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 $\mathbf{x} = (x_1, \ldots, x_n)$ 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

$$\widehat{u}(t,k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u(t,x)\, e^{-\,\mathrm{i}\,k\,x}\, dx \tag{8.10}$$

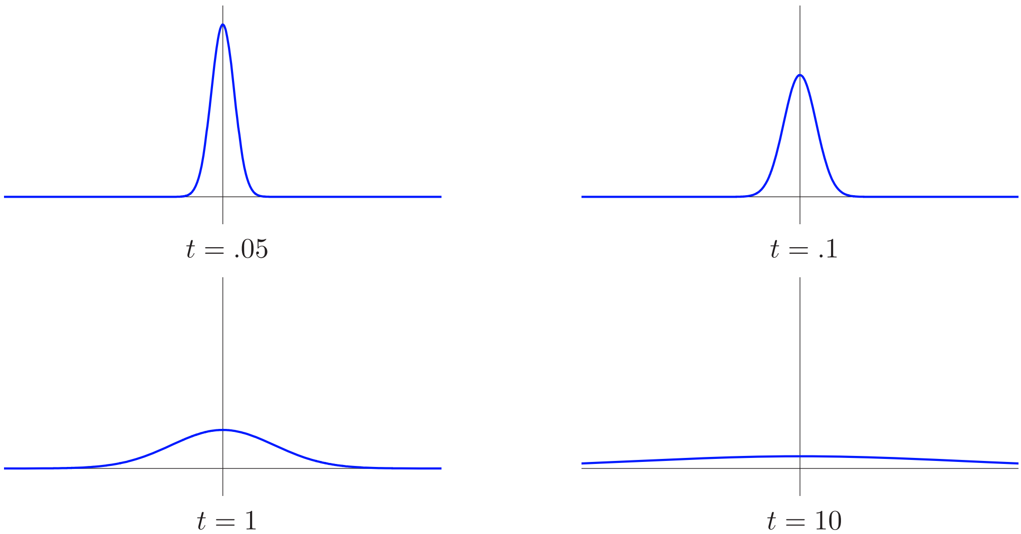

Figure 8.1. The fundamental solution to the one-dimensional heat equation.

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

$$u(t,x) = \frac{1}{2\sqrt{\pi\,t}} \int_{-\infty}^{\infty} e^{-(x-\xi)^2/(4\,t)}\, f(\xi)\,d\xi \tag{8.16}$$

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:

$$u(t,x) = F_0(t,x) * f(x), \qquad \text{where} \qquad F_0(t,x) = F(t,x;0) = \frac{e^{-x^2/(4\,t)}}{2\sqrt{\pi\,t}}\,.$$

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)

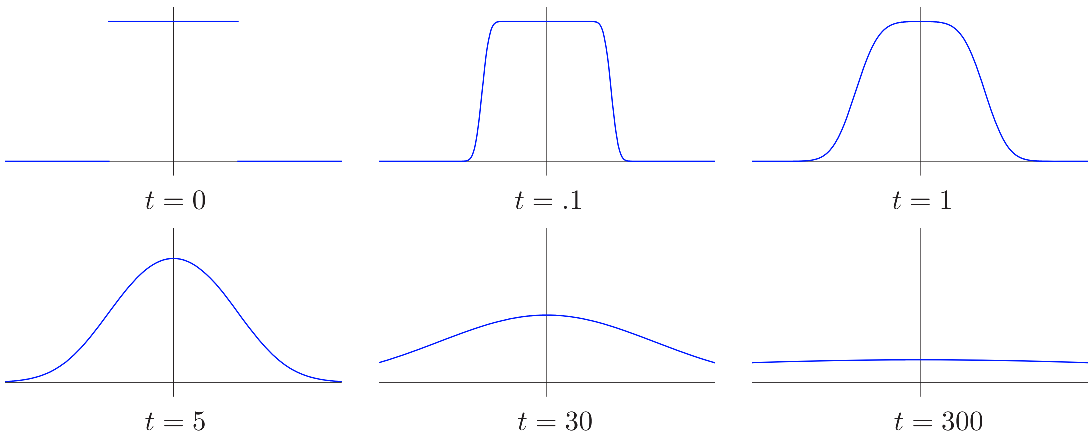

Figure 8.2. Error function solution to the heat equation.

