2^0(s) = 1 - 2\,s + \tfrac{1}{2}\,s^2, \quad L_3^0(s) = 1 - 3\,s + \tfrac{3}{2}\,s^2 - \tfrac{1}{6}\,s^3,\\
&L_0^1(s) = 1, \quad L_1^1(s) = 2 - s, \quad L_2^1(s) = 3 - 3\,s + \tfrac{1}{2}\,s^2, \quad L_3^1(s) = 4 - 6\,s + 2\,s^2 - \tfrac{1}{6}\,s^3,\\
&L_0^2(s) = 1, \quad L_1^2(s) = 3 - s, \quad L_2^2(s) = 6 - 4\,s + \tfrac{1}{2}\,s^2, \quad L_3^2(s) = 10 - 10\,s + \tfrac{5}{2}\,s^2 - \tfrac{1}{6}\,s^3.
\end{aligned}$$

Note that $L_k^j(s)$ has degree k. A few graphs, on the interval $0 \le t \le 6$, appear in Figure 12.13. See [**86**] for details on their properties.

Atomic Eigenstates and Quantum Numbers

The integer n, whose physical value was noted in (12.183), is known as the *principal quantum number*. We further note that the scaling factor in (12.181) can be written as

$$\sigma = \frac{2\,M\,\alpha^2}{n\,\hbar^2} = \frac{2}{n\,a}, \qquad \text{where} \qquad a = \frac{\hbar^2}{M\,\alpha^2} \approx .529 \times 10^{-10} \text{ meter},$$

which approximates the radius of the electron's lowest energy level, is known as the *Bohr radius*, in honor of the pioneering Danish quantum physicist Niels Bohr. Reverting to phys-

ical coordinates, the bound state solutions (12.185) become, up to an inessential constant multiple, the *radial wave functions*

$$\beta_l^n(r) = \left(\frac{2r}{na}\right)^l e^{-r/(na)}\, L_{n-l-1}^{2l+1}\left(\frac{2r}{na}\right). \tag{12.187}$$

Combining them with the spherical harmonics (12.179) yields the *atomic eigenfunctions* or *eigenstates*

$$v_{lmn}(r,\varphi,\theta) = \sqrt{\frac{(2l+1)\,(l-m)\,!\,(n-l-1)\,!}{\pi\, a^3 n^4 (l+m)\,!\,(l+n)\,!}}\; \beta_l^n(r)\, \mathcal{Y}_l^m(\varphi,\theta), \tag{12.188}$$

where the initial factor is selected so as to make $\| v_{lmn} \| = 1$, and hence a bona fide wave function. (A proof of this fact is outlined in Exercise 12.7.8.) The eigenstates depend on three integers, which have the following physical designations:

- $n = 1, 2, 3, \dots$: the *principal quantum number*;
- $l = 0, 1, \dots, n-1$: the *angular quantum number*;
- $m = -l, -l+1, \dots, l-1, l$: the *magnetic quantum number*.

The energy is the associated eigenvalue:

$$\lambda_n = -\frac{\alpha^4 M}{2\,\hbar^2}\,\frac{1}{n^2} = -\frac{\alpha^2}{2a}\,\frac{1}{n^2}, \qquad n = 1, 2, 3, \dots. \tag{12.189}$$

The fact that the ratios $\lambda_n/\lambda_1 = 1/n^2$ between the energy levels of an atom are inverse squares of integers was one of the key experimental discoveries that precipitated the discovery of quantum mechanics. Observe that the n^{th} energy level has a total of

$$\sum_{l=0}^{n-1} (2l+1) = n^2 \tag{12.190}$$

linearly independent bound states (12.188). The dimension of the eigenspace corresponds to the number of orbital subshells in the atom for the corresponding energy level. The shells indexed by the angular quantum number, i.e., the order $l = 0, 1, 2, \dots$ of the spherical harmonic, are traditionally labeled by a letter in the sequence $s, p, d, f, g, \dots$, where each successive shell contains $2l + 1$ individual subshells, indexed by the magnetic quantum number m.

