

Chapter 8

Linear and Nonlinear Evolution Equations

In this chapter, we analyze some of the most important evolution equations, both linear and nonlinear, in one spatial dimension. Our first stop is to revisit the heat equation. Armed with the Fourier transform, we introduce the notion of a fundamental solution, which plays the role of the Green's function for a dynamical partial differential equation. On an unbounded interval, the Fourier transform enables us to construct an explicit formula in terms of a Gaussian kernel; on finite intervals, we recover the Fourier series solutions constructed in Chapter 4. We next introduce the Maximum Principle for solutions to the heat equation, that rigorously justifies the entropic decay of temperature in a heated body. In the following section, we discuss solution techniques based on symmetry methods, particularly similarity solutions. Finally, we discuss the Black–Scholes equation, which is the paradigmatic model for the behavior of financial portfolios, first proposed in the early 1970's, and now lying at the foundation of much of modern financial modeling and investment. Using the fact that it can be transformed into the linear heat equation, one applies the latter's fundamental solution to establish the celebrated Black–Scholes formula for option pricing.

Our second example is a paradigmatic nonlinear diffusion equation, named after the mid-twentieth century applied mathematician J.M. Burgers, but actually of much old vintage. It can be viewed as a regularization of the nonlinear transport equation studied in Section 2.3 by the inclusion of diffusive (or viscous) effects modeled by the heat equation. Burgers' equation can be regarded as a very simplified model of viscous fluid dynamics, combining both nonlinearity and diffusion. We discover a remarkable nonlinear change of variables that maps Burgers' equation to the linear heat equation, and thereby facilitates its analysis. We investigate particular solutions of Burgers' equation, and discuss how its viscosity-dependent solutions converge to the shock wave solutions of the nonlinear transport equation as the diffusion term becomes vanishingly small.

Next, we turn our attention to the simplest third order linear evolution equation, which arises as a model for wave mechanics. Unlike the wave equation and first order transport equations, its solutions exhibit dispersion, in which waves of different frequencies move at different speeds. As a result, initially localized disturbances will spread out even though they conserve the underlying energy. Dispersion implies that the individual wave velocity is different than the group velocity, 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.

Our final example is the celebrated Korteweg–deVries 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 nonlinear transport and dispersion. Unlike the linearly dispersive model, the Korteweg–deVries equation admits explicit, localized traveling wave solutions known as “solitons”. Even more remarkably, two solitons interact by preserving their overall form, and only experiencing a slight phase shift as the remnant of the inherent nonlinearity. The Korteweg–deVries equation is the prototype of what is known as a “completely integrable” partial differential equation, whose many remarkable properties were first discovered in the mid 1960’s. A surprisingly large number of such completely integrable nonlinear systems arise in a wide range of applications, and their analysis remains an active area of contemporary research, [1, 43].

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 either 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 when solving initial value problems. The fundamental solution(s) measure the effect of a concentrated, instantaneous impulse, either in the initial conditions or as an external force to the system.

We restrict our attention to homogeneous boundary conditions — keeping in mind that these can be included by use of linear superposition. The idea is to analyze the case when 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 y along the bar. The heat will diffuse away from its initial concentration, and the resulting *fundamental solution* is denoted by

$$u(t, x) = F(t, x; \xi), \quad \text{with} \quad F(0, x; \xi) = \delta(x - \xi). \quad (8.1)$$

For each fixed ξ , the fundamental solution $F(t, x; \xi)$, considered as a function of $t > 0$ and x , must satisfy the differential equation, so

$$\frac{\partial F}{\partial t} = \gamma \frac{\partial^2 F}{\partial x^2}, \quad (8.2)$$

as well as 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

$$u(0, x) = f(x) = \int_a^b \delta(x - \xi) f(\xi) d\xi \quad (8.3)$$

as a superposition of delta functions, as in (5.16). Linearity implies that the solution can be expressed as the corresponding superposition of the responses to those concentrated delta profiles:

$$u(t, x) = \int_a^b F(t, x; \xi) f(\xi) d\xi. \quad (8.4)$$

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 y immediately implies that the integral (8.4) 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

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \quad \text{for} \quad -\infty < x < \infty, \quad t > 0, \quad (8.5)$$

on the entire real line. 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

$$u(0, x) = f(x) \quad \text{for} \quad -\infty < x < \infty. \quad (8.6)$$

In order to specify the solution uniquely, we shall require that the temperature be square-integrable (in L^2) at all times, so that

$$\int_{-\infty}^{\infty} |u(t, x)|^2 dx < \infty \quad \text{for all} \quad t \geq 0. \quad (8.7)$$

Roughly speaking, we are requiring that the temperature be vanishingly small at large distances; this requirement assumes the role of boundary conditions in this context.