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):

$$u(t,x) = \frac{1}{2\sqrt{\pi\, t}} \int_a^b e^{-(x-\xi)^2/(4t)}\, d\xi = \frac{1}{2}\left[\operatorname{erf}\left(\frac{x-a}{2\sqrt{t}}\right) - \operatorname{erf}\left(\frac{x-b}{2\sqrt{t}}\right) \right], \tag{8.17}$$

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

$$\frac{\partial \widehat{u}}{\partial t} + k^2\, \widehat{u} = \frac{1}{\sqrt{2\pi}}\, e^{-\,i\,k\,\xi}\, \delta(t - \tau), \tag{8.26}$$

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

$$\begin{aligned} u(t,x) &= \int_0^t \int_{-\infty}^{\infty} \frac{\delta(\xi)}{2\sqrt{\pi(t-\tau)}}\, \exp\left[-\, \frac{(x-\xi)^2}{4\,(t-\tau)} \right] d\xi\, d\tau \\ &= \int_0^t \frac{1}{2\sqrt{\pi(t-\tau)}}\, \exp\left[-\, \frac{x^2}{4\,(t-\tau)} \right] d\tau = \sqrt{\frac{t}{\pi}}\, \exp\left[-\, \frac{x^2}{4\,t} \right] + \frac{x \operatorname{erf}\left(\dfrac{x}{2\sqrt{t}} \right) - |\,x\,|}{2}\,. \end{aligned}$$

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

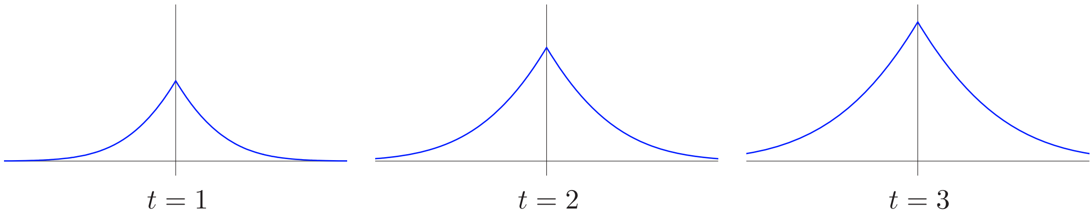

Figure 8.3. Effect of a concentrated heat source.

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

$$\begin{aligned} z(\tau, y) = \frac{1}{2}\bigg[\, & e^{(\kappa+1)^2\tau/4 + (\kappa+1)y/2}\, \mathrm{erfc}\left(\frac{\log p - (\kappa+1)\tau - y}{2\sqrt{\tau}}\right) \\ & - p\, e^{(\kappa-1)^2\tau/4 + (\kappa-1)y/2}\, \mathrm{erfc}\left(\frac{\log p - (\kappa-1)\tau - y}{2\sqrt{\tau}}\right)\bigg], \end{aligned} \tag{8.42}$$

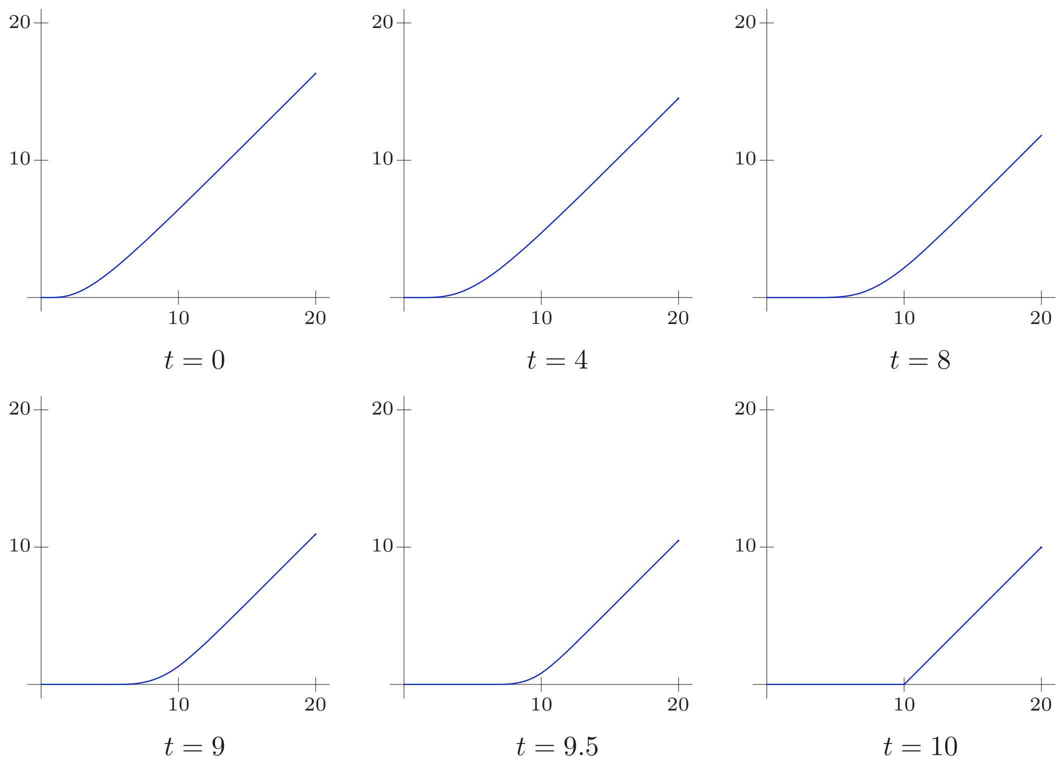

Figure 8.4. Solution to the Black–Scholes equation.