The one missing ingredient in this simple model is the electron's *spin*. Since electrons can have one of two possible spins, the *Pauli Exclusion Principle*, first formulated by the Austrian physicist Wolfgang Pauli, tells us that each atomic energy shell can be occupied by at most two electrons. Consequently, the atomic shell with angular quantum number l may contain up to $2(2l + 1)$ electrons. Keep in mind that, since $0 \le l < n$, the l^{th} shell appears only when n is sufficiently large, so that, according to (12.190), the n^{th} energy level contains up to $2n^2$ electrons.

The resulting atomic configuration of electronic energy shells is the explanation for Mendeleev's periodic table. Its rows are indexed by the principal quantum number n, while the columns are labeled by the angular and magnetic quantum numbers l, m, and the spin. As one moves up the periodic table, the electrons in each successive element's atom progressively fill up the lower energy levels, each new shell containing first a single electron, then two electrons with opposite spins. Thus, hydrogen (in its ground state) has a single electron in the $1s$ shell. Helium has two electrons in the $1s$ shell. Lithium has

three electrons, with two of them filling the $1s$ shell and the third in the $2s$ shell. Neon has ten electrons filling the first two energy levels, with two electrons in the $1s$ shell, two in the $2s$ shell, and six in the $2p$ shell. And so on. The one complication is that, owing to the orbital's geometry, as prescribed by the associated spherical harmonic, the angular and, to a lesser extent, magnetic quantum numbers also affect the physically observed energy, and this can cause shells to fill later than might initially be expected. For example, in potassium and calcium, the $4s$ shell is successively filled, followed by scandium, which begins the process of filling the $3d$ subshells. The chemical properties of the elements are, to a very large extent, determined by the placement of their atom's electrons within the outermost energy level. The interested reader can consult, for example, [**67**, **79**] for further details.

Exercises

12.7.1. If the nucleus contains Z protons circled by a single electron, then its atomic potential $V(\mathbf{x})$ is rescaled accordingly, replacing α^2/r by $Z\,\alpha^2/r$. Discuss the induced effect on the energy levels of such an atomic ion.

♡ 12.7.2. (*a*) Write down the time-dependent wave function for a single electron atom when the electron is in its ground state, i.e., the lowest energy level. (*b*) What is the probability density of the electron? (*c*) What is the probability of finding the electron within 1 Bohr radius of the atom? (*d*) Find the distance d (measured in Bohr radii) so that there is a 95% probability of finding the electron within a distance d of the nucleus.

♢ 12.7.3. Prove that the two expressions for the Laguerre polynomials in (12.186) agree.

♢ 12.7.4. (*a*) Let $k = 0, 1, 2, \dots$ be a nonnegative integer. The *Laguerre differential equation* of *order* k is

$$x\,u'' + (1 - x)\,u' + k\,u = 0. \tag{12.191}$$

Show that $x = 0$ is a regular singular point. Then prove that the Frobenius solution based at $x = 0$ is a polynomial of degree j that coincides with the Laguerre polynomial $L_k^0(x)$.

(*b*) Given nonnegative integers $j, k \geq 0$, use the Frobenius method to prove that the *generalized Laguerre differential equation*

$$x\,u'' + (j + 1 - x)\,u' + k\,u = 0 \tag{12.192}$$

has a polynomial solution that can be identified with the generalized Laguerre polynomial $L_k^j(x)$ in (12.186).

♢ 12.7.5. Suppose that $P(s)$ solves the ordinary differential equation (12.182). Prove that $Q(s) = s^{-l}\,e^{s/2}\,P(s)$ solves the differential equation

$$s\,\frac{d^2Q}{ds^2} + [2(l + 1) - s]\,\frac{dQ}{ds} + (n - l - 1)Q = 0. \tag{12.193}$$

Then apply the result of Exercise 12.7.4 to complete the proof of Theorem 12.19.

♡ 12.7.6. Suppose $f(x)$ is a polynomial, and let $L_k^j(s)$ denote the generalized Laguerre polynomials (12.186). (*a*) Prove that, for $j, k \geq 0$,

$$\int_0^\infty f(s)\,L_k^j(s)\,s^j\,e^{-s}\,ds = \frac{(-1)^k}{k!}\int_0^\infty f^{(k)}(s)\,s^{j+k}\,e^{-s}\,ds.$$

(*b*) For fixed j, prove that the generalized Laguerre polynomials $L_k^j(s)$, $k = 0, 1, 2, \dots$, are orthogonal with respect to the weighted inner product $\langle\, f\,, g\,\rangle = \displaystyle\int_0^\infty f(s)\,g(s)\,s^j\,e^{-s}\,ds$.