To solve the initial value problem (8.5–6), we apply the Fourier transform, in the x variable, to both sides of the heat equation. In view of the effect of the Fourier transform on derivatives, cf. (7.40), the result is

$$\frac{\partial \hat{u}}{\partial t} = -k^2 \hat{u}, \quad (8.8)$$

where

$$\hat{u}(t, k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u(t, x) e^{-ikx} dx \quad (8.9)$$

is the Fourier transformed solution. For each fixed k , (8.8) can be viewed as a linear ordinary differential equation for $\hat{u}(t, k)$, with initial conditions

$$\hat{u}(0, k) = \hat{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ikx} dx \quad (8.10)$$

given by Fourier transforming the initial data (8.6). The solution to the initial value problem (8.8, 10) is immediate:

$$\hat{u}(t, k) = e^{-k^2 t} \hat{f}(k). \quad (8.11)$$

We can thus recover the physical solution to the initial value problem (8.5–6) by applying the inverse Fourier transform to (8.11), leading to the explicit integral formula

$$u(t, x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx} \hat{u}(t, k) dk = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx - k^2 t} \hat{f}(k) dk. \quad (8.12)$$

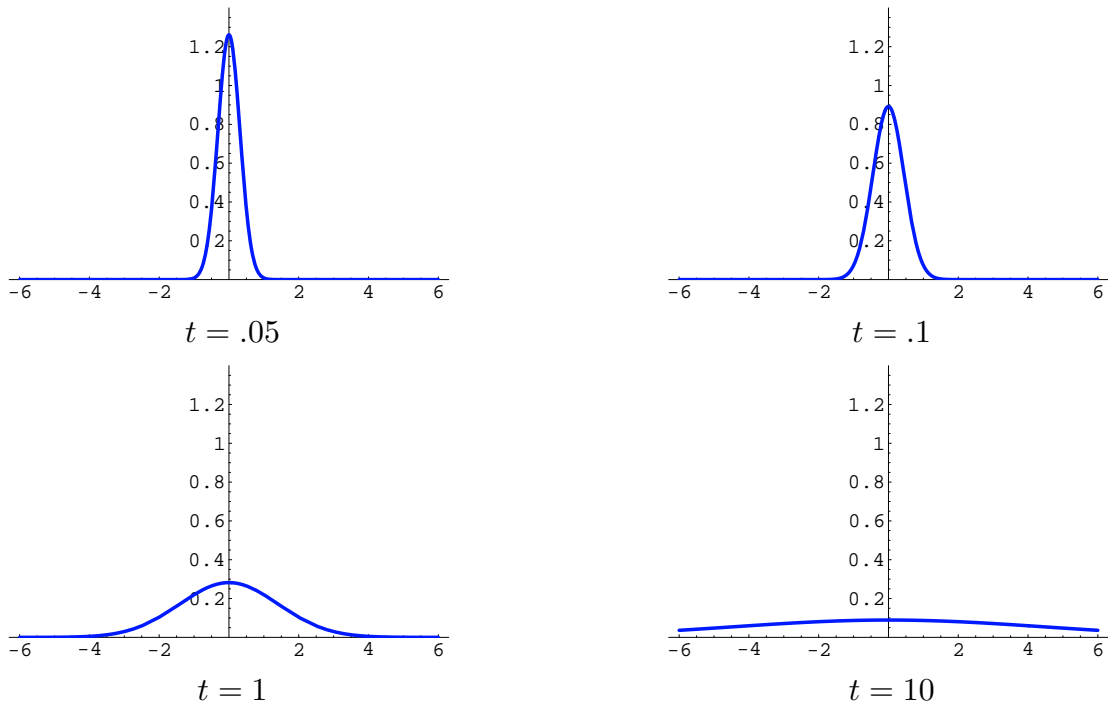


Figure 8.1. The Fundamental Solution to the Heat Equation.

In particular, to recover 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 (7.36), its Fourier transform is

$$\widehat{\delta}_\xi(k) = \frac{e^{-ik\xi}}{\sqrt{2\pi}}.$$

Plugging this into (8.12), and then referring to our table of Fourier transforms, we are led to the following explicit formula for the fundamental solution:

$$F(t, x; \xi) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ik(x-\xi) - k^2 t} dk = \frac{1}{2\sqrt{\pi t}} e^{-(x-\xi)^2/(4t)}. \quad (8.13)$$

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,

$$\lim_{t \rightarrow 0^+} F(t, x; \xi) = \begin{cases} 0, & x \neq \xi, \\ \infty, & x = \xi. \end{cases}$$

Furthermore, its integral

$$\int_{-\infty}^{\infty} F(t, x; \xi) dx = 1, \quad (8.14)$$

is constant — in accordance with the law of conservation of heat energy. Therefore, as $t \rightarrow 0^+$, the fundamental solution satisfies the original limiting definition (5.8–9) 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 smoothes 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 total heat energy (8.14), 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*. The width of the bell curve corresponds to the *standard deviation*. For this reason, the fundamental solution to the heat equation is sometimes referred to as a “Gaussian filter”.

Remark: One of the striking, non-physical artifacts of the heat equation is that the heat 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 energy 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 that postulates that signals cannot propagate faster than the speed of light. Despite this non-physical effect, the heat equation remains an accurate model for heat propagation and similar diffusive phenomena, and so continues to be successfully used in many applications.

With the fundamental solution in hand, we can adapt the linear superposition formula (8.4) to reconstruct the general solution

$$u(t, x) = \frac{1}{2\sqrt{\pi t}} \int_{-\infty}^{\infty} e^{-(x-\xi)^2/(4t)} f(\xi) d\xi \quad (8.15)$$

to our initial value problem (8.5). This solution formula is merely a restatement of (8.12) when combined with the Fourier transform formula (8.10). Comparing with (7.51), we see that the solutions are obtained by convolution,

$$u(t, x) = g(t, x) * f(x), \quad \text{where} \quad g(t, x) = F(t, x; 0) = e^{-x^2/(4t)},$$

of the initial data with a one-parameter family of progressively wider and shorter Gaussian filters. 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 (8.15) 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 (8.15) provides an effective method for denoising signals and images.