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:

$$\begin{aligned} u(t,x) = \frac{1}{2} \Bigg[\, x \operatorname{erfc}\left(- \frac{\left(r + \frac{1}{2}\sigma^2\right)(t_\star - t) + \log(x/p)}{\sqrt{2\,\sigma^2 (t_\star - t)}} \right) \\ - p\, e^{-r(t_\star - t)} \operatorname{erfc}\left(- \frac{\left(r - \frac{1}{2}\sigma^2\right)(t_\star - t) + \log(x/p)}{\sqrt{2\,\sigma^2 (t_\star - t)}} \right) \Bigg]. \end{aligned} \tag{8.44}$$

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, $E(t) = \int_{-\infty}^{\infty} u(t,x)\, dx < \infty$, 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 $\int_{-\infty}^{\infty} u(t,x)\, dx = 1$ 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,

$$t \longmapsto t+a, \qquad\qquad x \longmapsto x+b, \tag{8.45}$$

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

$$t \longmapsto \beta^2\, t, \qquad x \longmapsto \beta\, x, \tag{8.52}$$

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 $a \le x \le b$, then $U(t,x)$, as given in (8.53), is defined for $\beta\, a \le x \le \beta\, b$ (or, when $\beta < 0$, for $\beta\, b \le x \le \beta\, a$).

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 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

$$\begin{aligned} u_t &= -\tfrac{1}{2}\, x\, t^{c/2-3/2}\, v'(\xi) + \tfrac{1}{2}\, c t^{c/2-1}\, v(\xi) = t^{c/2-1}\big[-\tfrac{1}{2}\,\xi\, v'(\xi) + \tfrac{1}{2}\, c v(\xi)\big], \\ u_{xx} &= t^{c/2-1}\, v''(\xi). \end{aligned}$$

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 $U(t,x) = \dfrac{e^{-c x^2/(4(1+ct))}}{\sqrt{1+ct}}\, u\left(\dfrac{t}{1+ct}, \dfrac{x}{1+ct}\right)$ 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 $(t,x) \in \widehat{R} = \{\, a < x < b,\ 0 < t \le c \,\}$ 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 $m = 0 \le v(t,x) \le 0 = M$ for $t = 0$, $a \le x \le b$, and also $x = a$ or b, $0 \le t \le c$. But then Corollary 8.7 implies that $0 \le v(t,x) \le 0$ 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 $u \le v$ when $t = 0$ and when $x = a$ or $x = b$. Prove that $u(t,x) \le v(t,x)$ for all $a \le x \le b$ and all $t \ge 0$. Provide a physical interpretation 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 $M(t) = \max\{\, u(t,x) \mid a \le x \le b \,\}$ 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 $0 < x < \ell$ when subject to either 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

$$u(t,x) = v(\xi) = v(x - c\,t), \qquad \text{where} \qquad \xi = x - c\,t, \tag{8.72}$$

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

$$v(\xi) = \frac{a\,e^{(b-a)(\xi-\delta)/(2\,\gamma)} + b}{e^{(b-a)(\xi-\delta)/(2\,\gamma)} + 1}\,,$$

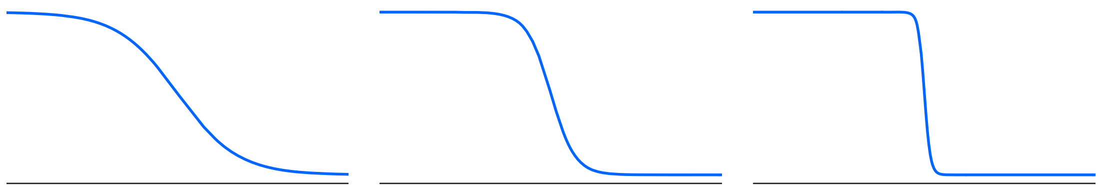

Figure 8.5. Traveling-wave solutions to Burgers' equation.

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*.