(*c*) Prove the formula for their corresponding norms: $\| L_k^j \| = \sqrt{\dfrac{(j + k)!}{k!}}$.

♢ 12.7.7. (*a*) Prove that the generalized Laguerre polynomials satisfy the following recurrence relation:

$$(k+1)\,L^j_{k+1}(s) - (j+2\,k+1-s)\,L^j_k(s) + (j+k)\,L^j_{k-1}(s) = 0. \tag{12.194}$$

(*b*) Prove that

$$\int_0^\infty s^{j+1} e^{-s} \left[L^j_k(s)\right]^2 ds = \frac{(j+2\,k+1)\,(j+k)!}{k!}\,. \tag{12.195}$$

Hint: Use part (a) and Exercise 12.7.6.

♡ 12.7.8. Prove that the atomic eigenfunctions (12.188) form an orthonormal system of wave functions with respect to the L^2 inner product on $\mathbb{R}^3$. *Hint*: Use Theorem 9.33 and equation (12.195).

References

[**1**] Abdulloev, K. O., Bogolubsky, I. L., and Makhankov, V. G., One more example of inelastic soliton interaction, *Phys. Lett. A* **56** (1976), 427–428.

[**2**] Ablowitz, M. J., and Clarkson, P. A., *Solitons, Nonlinear Evolution Equations and the Inverse Scattering Transform*, L.M.S. Lecture Notes in Math., vol. 149, Cambridge University Press, Cambridge, 1991.

[**3**] Abraham, R., Marsden, J. E., and Ratiu, T., *Manifolds, Tensor Analysis, and Applications*, Springer–Verlag, New York, 1988.

[**4**] Airy, G. B., On the intensity of light in the neighborhood of a caustic, *Trans. Cambridge Phil. Soc.* **6** (1838), 379–402.

[**5**] Aki, K., and Richards, P. G., *Quantitative Seismology*, W. H. Freeman, San Francisco, 1980.

[**6**] Ames, W. F., *Numerical Methods for Partial Differential Equations*, 3rd ed., Academic Press, New York, 1992.

[**7**] Antman, S. S., *Nonlinear Problems of Elasticity*, Appl. Math. Sci., vol. 107, Springer–Verlag, New York, 1995.

[**8**] Apostol, T. M., *Calculus*, Blaisdell Publishing Co., Waltham, Mass., 1967–1969.

[**9**] Apostol, T. M., *Introduction to Analytic Number Theory*, Springer–Verlag, New York, 1976.

[**10**] Atkinson, K., and Han, W., *Spherical Harmonics and Approximations on the Unit Sphere*: *An Introduction*, Lecture Notes in Math., vol. 2044, Springer, Berlin, 2012.

[**11**] Bank, S. B., and Kaufman, R. P., A note on Hölder's theorem concerning the gamma function, *Math. Ann.* **232** (1978), 115–120.

[**12**] Batchelor, G. K., *An Introduction to Fluid Dynamics*, Cambridge University Press, Cambridge, 1967.

[**13**] Bateman, H., Some recent researches on the motion of fluids, *Monthly Weather Rev.* **43** (1915), 63–170.

[**14**] Benjamin, T. B., Bona, J. L., and Mahony, J. J., Model equations for long waves in nonlinear dispersive systems, *Phil. Trans. Roy. Soc. London A* **272** (1972), 47–78.

[**15**] Berest, Y., and Winternitz, P., Huygens' principle and separation of variables, *Rev. Math. Phys.* **12** (2000), 159–180.

[**16**] Berry, M. V., Marzoli, I., and Schleich, W., Quantum carpets, carpets of light, *Physics World* **14**(6) (2001), 39–44.

[**17**] Birkhoff, G., *Hydrodynamics — A Study in Logic, Fact and Similitude*, 2nd ed., Princeton University Press, Princeton, 1960.

[**18**] Birkhoff, G., and Rota, G.–C., *Ordinary Differential Equations*, Blaisdell Publ. Co., Waltham, Mass., 1962.