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

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

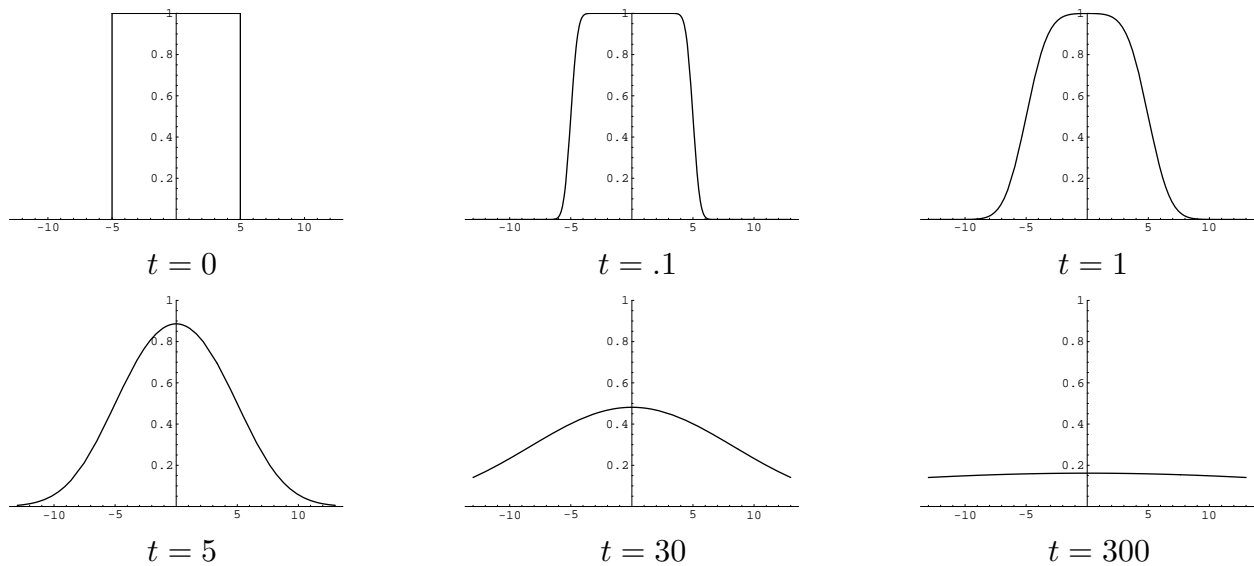


Figure 8.2. Error Function Solution to the Heat Equation.

The subsequent temperature is provided by the solution to the heat equation obtained by the integral formula (8.15):

$$u(t, x) = \frac{1}{2\sqrt{\pi t}} \int_a^b e^{-(x-\xi)^2/(4t)} d\xi = \frac{1}{2} \left[\operatorname{erf} \left(\frac{x-a}{2\sqrt{t}} \right) - \operatorname{erf} \left(\frac{x-b}{2\sqrt{t}} \right) \right], \quad (8.16)$$

where $\operatorname{erf}(z)$ denotes the error function defined in (2.81). Graphs of the solution (8.16) for $a = -5$, $b = 5$ at the indicated times are displayed in Figure 8.2. Observe the instantaneous smoothing of the sharp interface and infinitely fast propagation of the disturbance out to infinity, followed by a gradual decay to thermal equilibrium.

The Forced Heat Equation

The fundamental solution approach can be also applied to solve the inhomogeneous heat equation

$$u_t = u_{xx} + h(t, x), \quad (8.17)$$

modeling a bar under an external heat source $h(t, x)$, that might depend upon both position and time. We begin by solving the particular case

$$u_t = u_{xx} + \delta(t - \tau) \delta(x - \xi), \quad (8.18)$$

whose inhomogeneity represents a heat source of unit magnitude that is concentrated at a position $x = \xi$ and applied at a single time $t = \tau$. 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

$$u(0, x) = 0 \quad (8.19)$$

as well as homogeneous boundary conditions of one of our standard types. The resulting solution

$$u(t, x) = G(t, x; \tau, \xi) \quad (8.20)$$

will be referred to as the *general fundamental solution* to the heat equation. Since a heat source which is applied at time τ will only affect the solution at later times $t \geq \tau$, we expect that

$$G(t, x; \tau, \xi) = 0 \quad \text{for all} \quad t < \tau. \quad (8.21)$$

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 (8.20), we are able to solve the problem for a general external heat source (8.17). We first write the forcing as a superposition

$$h(t, x) = \int_0^\infty \int_a^b h(\tau, \xi) \delta(t - \tau) \delta(x - \xi) d\xi d\tau \quad (8.22)$$

of concentrated instantaneous heat sources. Linearity allows us to conclude that the solution is given by the self-same superposition formula

$$u(t, x) = \int_0^t \int_a^b h(\tau, \xi) G(t, x; \tau, \xi) d\xi d\tau. \quad (8.23)$$

The fact that we only need to integrate over times $0 \leq \tau \leq t$ is a consequence of (8.21).