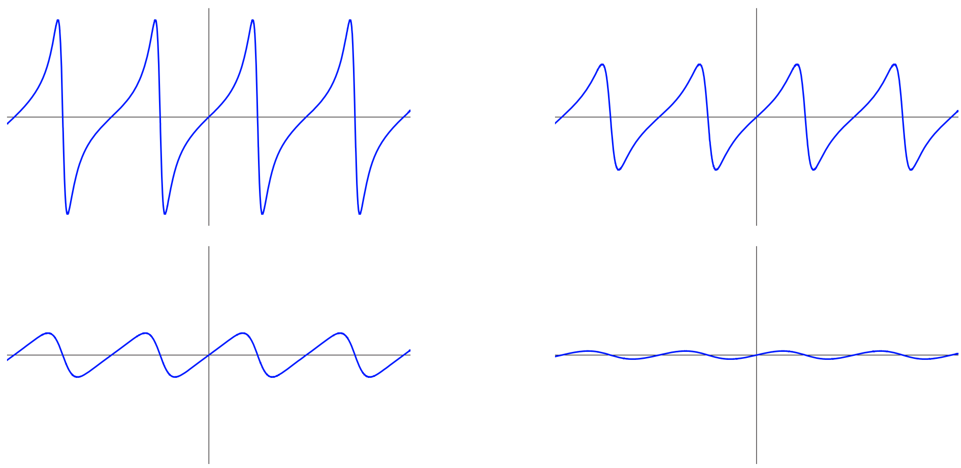

Figure 8.6. Trignometric solution to Burgers' equation.

Theorem 8.9. *If $v(t,x) > 0$ is any positive solution to the linear heat equation $v_t = \gamma\, v_{xx}$, then*

$$u(t,x) = \frac{\partial}{\partial x}\left[\, -2\gamma \log v(t,x)\,\right] = -2\gamma\, \frac{v_x}{v} \tag{8.80}$$

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

$$u(t,x) = -\frac{2\gamma}{\widehat{v}(t,x)}\,\frac{\partial\widehat{v}}{\partial x}\,, \qquad \text{where} \qquad \left\{ \begin{aligned} \widehat{v}(t,x) &= \int_{-\infty}^{\infty} e^{-H(t,x;\xi)}\,d\xi, \\ H(t,x;\xi) &= \frac{(x-\xi)^2}{4\gamma t} + \frac{1}{2\gamma}\int_0^\xi f(\eta)\,d\eta. \end{aligned} \right. \tag{8.84}$$

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,

$$H(t,x;\xi) = \frac{(x-\xi)^2}{4\gamma t} + \begin{cases} \dfrac{a\,\xi}{2\gamma}\,, & \xi < 0, \\[2ex] \dfrac{b\,\xi}{2\gamma}\,, & \xi > 0. \end{cases} \tag{8.86}$$

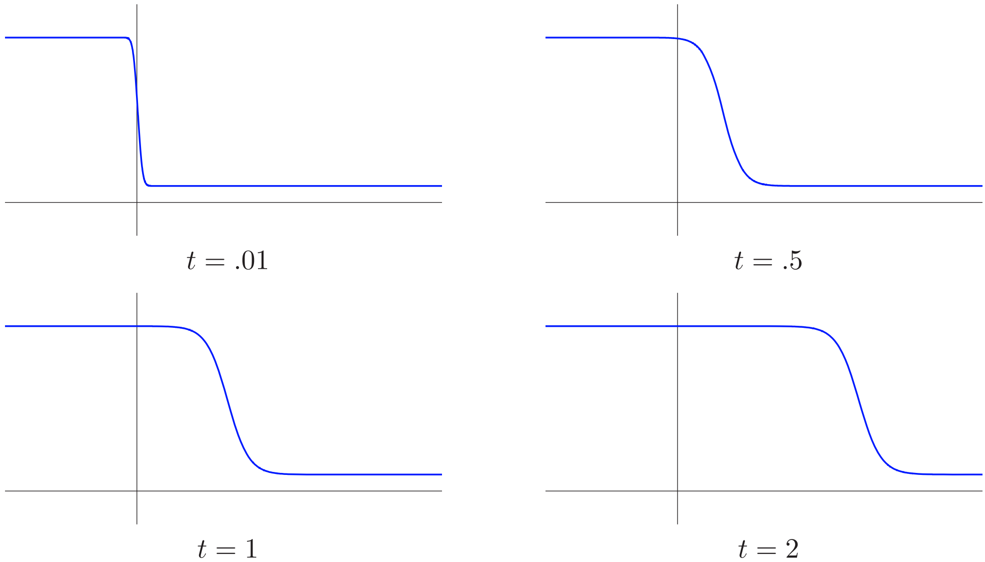

Figure 8.7. Shock-wave solution to Burgers' equation.