P.J. Olver, *Introduction to Partial Differential Equations*, Undergraduate Texts in Mathematics,
DOI 10.1007/978-3-319-02099-0, © Springer International Publishing Switzerland 2014

[**19**] Black, F., and Scholes, M., The pricing of options and corporate liabilities, *J. Political Economy* **81** (1973), 637–654.

[**20**] Blanchard, P., Devaney, R. L., and Hall, G. R., *Differential Equations*, Brooks–Cole Publ. Co., Pacific Grove, Calif., 1998.

[**21**] Boussinesq, J., Théorie des ondes et des remous qui se propagent le long d'un canal rectangulaire horizontal, en communiquant au liquide contenu dans ce canal des vitesses sensiblement pareilles de la surface au fond, *J. Math. Pures Appl.* **17** (2) (1872), 55–108.

[**22**] Boussinesq, J., Essai sur la théorie des eaux courants, *Mém. Acad. Sci. Inst. Nat. France* **23** (1) (1877), 1–680.

[**23**] Boyce, W. E., and DiPrima, R. C., *Elementary Differential Equations and Boundary Value Problems*, 7th ed., John Wiley & Sons, Inc., New York, 2001.

[**24**] Bradie, B., *A Friendly Introduction to Numerical Analysis*, Prentice–Hall, Inc., Upper Saddle River, N.J., 2006.

[**25**] Bronstein, M., *Symbolic integration I*: *Transcendental Functions*, Springer–Verlag, New York, 1997.

[**26**] Burgers, J. M., A mathematical model illustrating the theory of turbulence, *Adv. Appl. Mech.* **1** (1948), 171–199.

[**27**] Cantwell, B. J., *Introduction to Symmetry Analysis*, Cambridge University Press, Cambridge, 2003.

[**28**] Carleson, L., On the convergence and growth of partial sums of Fourier series, *Acta Math.* **116** (1966), 135–157.

[**29**] Carmichael, R., *The Theory of Numbers*, Dover Publ., New York, 1959.

[**30**] Chen, G., and Olver, P. J., Dispersion of discontinuous periodic waves, *Proc. Roy. Soc. London* **469** (2012), 20120407.

[**31**] Coddington, E. A., and Levinson, N., *Theory of Ordinary Differential Equations*, McGraw–Hill, New York, 1955.

[**32**] Cole, J. D., On a quasilinear parabolic equation occurring in aerodynamics, *Q. Appl. Math.* **9** (1951), 225–236.

[**33**] Courant, R., Friedrichs, K. O., and Lewy, H., Über die partiellen Differenzengleichungen der mathematischen Physik, *Math. Ann.* **100** (1928), 32–74.

[**34**] Courant, R., and Hilbert, D., *Methods of Mathematical Physics*, vol. I, Interscience Publ., New York, 1953.

[**35**] Courant, R., and Hilbert, D., *Methods of Mathematical Physics*, vol. II, Interscience Publ., New York, 1953.

[**36**] Drazin, P. G., and Johnson, R. S., *Solitons*: *An Introduction*, Cambridge University Press, Cambridge, 1989.

[**37**] Dym, H., and McKean, H. P., *Fourier Series and Integrals*, Academic Press, New York, 1972.

[**38**] Evans, L. C., *Partial Differential Equations*, Grad. Studies Math. vol. 19, Amer. Math. Soc., Providence, R.I., 1998.

[**39**] Feller, W., *An Introduction to Probability Theory and Its Applications*, 3rd ed., J. Wiley & Sons, New York, 1968.

[**40**] Fermi, E., Pasta, J., and Ulam, S., Studies of nonlinear problems. I., preprint, Los Alamos Report LA 1940, 1955; in: *Nonlinear Wave Motion*, A. C. Newell, ed., Lectures in Applied Math., vol. 15, American Math. Soc., Providence, R.I., 1974, pp. 143–156.

[**41**] Forsyth, A. R., *The Theory of Differential Equations*, Cambridge University Press, Cambridge, 1890, 1900, 1902, 1906.

[**42**] Fourier, J., *The Analytical Theory of Heat*, Dover Publ., New York, 1955.