Remark: If we have a nonzero initial condition, $u(0, x) = f(x)$, then, by linear superposition, the solution

$$u(t, x) = \int_a^b F(t, x; \xi) f(\xi) d\xi + \int_0^t \int_a^b h(\tau, \xi) G(t, x; \tau, \xi) d\xi d\tau \quad (8.24)$$

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 with respect to x . In view of (7.36, 40), we find

$$\frac{\partial \hat{u}}{\partial t} + k^2 \hat{u} = \frac{1}{\sqrt{2\pi}} e^{-ik\xi} \delta(t - \tau). \quad (8.25)$$

which is an inhomogeneous first order ordinary differential equation for the Fourier transform $\hat{u}(t, k)$ of $u(t, x)$. Assuming $\tau > 0$, by (8.21), the initial condition is

$$\hat{u}(0, k) = 0. \quad (8.26)$$

We solve the initial value problem (8.25–26) by the usual method, [23]: Multiplying the differential equation by the integrating factor $e^{k^2 t}$ yields

$$\frac{\partial}{\partial t} \left(e^{k^2 t} \hat{u} \right) = \frac{1}{\sqrt{2\pi}} e^{k^2 t - ik\xi} \delta(t - \tau).$$

Integrating both sides from 0 to t and using the initial condition, we find

$$\widehat{u}(t, k) = \frac{1}{\sqrt{2\pi}} e^{-k^2(t-\tau) - ik\xi} \sigma(t - \tau),$$

where $\sigma(s)$ is the usual step function (5.23). Finally, we apply the inverse Fourier transform formula (7.9), and then (8.13), to deduce that

$$\begin{aligned} u(t, x) = G(t, x; \tau, \xi) &= \frac{\sigma(t - \tau)}{2\pi} \int_{-\infty}^{\infty} e^{-k^2(t-\tau) + ik(x-\xi)} dk \\ &= \frac{\sigma(t - \tau)}{2\sqrt{\pi(t - \tau)}} \exp\left[-\frac{(x - \xi)^2}{4(t - \tau)}\right] = \sigma(t - \tau)F(t - \tau, x; \xi). \end{aligned}$$

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 $t = 0$. In other words, an initial condition has the same after-effect on the temperature as an instantaneous applied heat source of the same magnitude. Finally, the superposition principle (8.23) produces the solution,

$$u(t, x) = \int_0^t \int_{-\infty}^{\infty} \frac{h(\tau, \xi)}{2\sqrt{\pi(t - \tau)}} \exp\left[-\frac{(x - \xi)^2}{4(t - \tau)}\right] d\xi d\tau, \quad (8.27)$$

to the heat equation with source term and zero initial condition on an infinite bar. A non-zero initial condition $u(0, x) = f(x)$ leads, as in the superposition formula (8.24), to an additional term of the form (8.15) in the solution formula.

The Black–Scholes Equation and Mathematical Finance

The most important partial differential equation in financial modeling and investment is the celebrated *Black–Scholes equation*

$$\frac{\partial u}{\partial t} + \frac{\sigma^2}{2} x^2 \frac{\partial^2 u}{\partial x^2} + r x \frac{\partial u}{\partial x} - r u = 0, \quad (8.28)$$

first proposed in 1973 by the American economists Fischer Black and Myron Scholes, [16], and Robert C. Merton, [88]. The dependent variable $u(t, x)$ represents the monetary value of a portfolio consisting of a single financial *option*, that is, a contract to either buy or sell an asset at a specified *exercise price* p at a certain future time t_* . The value $u(t, x)$ of the option will depend on the current time $t \leq t_*$ and the current price $x \geq 0$ of the underlying asset. The constant $\sigma > 0$ represents the asset's *volatility*, while r denotes the (assumed fixed) *interest rate* for bank deposits, where the investor 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, and so would take us too far afield to explain here; instead, we refer the interested reader to [135] for details. The Black–Scholes equation and its generalizations form the basis of much of the modern financial world, particularly hedge funds.

Observe first that the Black–Scholes equation is a *backwards* diffusion process, since, upon solving for

$$\frac{\partial u}{\partial t} = -\frac{\sigma^2}{2} x^2 \frac{\partial^2 u}{\partial x^2} - r x \frac{\partial u}{\partial x} + r u, \quad (8.29)$$

the coefficient of the diffusion term u_{xx} is *negative*. This implies that the initial value problem is well-posed only in *backwards* time. In other words, given a prescribed value of the option at some time in the future, we can use the Black–Scholes equation to determine its current value. However, we cannot predict the future value based on the current worth of the portfolio. The “final value problem” for the Black–Scholes equation is, given the exercise time of the option t_* , to determine its value $u(t, x)$ for times $t < t_*$ and asset values $x \geq 0$, subject to the *final condition*

$$u(t_*, x) = f(x). \quad (8.30)$$

For a so-called *European call option*, where the asset is to be bought at the exercise price $p > 0$ at the specified time, the final condition is

$$u(t_*, x) = \max\{x - p, 0\}. \quad (8.31)$$

For a *put option*, where the asset is to be sold, the final condition is

$$u(t_*, x) = \max\{p - x, 0\}. \quad (8.32)$$

The solution $u(t, x)$ will be defined for all $t < t_*$ and all $x > 0$, subject to the boundary conditions[†]

$$u(t, 0) = 0, \quad u(t, x) \sim x \quad \text{as} \quad x \rightarrow \infty.$$