After some algebraic manipulations, the solution (8.84) is found to have the explicit form

$$u(t,x) = a + \frac{b-a}{1 + \exp\left(\dfrac{b-a}{2\gamma}\,(x-c\,t)\right) \operatorname{erfc}\left(\dfrac{x-a\,t}{2\sqrt{\gamma\,t}}\right) \Big/ \operatorname{erfc}\left(\dfrac{b\,t-x}{2\sqrt{\gamma\,t}}\right)}, \tag{8.87}$$

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 $x < b\,t$, to 1 when $b\,t < x < a\,t$, and to 0 when $x > a\,t$. On the other hand, since $a > b$, the exponential term tends to ∞ when $x < c\,t$, and to 0 when $x > c\,t$. Put together, these imply that the solution $u(t,x) \to a$ when $x < c\,t$, while $u(t,x) \to b$, when $x > c\,t$, 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

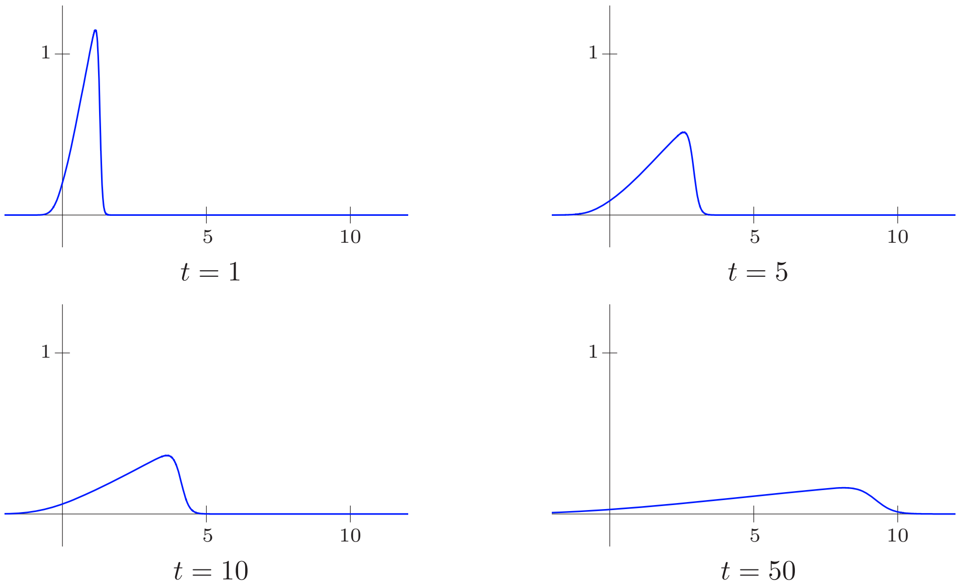

Figure 8.8. Triangular-wave solution to Burgers' equation.

other starting point, e.g., $-\infty$. Thus, we take

$$H(t,x;\xi) = \frac{(x-\xi)^2}{4\gamma t} + \frac{1}{2\gamma}\int_{-\infty}^{\xi} \delta(\eta)\,d\eta = \begin{cases} \dfrac{(x-\xi)^2}{4\gamma t}, & \xi < 0, \\[2mm] \dfrac{1}{2\gamma} + \dfrac{(x-\xi)^2}{4\gamma t}, & \xi > 0. \end{cases}$$

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 $\lim\limits_{z\to\infty} z\, e^{z^2} \operatorname{erfc} z = 1/\sqrt{\pi}$. (*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 $U(t,x) = u(t, x - c t) + c$ 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.

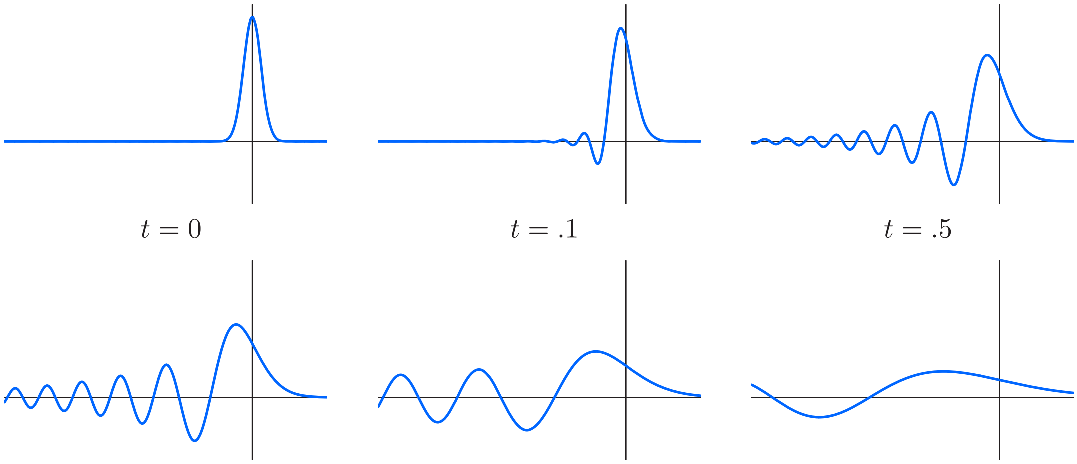

Figure 8.9. Gaussian solution to the dispersive wave equation.