[**43**] Gander, M. J., and Kwok, F., Chladni figures and the Tacoma bridge: motivating PDE eigenvalue problems via vibrating plates, *SIAM Review* **54** (2012), 573–596.

[**44**] Garabedian, P., *Partial Differential Equations*, 2nd ed., Chelsea Publ. Co., New York, 1986.

[**45**] Gardner, C. S., Greene, J. M., Kruskal, M. D., and Miura, R. M., Method for solving the Korteweg–deVries equation, *Phys. Rev. Lett.* **19** (1967), 1095–1097.

[**46**] Gonzalez, R. C., and Woods, R. E., *Digital Image Processing*, 2nd ed., Prentice–Hall, Inc., Upper Saddle River, N.J., 2002.

[**47**] Gordon, C., Webb, D. L., and Wolpert, S., One cannot hear the shape of a drum, *Bull. Amer. Math. Soc.* **27** (1992), 134–138.

[**48**] Gradshteyn, I. S., and Ryzhik, I. W., *Table of Integrals, Series and Products*, Academic Press, New York, 1965.

[**49**] Gurtin, M. E., *An Introduction to Continuum Mechanics*, Academic Press, New York, 1981.

[**50**] Haberman, R., *Elementary Applied Partial Differential Equations*, 3rd ed., Prentice–Hall, Inc., Upper Saddle River, NJ, 1998.

[**51**] Hairer, E., Lubich, C., and Wanner, G., *Geometric Numerical Integration*, Springer–Verlag, New York, 2002.

[**52**] Hale, J. K., *Ordinary Differential Equations*, 2nd ed., R.E. Krieger Pub. Co., Huntington, N.Y., 1980.

[**53**] Henrici, P., *Applied and Computational Complex Analysis*, vol. 1, J. Wiley & Sons, New York, 1974.

[**54**] Hille, E., *Ordinary Differential Equations in the Complex Domain*, John Wiley & Sons, New York, 1976.

[**55**] Hobson, E. W., *The Theory of Functions of a Real Variable and the Theory of Fourier's Series*, Dover Publ., New York, 1957.

[**56**] Hopf, E., The partial differential equation $u_t + u\,u_x = \mu\,u$, *Commun. Pure Appl. Math.* **3** (1950), 201–230.

[**57**] Howison, S., *Practical Applied Mathematics*: *Modelling, Analysis, Approximation*, Cambridge University Press, Cambridge, 2005.

[**58**] Hydon, P. E., *Symmetry Methods for Differential Equations*, Cambridge Texts in Appl. Math., Cambridge University Press, Cambridge, 2000.

[**59**] Ince, E. L., *Ordinary Differential Equations*, Dover Publ., New York, 1956.

[**60**] Iserles, A., *A First Course in the Numerical Analysis of Differential Equations*, Cambridge University Press, Cambridge, 1996.

[**61**] Jost, J., *Partial Differential Equations*, Graduate Texts in Mathematics, vol. 214, Springer–Verlag, New York, 2007.

[**62**] Kamke, E., *Differentialgleichungen Lösungsmethoden und Lösungen*, vol. 1, Chelsea, New York, 1971.

[**63**] Keller, H. B., *Numerical Methods for Two-Point Boundary-Value Problems*, Blaisdell, Waltham, MA, 1968.

[**64**] Knobel, R., *An Introduction to the Mathematical Theory of Waves*, American Mathematical Society, Providence, RI, 2000.

[**65**] Korteweg, D. J., and de Vries, G., On the change of form of long waves advancing in a rectangular channel, and on a new type of long stationary waves, *Phil. Mag.* (5) **39** (1895), 422–443.

[**66**] Landau, L. D., and Lifshitz, E. M., *Quantum Mechanics* (*Non-relativistic Theory*), Course of Theoretical Physics, vol. 3, Pergamon Press, New York, 1977.

[**67**] Levine, I. N., *Quantum Chemistry*, 5th ed., Prentice–Hall, Inc., Upper Saddle River, N.J., 2000.

[**68**] Lighthill, M. J., *Introduction to Fourier Analysis and Generalised Functions*, Cambridge University Press, Cambridge, 1970.

[**69**] Lin, C. C., and Segel, L. A., *Mathematics Applied to Deterministic Problems in the Natural Sciences*, SIAM, Philadelphia, 1988.