It turns out that the Black–Scholes equation can be solved explicitly by transforming it into the heat equation. The first step is to convert it to a forwards diffusion process, by setting

$$\tau = \frac{1}{2} \sigma^2 (t_* - t), \quad v(\tau, x) = u(t_* - 2\tau/\sigma^2, x).$$

A simple chain rule computation shows that v satisfies

$$\frac{\partial v}{\partial \tau} = x^2 \frac{\partial^2 v}{\partial x^2} + \kappa x \frac{\partial v}{\partial x} - \kappa v, \quad \text{where} \quad \kappa = \frac{2r}{\sigma^2}.$$

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 as an Euler ordinary differential equation (A.11). According to Exercise ■, these terms can be placed into constant coefficient form by the change of independent variables $x = e^y$. Indeed, writing

$$w(\tau, y) = v(\tau, e^y),$$

[†] The asymptotic boundary condition means that the ratio $u(t, x)/x$ tends to a constant as $x \rightarrow \infty$.

we apply the chain rule to compute the derivatives

$$\frac{\partial w}{\partial \tau} = \frac{\partial v}{\partial \tau}, \quad \frac{\partial w}{\partial y} = e^y \frac{\partial v}{\partial x} = x \frac{\partial v}{\partial x}, \quad \frac{\partial^2 w}{\partial y^2} = e^{2y} \frac{\partial^2 v}{\partial x^2} + e^y \frac{\partial v}{\partial x} = x^2 \frac{\partial^2 v}{\partial x^2} + x \frac{\partial v}{\partial x}.$$

As a result, we find that w solves the partial differential equation

$$\frac{\partial w}{\partial \tau} = \frac{\partial^2 w}{\partial y^2} + (\kappa - 1) \frac{\partial w}{\partial y} - \kappa w. \quad (8.33)$$

This is getting closer to the heat equation, and, in fact, can be changed into it by setting

$$w(\tau, y) = e^{\alpha\tau + \beta x} z(\tau, y)$$

for suitable constants α, β . Indeed, differentiating and substituting into (8.33) yields

$$\frac{\partial z}{\partial \tau} + \alpha z = \frac{\partial^2 z}{\partial y^2} + 2\beta \frac{\partial z}{\partial y} + \beta^2 z + (\kappa - 1) \left(\frac{\partial z}{\partial y} + \beta z \right) + \kappa z.$$

We can eliminate the terms involving $\partial z/\partial y$ and z by setting

$$\alpha = -\frac{1}{4}(\kappa + 1)^2, \quad \beta = -\frac{1}{2}(\kappa - 1). \quad (8.34)$$

Thus, if

$$w(\tau, y) = e^{-\frac{1}{4}(\kappa+1)^2\tau - \frac{1}{2}(\kappa-1)y} z(\tau, y), \quad (8.35)$$

then $z(\tau, y)$ satisfies the heat equation

$$\frac{\partial z}{\partial \tau} = \frac{\partial^2 z}{\partial y^2}. \quad (8.36)$$

Unwinding the preceding argument, we have managed to prove the following:

Proposition 8.2. *If $z(\tau, y)$ is the solution to the initial value problem*

$$\frac{\partial z}{\partial \tau} = \frac{\partial^2 z}{\partial y^2}, \quad z(0, y) = h(y) = e^{(\kappa-1)y/2} f(e^y), \quad (8.37)$$

for $\tau > 0$, $-\infty < y < \infty$, then

$$u(t, x) = x^{-(\kappa-1)/2} e^{-(\kappa+1)^2\sigma^2(t_\star-t)/8} z\left(\frac{1}{2}\sigma^2(t_\star-t), \log x\right) \quad (8.38)$$

solves the final value problem (8.28, 30) for the Black–Scholes equation for $t < t_\star$ and $0 < x < \infty$.

Now, according to (8.15), the solution to the initial value problem (8.37) can be written as an convolution integral of the initial data with the heat equation's fundamental solution:

$$z(\tau, y) = \frac{1}{2\sqrt{\pi\tau}} \int_{-\infty}^{\infty} e^{-(y-\eta)^2/(4\tau)} h(\eta) d\eta = \frac{1}{2\sqrt{\pi\tau}} \int_{-\infty}^{\infty} e^{-(y-\eta)^2/(4\tau) + (\kappa-1)\eta/2} f(e^\eta) d\eta. \quad (8.39)$$