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

$$u(t,x) = \frac{1}{2\sqrt{\pi}} \int_{-\infty}^{\infty} e^{\,\mathrm{i}\,(kx+k^3 t)-k^2/4}\, dk = \frac{1}{2\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-k^2/4} \cos(kx+k^3 t)\, dk;$$

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 $M = \int_{-\infty}^{\infty} u(t,x)\, dx$ and the energy $E = \int_{-\infty}^{\infty} u(t,x)^2\, dx$ 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

$$u(t,x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{\,\mathrm{i}\,(kx+k^3 t)}\, dk = \frac{1}{\pi} \int_{0}^{\infty} \cos(kx+k^3 t)\, dk, \tag{8.95}$$

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:

$$\begin{aligned} \int_0^l \cos(kx+k^3 t)\, dk &= \int_0^l \frac{1}{x+3k^2 t} \frac{d}{dk} \sin(kx+k^3 t)\, dk \qquad (8.96) \\ &= \frac{\sin(kx+k^3 t)}{x+3k^2 t} \bigg|_{k=0}^{l} - \int_0^l \frac{d}{dk} \left(\frac{1}{x+3k^2 t} \right) \sin(kx+k^3 t)\, dk \\ &= \frac{\sin(lx+l^3 t)}{x+3l^2 t} + \int_0^l \frac{6kt \sin(kx+k^3 t)}{(x+3k^2 t)^2}\, dk. \end{aligned}$$

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.

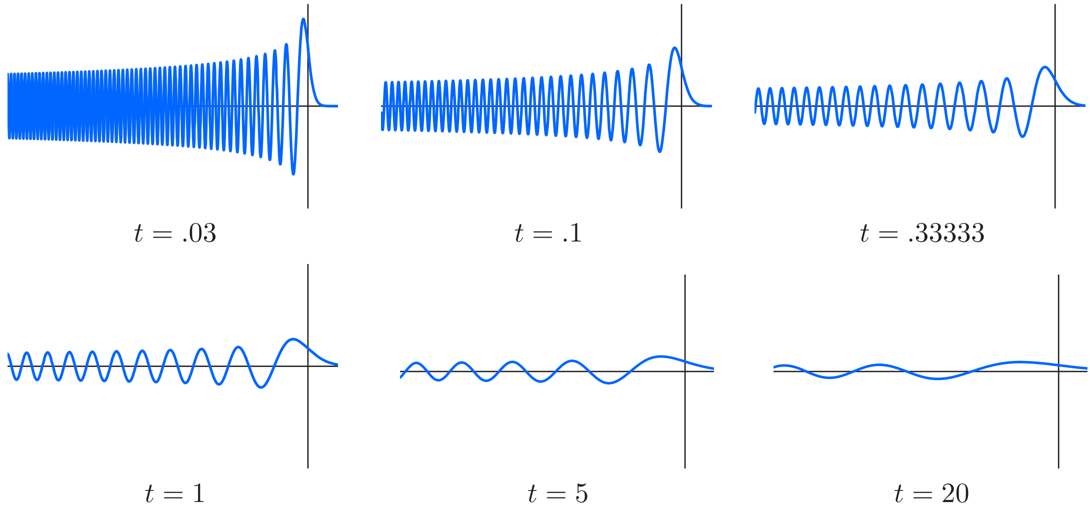

Figure 8.10. Fundamental solution to the dispersive wave equation.

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*

$$\operatorname{Ai}(z) = \frac{1}{\pi} \int_0^\infty \cos\left(s\,z + \tfrac{1}{3}\,s^3\right) ds, \tag{8.97}$$

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

$$s = k\,\sqrt[3]{3\,t}\,, \qquad\qquad z = \frac{x}{\sqrt[3]{3\,t}}\,,$$

to the Airy function integral (8.97), we deduce that the fundamental solution to the dispersive wave equation (8.90) can be written as

$$u(t,x) = \frac{1}{\sqrt[3]{3\,t}}\,\operatorname{Ai}\left(\frac{x}{\sqrt[3]{3\,t}}\right). \tag{8.98}$$

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 implpuse 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