[**70**] McOwen, R. C., *Partial Differential Equations*: *Methods and Applications*, Prentice–Hall, Inc., Upper Saddle River, N.J., 2002.

[**71**] Merton, R. C., Theory of rational option pricing, *Bell J. Econ. Management Sci.* **4** (1973), 141–183.

[**72**] Messiah, A., *Quantum Mechanics*, John Wiley & Sons, New York, 1976.

[**73**] Miller, W., Jr., *Symmetry and Separation of Variables*, Encyclopedia of Mathematics and Its Applications, vol. 4, Addison–Wesley Publ. Co., Reading, Mass., 1977.

[**74**] Milne–Thompson, L. M., *The Calculus of Finite Differences*, Macmillan and Co., Ltd., London, 1951.

[**75**] Misner, C. W., Thorne, K. S., and Wheeler, J. A., *Gravitation*, W. H. Freeman, San Francisco, 1973.

[**76**] Miura, R. M., Gardner, C. S., and Kruskal, M. D., Korteweg–deVries equation and generalizations. II. Existence of conservation laws and constants of the motion, *J. Math. Phys.* **9** (1968), 1204–1209.

[**77**] Moon, F. C., *Chaotic Vibrations*, John Wiley & Sons, New York, 1987.

[**78**] Moon, P., and Spencer, D. E., *Field Theory Handbook*, Springer-Verlag, New York, 1971.

[**79**] Morse, P. M., and Feshbach, H., *Methods of Theoretical Physics*, McGraw–Hill, New York, 1953.

[**80**] Morton, K. W., and Mayers, D. F., *Numerical Solution of Partial Differential Equations*, 2nd ed., Cambridge University Press, Cambridge, 2005.

[**81**] Murray, J. D., *Mathematical Biology*, 3rd ed., Springer-Verlag, New York, 2002–2003.

[**82**] Oberhettinger, F., *Tables of Fourier Transforms and Fourier Transforms of Distributions*, Springer-Verlag, New York, 1990.

[**83**] Øksendal, B., *Stochastic Differential Equations*: *An Introduction with Applications*, Springer–Verlag, New York, 1985.

[**84**] Okubo, A., *Diffusion and Ecological Problems*: *Mathematical Models*, Springer-Verlag, New York, 1980.

[**85**] Olver, F. W. J., *Asymptotics and Special Functions*, Academic Press, New York, 1974.

[**86**] Olver, F. W. J., Lozier, D. W., Boisvert, R. F., and Clark, C. W., eds., *NIST Handbook of Mathematical Functions*, Cambridge University Press, Cambridge, 2010.

[**87**] Olver, P. J., *Applications of Lie Groups to Differential Equations*, 2nd ed., Graduate Texts in Mathematics, vol. 107, Springer–Verlag, New York, 1993.

[**88**] Olver, P. J., Dispersive quantization, *Amer. Math. Monthly* **117** (2010), 599–610.

[**89**] Olver, P. J., and Shakiban, C., *Applied Linear Algebra*, Prentice–Hall, Inc., Upper Saddle River, N.J., 2005.

[**90**] Oskolkov, K. I., A class of I. M. Vinogradov's series and its applications in harmonic analysis, *in*: *Progress in Approximation Theory* , Springer Ser. Comput. Math., 19, Springer, New York, 1992, pp. 353–402.

[**91**] Pinchover, Y., and Rubinstein, J., *An Introduction to Partial Differential Equations*, Cambridge University Press, Cambridge, 2005.

[**92**] Pinsky, M. A., *Partial Differential Equations and Boundary–Value Problems with Applications*, 3rd ed., McGraw–Hill, New York, 1998.

[**93**] Polyanin, A. D., and Zaitsev, V. F., *Handbook of Exact Solutions for Ordinary Differential Equations*, 2nd ed., Chapman & Hall/CRC, Boca Raton, Fl., 2003.

[**94**] Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P., *Numerical Recipes*: *The Art of Scientific Computing*, 3rd ed., Cambridge University Press, Cambridge, 2007.

[**95**] Reed, M., and Simon, B., *Methods of Modern Mathematical Physics*, Academic Press, New York, 1972.