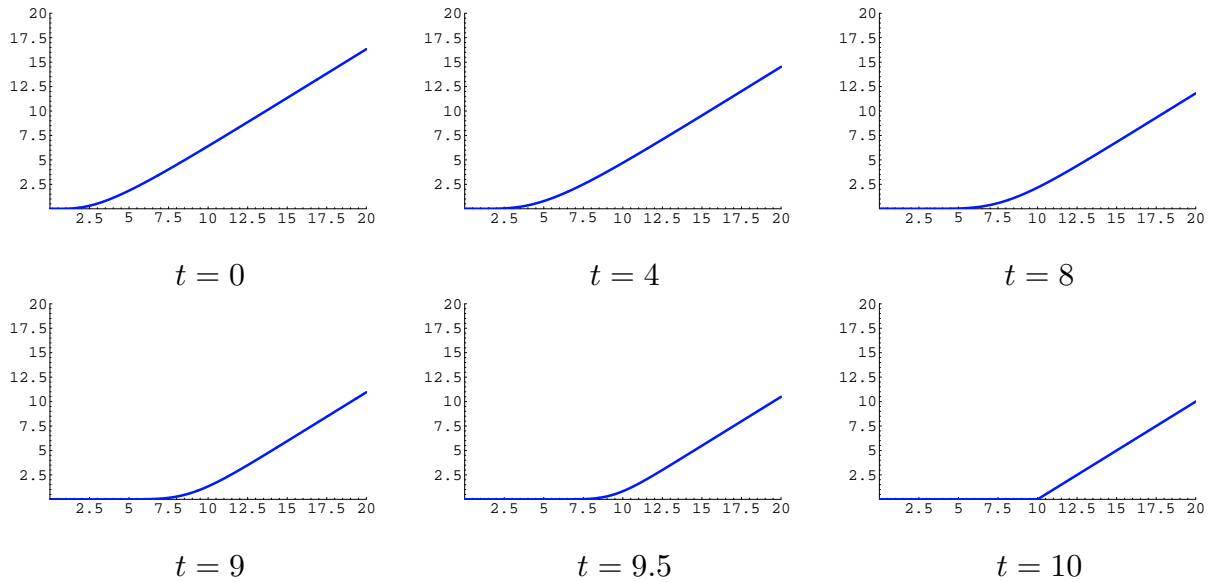


Figure 8.3. Solution to the Black–Scholes Equation.

Combining this formula with (8.38) produces an explicit solution formula to the general final value problem for the Black–Scholes equation. In particular, for the European call option (8.31), the initial condition is

$$z(0, y) = h(y) = e^{(\kappa-1)y/2} \max\{e^y - p, 0\},$$

and so

$$\begin{aligned} z(\tau, y) &= \frac{1}{2\sqrt{\pi\tau}} \int_{\log p}^{\infty} e^{-(y-\eta)^2/(4\tau) + (\kappa-1)\eta/2} (e^\eta - p) d\eta \\ &= \frac{1}{2} \left[e^{(\kappa+1)^2\tau/4 + (\kappa+1)y/2} \operatorname{erfc} \left(\frac{\log p - (\kappa+1)\tau - y}{2\sqrt{\tau}} \right) \right. \\ &\quad \left. - p e^{(\kappa-1)^2\tau/4 + (\kappa-1)y/2} \operatorname{erfc} \left(\frac{\log p - (\kappa-1)\tau - y}{2\sqrt{\tau}} \right) \right], \end{aligned} \quad (8.40)$$

where

$$\operatorname{erfc} x = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-z^2} dz = 1 - \operatorname{erf} x \quad (8.41)$$

is the *complementary error function*, cf. (2.81). (The integrals in (8.40) are evaluated by completing the square inside the exponential; see Exercise ■.) Substituting (8.40) into (8.38) results in the *Black–Scholes formula* for a European call option:

$$\begin{aligned} u(t, x) &= \frac{1}{2} \left[x \operatorname{erfc} \left(- \frac{(r + \frac{1}{2}\sigma^2)(t_* - t) + \log(x/p)}{\sqrt{2\sigma^2(t_* - t)}} \right) \right. \\ &\quad \left. - p e^{-r(t_* - t)} \operatorname{erfc} \left(- \frac{(r - \frac{1}{2}\sigma^2)(t_* - t) + \log(x/p)}{\sqrt{2\sigma^2(t_* - t)}} \right) \right]. \end{aligned} \quad (8.42)$$

A graph of the solution for the specific values $t_* = 10$, $r = .1$, $\sigma = .2$, $p = 10$, appears in Figure 8.3. Observe that the option's value slowly decreases as the time gets closer and closer to the exercise time t_* .

8.2. Symmetry and Similarity.

One of the key insights of the geometric approach to partial differential equations is to exploit the inherent symmetry properties to construct solutions. Unlike separation of variables which are restricted to special types of linear partial differential equations[†], symmetry methods can also be successfully applied to produce solutions to a broad range of nonlinear partial differential equations. While we do not have the space or the mathematical tools to develop the full apparatus of symmetry techniques, we can exploit some of the basic symmetries such as 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, if we have a symmetry, and know 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 and 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 provides an algorithm for completely determining all the (continuous) symmetries of a given differential equation, but this is beyond the scope of this introductory text. However, direct inspection and/or physical intuition will often detect the most important symmetries without appealing to such a sophisticated theory. Modern applications of Lie’s symmetry methods to partial differential equations arising in physics and engineering can be traced back to the influential book on hydrodynamics by the author’s thesis advisor, Garrett Birkhoff, [14]. A complete and comprehensive treatment of symmetry methods can be found in the first author’s book [103], and, at a more introductory level, in the recent books by Hydon, [69], and Cantwell, [30], the latter having particular emphasis on applications to fluid mechanics.

The heat equation serves as an excellent testing ground for the general symmetry methodology, as 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,

$$t \longmapsto t + a, \quad x \longmapsto x + b, \quad (8.43)$$

where a, b are constants, changes the function $u(t, x)$ into the translated function[§]

$$U(t, x) = u(t - a, x - b). \quad (8.44)$$

[†] This is not quite fair: separation of variables can be applied to some special nonlinear partial differential equations such as Hamilton–Jacobi equations, [90].

[§] The minus signs arise because when we set $\hat{t} = t + a$, $\hat{x} = x + b$, then the translated function is $U(\hat{t}, \hat{x}) = u(t, x) = u(\hat{t} - a, \hat{x} - b)$. Dropping the hats produces the stated formula.