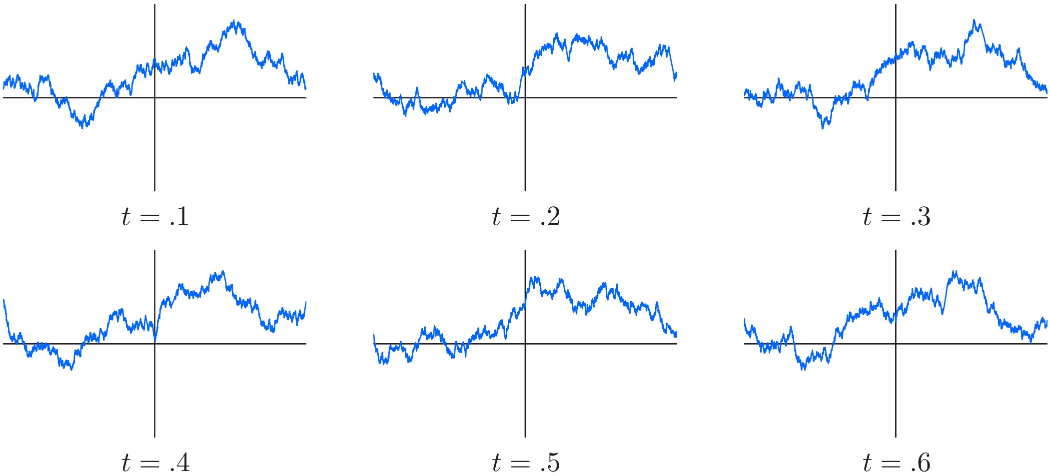

Figure 8.11. Periodic dispersion at irrational (with respect to π) times.

origin. By translation invariance, we immediately deduce that

$$F(t,x;\xi) = \frac{1}{\sqrt[3]{3\,t}}\ \mathrm{Ai}\left(\frac{x-\xi}{\sqrt[3]{3\,t}}\right)$$

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:

$$u(t,x) = \frac{1}{\sqrt[3]{3\,t}} \int_{-\infty}^{\infty} f(\xi)\ \mathrm{Ai}\left(\frac{x-\xi}{\sqrt[3]{3\,t}}\right) d\xi. \tag{8.99}$$

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 $f(x)$.

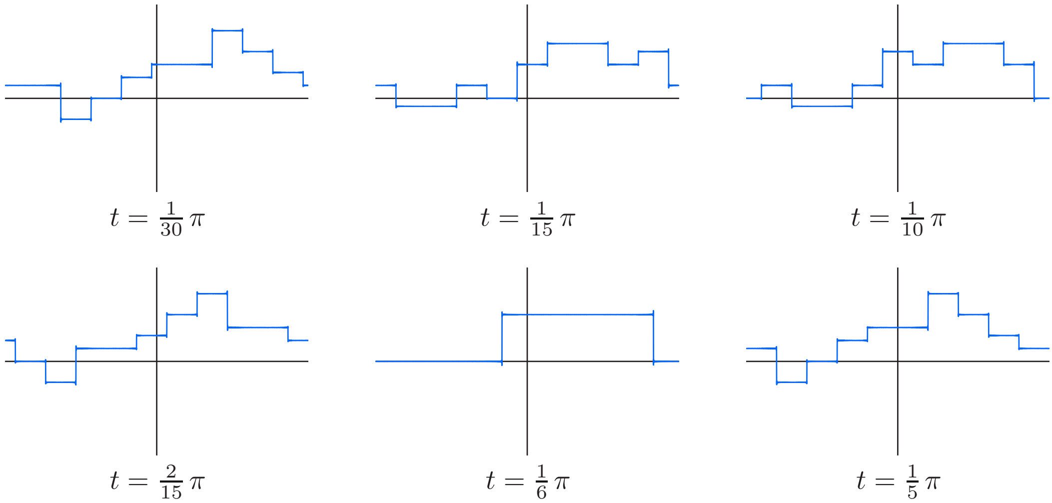

Figure 8.12. Periodic dispersion at rational (with respect to π) times.