[**96**] Royden, H. L., and Fitzpatrick, P. M., *Real Analysis*, 4th ed., Pearson Education Inc., Boston, MA, 2010.

[**97**] Rudin, W., *Principles of Mathematical Analysis*, 3rd ed., McGraw–Hill, New York, 1976.

[**98**] Rudin, W., *Real and Complex Analysis*, 3rd ed., McGraw–Hill, New York, 1987.

[**99**] Salsa, S., *Partial Differential Equations in Action*: *From Modelling to Theory*, Springer–Verlag, New York, 2008.

[**100**] Sapiro, G., *Geometric Partial Differential Equations and Image Analysis*, Cambridge University Press, Cambridge, 2001.

[**101**] Schrödinger, E., *Collected Papers on Wave Mechanics*, Chelsea Publ. Co., New York, 1982.

[**102**] Schumaker, L. L., *Spline Functions*: *Basic Theory*, John Wiley & Sons, New York, 1981.

[**103**] Schwartz, L., *Théorie des distributions*, Hermann, Paris, 1957.

[**104**] Scott Russell, J., On waves, *in*: *Report of the 14*$^{\text{th}}$ *Meeting*, British Assoc. Adv. Sci., 1845, pp. 311–390.

[**105**] Sethares, W. A., *Tuning, Timbre, Spectrum, Scale*, Springer–Verlag, New York, 1999.

[**106**] Siegel, C. L., Über einige Anwendungen diophantischer Approximationen, *in*: *Gesammelte Abhandlungen*, vol. 1, Springer–Verlag, New York, 1966, pp. 209–266.

[**107**] Smoller, J., *Shock Waves and Reaction–Diffusion Equations*, 2nd ed., Springer-Verlag, New York, 1994.

[**108**] Stewart, J., *Calculus*: *Early Transcendentals*, vols. 1 & 2, 7th ed., Cengage Learning, Mason, OH, 2012.

[**109**] Stokes, G. G., On a difficulty in the theory of sound, *Phil. Mag.* **33**(3) (1848), 349–356.

[**110**] Stokes, G. G., *Mathematical and Physical Papers*, Cambridge University Press, Cambridge, 1880–1905.

[**111**] Stokes, G. G., *Mathematical and Physical Papers*, 2nd ed., Johnson Reprint Corp., New York, 1966.

[**112**] Strang, G., *Introduction to Applied Mathematics*, Wellesley Cambridge Press, Wellesley, Mass., 1986.

[**113**] Strang, G., and Fix, G. J., *An Analysis of the Finite Element Method*, Prentice–Hall, Inc., Englewood Cliffs, N.J., 1973.

[**114**] Strauss, W. A., *Partial Differential Equations*: *An Introduction*, John Wiley & Sons, New York, 1992.

[**115**] Thaller, B., *Visual Quantum Mechanics*, Springer–Verlag, New York, 2000.

[**116**] Thijssen, J., *Computational Physics*, Cambridge University Press, Cambridge, 1999.

[**117**] Titchmarsh, E. C., *Theory of Functions*, Oxford University Press, London, 1968.

[**118**] Varga, R. S., *Matrix Iterative Analysis*, 2nd ed., Springer–Verlag, New York, 2000.

[**119**] Watson, G. N., *A Treatise on the Theory of Bessel Functions*, Cambridge University Press, Cambridge, 1952.

[**120**] Weinberger, H. F., *A First Course in Partial Differential Equations*, Dover Publ., New York, 1995.

[**121**] Wiener, N., *I Am a Mathematician*, Doubleday, Garden City, N.Y., 1956.

[**122**] Whitham, G. B., *Linear and Nonlinear Waves*, John Wiley & Sons, New York, 1974.

[**123**] Wilmott, P., Howison, S., and Dewynne, J., *The Mathematics of Financial Derivatives*, Cambridge University Press, Cambridge, 1995.

[**124**] Yong, D., Strings, chains, and ropes, *SIAM Review* **48** (2006), 771–781.

[**125**] Zabusky, N. J., and Kruskal, M. D., Interaction of "solitons" in a collisionless plasma and the recurrence of initial states, *Phys. Rev. Lett.* **15** (1965), 240–243.