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

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

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, the translation symmetries formalize 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 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 $U(t, x)$ 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 are the scaling invariances. We already know that if $u(t, x)$ is a solution, so is any scalar multiple $cu(t, x)$; this is a simple consequence of linearity of the heat equation. We can also add an arbitrary constant to the temperature, noting that

$$U(t, x) = cu(t, x) + k \tag{8.45}$$

is a solution for any choice of constants c, k . Physically, the transformation (8.45) amounts to a change in the scale used to measure temperature. For instance, if u is measured 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 upon our choice of thermometer.

More interestingly, suppose we rescale the space and time variables:

$$t \mapsto \alpha t, \quad x \mapsto \beta x, \tag{8.46}$$

where $\alpha, \beta > 0$ are positive constants. The effect of such a scaling transformation is to convert $u(t, x)$ into a rescaled function[†]

$$U(t, x) = u(\alpha^{-1}t, \beta^{-1}x). \tag{8.47}$$

[†] As before, setting $\hat{t} = \alpha t$, $\hat{x} = \beta x$, produces the rescaled function $U(\hat{t}, \hat{x}) = u(t, x) = u(\alpha^{-1}\hat{t}, \beta^{-1}\hat{x})$, and we then drop the hats.

The derivatives of U are related to those of u according to the formulae

$$\frac{\partial U}{\partial t} = \frac{1}{\alpha} \frac{\partial u}{\partial t}, \quad \frac{\partial U}{\partial x} = \frac{1}{\beta} \frac{\partial u}{\partial x}, \quad \frac{\partial^2 U}{\partial x^2} = \frac{1}{\beta^2} \frac{\partial^2 u}{\partial x^2},$$

which are direct consequences of the multi-variable chain rule. Therefore, if u satisfies the heat equation $u_t = \gamma u_{xx}$, then U satisfies the rescaled heat equation

$$U_t = \frac{1}{\alpha} u_t = \frac{\gamma}{\alpha} u_{xx} = \frac{\beta^2 \gamma}{\alpha} U_{xx},$$

which we rewrite as

$$U_t = \Gamma U_{xx}, \quad \text{where} \quad \Gamma = \frac{\beta^2 \gamma}{\alpha}. \quad (8.48)$$

Thus, the net effect of scaling space and time is merely to rescale the diffusion coefficient in the heat equation. Physically, the scaling symmetry (8.46) corresponds to a change in the physical units used to measure time and distance. For instance, to change from seconds to minutes, set $\alpha = 60$, and from meters to yards, set $\beta = 1.0936$. The net effect (8.48) on the diffusion coefficient γ is a reflection of its physical units, namely distance²/time.

In particular, if we choose

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

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

$$u(t, x) = U(\gamma t, x) \quad (8.49)$$

solves the heat equation for the diffusivity γ . Thus, the only effect of the diffusion coefficient γ 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 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

$$t \mapsto \beta^2 t, \quad x \mapsto \beta x, \quad (8.50)$$

does not alter the equation, and hence defines a *scaling symmetry*, also known as a *similarity transformation*, for the heat equation. Combining (8.50) with the linear rescaling $u \mapsto cu$, we make the elementary, but important observation that if $u(t, x)$ is any solution to the heat equation, then so is the function

$$U(t, x) = c u(\beta^{-2} t, \beta^{-1} x), \quad (8.51)$$

for the *same* diffusion coefficient γ . For example, rescaling the solution $u(t, x) = e^{-\gamma t^2} \cos x$ leads to the solution $U(t, x) = c e^{-\gamma t^2 / \beta^2} \cos x / \beta$.

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 $0 \leq x \leq \ell$, then $U(t, x)$, as given in (8.51), is defined for $0 \leq x \leq \beta \ell$.

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 $U(t, x) = u(\frac{1}{4}t, \frac{1}{2}x)$ 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. Thus, 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 group of scaling symmetries. For a partial differential equation in two variables — say x and t — the similarity solutions can be found by solving an *ordinary differential equation*.

Suppose our partial differential equation admits a one-parameter family of scaling symmetries

$$t \mapsto \beta^a t, \quad x \mapsto \beta^b x, \quad u \mapsto \beta^c u, \quad (8.52)$$

for constants a, b, c , 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

$$U(t, x) = \beta^c u(\beta^{-a} t, \beta^{-b} x) \quad (8.53)$$

for all values of $\beta > 0$. Checking that this indeed defines a symmetry is a simple matter of applying the chain rule, which implies that the derivatives scale according to

$$u_t \mapsto \beta^{c-a} u_t, \quad u_x \mapsto \beta^{c-b} u_x, \quad u_{tt} \mapsto \beta^{c-2a} u_{tt}, \quad u_{xt} \mapsto \beta^{c-a-b} u_{xt}, \quad (8.54)$$

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 symmetry group (8.52), so that

$$u(t, x) = \beta^c u(\beta^{-a} t, \beta^{-b} x) \quad (8.55)$$

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. Since the left hand side of (8.55) 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

$$u(t, x) = t^{c/a} v(\xi), \quad \text{where} \quad \xi = x t^{-b/a} \quad \text{and} \quad v(\xi) = u(\xi, 1), \quad (8.56)$$

[†] This assumes $t > 0$; for $t < 0$, just replace t by $-t$.

are referred to as the *similarity variables*. 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 will provide a similarity solution to the original partial differential equation by use of the similarity ansatz (8.56).