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

$$u = v(\xi) = v(x - c\,t), \qquad \text{where} \qquad \xi = x - c\,t,$$

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

$$\frac{d}{d\xi}\left(v'' + \tfrac{1}{2}v^2 - c\,v\right) = 0, \qquad \text{and hence} \qquad v'' + \tfrac{1}{2}v^2 - c\,v = a,$$

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:

$$0 = v'\left(v'' + \tfrac{1}{2}v^2 - c\,v\right) = \frac{d}{d\xi}\left[\,\tfrac{1}{2}\,(v')^2 + \tfrac{1}{6}v^3 - \tfrac{1}{2}c\,v^2\,\right] = 0.$$

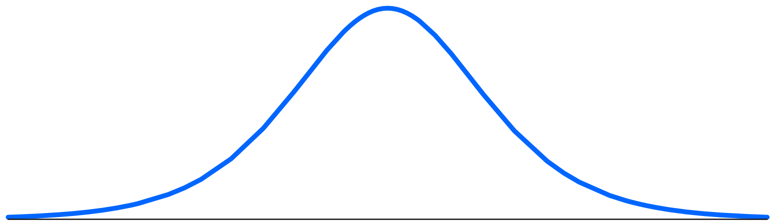

Figure 8.13. Solitary wave/soliton.

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:

$$\int \frac{dv}{v\sqrt{c - \tfrac{1}{3}v}} = \xi + \delta,$$

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

$$v(\xi) = 3c\,\mathrm{sech}^2\big(\tfrac{1}{2}\sqrt{c}\,\xi + \delta\big), \tag{8.116}$$

where

$$\mathrm{sech}\, y = \frac{1}{\cosh y} = \frac{2}{e^y + e^{-y}}$$

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

$$u(t,x) = 3c\,\mathrm{sech}^2\left[\tfrac{1}{2}\sqrt{c}\,(x - ct) + \delta\right], \tag{8.117}$$

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

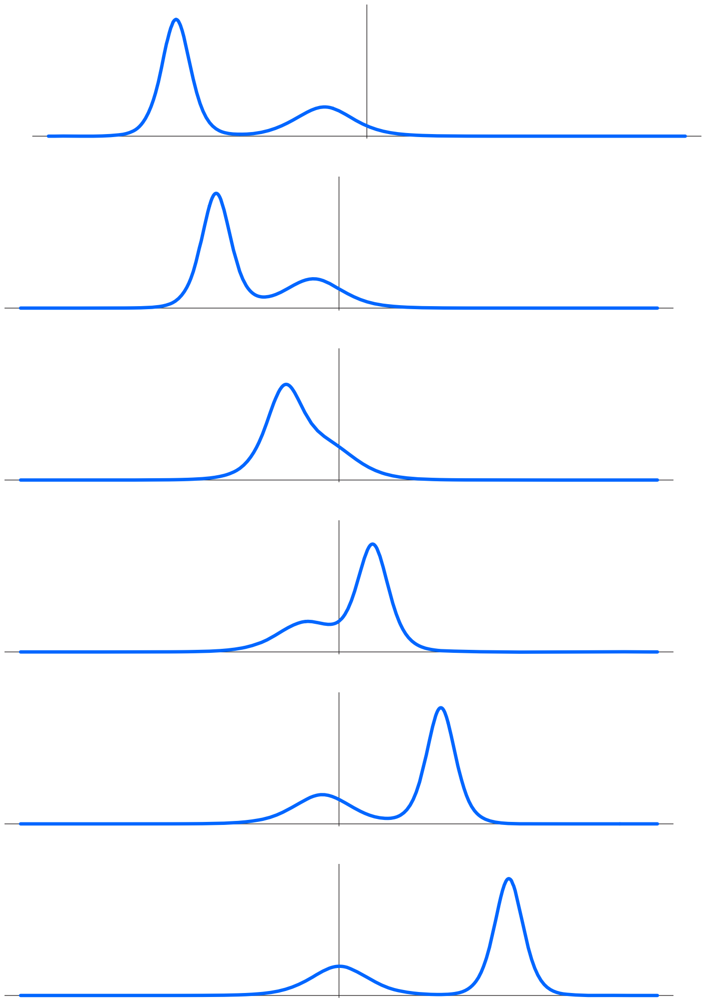

Figure 8.14. Interaction of two solitons.

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

$$\Delta(t,x) = \det\begin{pmatrix} 1+\varepsilon_1(t,x) & \dfrac{2\,b_1}{b_1+b_2}\,\varepsilon_2(t,x) \\ \dfrac{2\,b_2}{b_1+b_2}\,\varepsilon_1(t,x) & 1+\varepsilon_2(t,x) \end{pmatrix}, \tag{8.119}$$

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

$$\lim_{t\to\infty} \widetilde{u}(t,\xi) = \begin{cases} 3c_1 \operatorname{sech}^2\left[\,\tfrac12\sqrt{c_1}\,\xi + \delta_1\,\right], & c = c_1, \\ 3c_2 \operatorname{sech}^2\left[\,\tfrac12\sqrt{c_2}\,\xi + \delta_2\,\right], & c = c_2, \\ 0, & \text{otherwise}, \end{cases}$$

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 a solitary wave is related to its speed. *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 $u_t - u_{xxt} + u\, u_x = 0$.