[**126**] Zienkiewicz, O. C., and Taylor, R. L., *The Finite Element Method*, 4th ed., McGraw–Hill, New York, 1989.

[**127**] Zwillinger, D., *Handbook of Differential Equations*, Academic Press, Boston, 1992.

[**128**] Zygmund, A., *Trigonometric Series*, 3rd ed., Cambridge University Press, Cambridge, 2002.

Symbol Index

Symbol	Meaning	Page(s)
$c + d$	addition of scalars	575
$z + w$	complex addition	571
$A + B$	addition of matrices	575
$\mathbf{v} + \mathbf{w}$	addition of vectors	575
$f + g$	addition of functions	575
$z\,w$	complex multiplication	571
z/w	complex division	572
$c\,\mathbf{v},\ c\,A,\ c\,f$	scalar multiplication	575
$\overline{z}$	complex conjugate	571
$\overline{\Omega}$	closure of subset or domain	243
$\mathbf{0}$	zero vector	xvi, 575
> 0	positive definite	355, 578
≥ 0	positive semi-definite	355
f^{-1}	inverse function	xvi
A^{-1}	inverse matrix	xvi
$f(x^+),\ f(x^-)$	one-sided limits	xvi
$n!$	factorial	163, 453
$\binom{n}{k}$	binomial coefficient	163
$\lvert\,\cdot\,\rvert$	absolute value, modulus	94, 225, 571
$\lVert\,\cdot\,\rVert$	norm	73, 89, 106, 284, 356, 578, 579, 581
$\lVert\,\cdot\,\rVert$	double norm	380
$\lvert\lvert\lvert\,\cdot\,\rvert\rvert\rvert$	norm	356
$\mathbf{v} \cdot \mathbf{w}$	dot product	578
$\mathbf{z} \cdot \mathbf{w}$	Hermitian dot product	580
$\langle\,\cdot\,\rangle$	expected value	287
$\langle\,\cdot\,,\cdot\,\rangle$	inner product	73, 89, 107, 285, 341, 578, 579, 581
$\langle\!\langle\,\cdot\,,\cdot\,\rangle\!\rangle$	inner product	341
$[0, 1]$	closed interval	xvi
$\{\ f \mid C\ \}$	set	xvi
$\in$	element of	xvi

P.J. Olver, *Introduction to Partial Differential Equations*, Undergraduate Texts in Mathematics,
DOI 10.1007/978-3-319-02099-0, © Springer International Publishing Switzerland 2014

span	span	576
supp	support	407
t	time	xv, 3
T	conserved density	38, 256
A^T	transpose of matrix	341, 578
T_ν	finite element triangle	411
tan	tangent	1
tanh	hyperbolic tangent	135
u	dependent variable	xv, 3
$u_x, u_{xx}, \dots$	partial derivative	1
v	dependent variable	xv, 3
v	eigenvector/eigenfunction	371
$\mathbf{v}$	vector	xv, 575
$\mathbf{v}$	eigenvector	66, 582
$\mathbf{v}$	vector field	3, 242
V	vector space	575
V	potential function	6
$\mathbf{v}^\perp$	perpendicular vector	244
$v_{l\,m\,n}$	atomic eigenfunction	568
V_λ	eigenspace	371
w	dependent variable	xv, 3
w	heat flux	122
$\mathbf{w}$	heat flux vector	437
x	Cartesian space coordinate	xv, 3, 152, 504
x	real part of complex number	571
X	flux	38, 256
y	Cartesian space coordinate	xv, 3, 152, 504
y	imaginary part of complex number	571
Y	flux	256
Y_m	Bessel function of the second kind	470
$Y_n^m,\ \widetilde{Y}_n^m$	spherical harmonic	517
$\mathcal{Y}_n^m$	complex spherical harmonic	519
z	Cartesian space coordinate	xv, 3, 504
z	cylindrical coordinate	xv, 3, 508
z	complex number	571
$\mathbb{Z}$	integers	xv
α	electron charge	564
β_l^n	radial wave function	568
γ	thermal diffusivity	124, 438, 535
γ	Euler–Mascheroni constant	471
Γ	gamma function	454

ψ	time-dependent wave function	394, 564
ω	frequency	59, 330
Ω	domain	152, 242, 504