Example 8.3. As a first example, let's return to the nonlinear transport equation

$$u_t + uu_x = 0. \quad (8.57)$$

Under (8.52, 54), the equation rescales to

$$\beta^{c-a}u_t + \beta^{2c-b}uu_x = 0,$$

which is unchanged provided

$$c - a = 2c - b, \quad \text{and hence} \quad c = b - a.$$

Setting $a = 1$, we conclude that if $u(t, x)$ is any solution, so is the rescaled function

$$U(t, x) = \beta^{b-1} u(\beta^{-1} t, \beta^{-b} x)$$

for any b and any $\beta > 0$.

To find the associated similarity solutions, we use (8.56) to introduce the similarity ansatz

$$u(t, x) = t^{b-1} v(\xi), \quad \text{where} \quad \xi = x t^{-b}. \quad (8.58)$$

Differentiating, we find

$$\begin{aligned} u_t &= -b x t^{-b-1} v'(\xi) + (b-1) t^{b-2} v(\xi) = t^{b-2} [-b \xi v'(\xi) + (b-1) v(\xi)], \\ u_x &= t^{-1} v'(\xi). \end{aligned}$$

Substituting into the transport equation (8.57), we find

$$0 = u_t + uu_x = t^{b-2} [(v - b\xi) v' + (b-1) v],$$

and so

$$(v - b\xi) \frac{dv}{d\xi} + (b-1) v = 0. \quad (8.59)$$

Any solution to this ordinary differential equation will, when substituted into (8.58), produce a similarity solution to the nonlinear transport equation. If $b = 1$, then either $v = b\xi$, and we are led to our fan wave solution $u(t, x) = x/t$, or v is constant, and so is u . Otherwise, we can, in fact, linearize (8.59) by treating ξ as a function of v , whence

$$(b-1) v \frac{d\xi}{dv} - b\xi = -v.$$

The general solution to a linear, first order ordinary differential equation is found by the standard method, [15, 23], resulting in

$$\xi = v + k v^{b/(b-1)},$$

where k is the constant of integration. Recalling (8.58), 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 rotated parabola:

$$x = k u^2 + t u, \quad \text{so that} \quad u = \frac{-t \pm \sqrt{t^2 + 4kx}}{2k}.$$

Example 8.4. Let us return to the linear heat equation

$$u_t = u_{xx}. \tag{8.60}$$

Under rescaling, the equation becomes $\beta^{c-a} u_t = \beta^{c-2b} u_{xx}$, and thus (8.52) 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 is a reflection of the linearity of the equation.

To find the similarity solutions, we use (8.56) to introduce the similarity ansatz

$$u(t, x) = t^{c/2} v(\xi), \quad \text{where} \quad \xi = x/\sqrt{t}.$$

Differentiating,

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

Substituting into the heat equation and canceling a common power of t , we find v must satisfy the linear ordinary differential equation

$$v'' + \frac{1}{2} \xi v' - \frac{1}{2} c v = 0. \tag{8.61}$$

If $c = 0$, then the solution is

$$v(\xi) = c_1 + c_2 \operatorname{erf}\left(\frac{1}{2} \xi\right),$$

where c_1, c_2 are arbitrary constants and erf is the error function (2.81). 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 solution we already encountered in (8.16) 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 (8.61) 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.

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 temperature of an isolated body can only decrease. The *Maximum Principle* is the mathematical formulation of this thermodynamical law. It states that the temperature of a body cannot, in the absence of external sources, ever be larger than its initial or boundary values, and hence the maximal temperature must decrease. It 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 can only be achieved on its boundary.

The proof of the Maximum Principle will be facilitated if we generalize it to when the heat equation is subject to an external heat source that extracts heat energy from the body.

Theorem 8.5. *Let $\gamma > 0$. Suppose $u(t, x)$ is a solution to the forced heat equation*

$$\frac{\partial u}{\partial t} = \gamma \frac{\partial^2 u}{\partial x^2} + F(t, x) \quad (8.62)$$

on the rectangular domain

$$R = \{a \leq x \leq b, 0 \leq t \leq c\}.$$

Assume that the external heat source is nowhere positive: $F(t, x) \leq 0$ for all $(t, x) \in R$. Then the global maximum of $u(t, x)$ on the domain R occurs either at $t = 0$ or $x = a$ or $x = b$.

In other words, if we are everywhere either taking heat out of the bar or leaving it insulated, then the maximum overall temperature occurs either at the initial time, or at one of its endpoints. In particular, in the fully insulated case $F(t, x) \equiv 0$, (8.62) reduces to the heat equation, and Theorem 8.5 applies as stated.

Proof: First let us first prove the result under the stronger assumption $F(t, x) < 0$, which implies that

$$\frac{\partial u}{\partial t} < \gamma \frac{\partial^2 u}{\partial x^2} \quad (8.63)$$

[†] According to [103; Example 3.3], the general solution can be written in terms of parabolic cylinder functions, [3].

everywhere in the rectangle R . Suppose first that $u(t, x)$ has a (local) maximum at a point (t^*, x^*) in the interior of R . Then, by multivariable calculus, [8], its gradient must vanish there, $\nabla u(t^*, x^*) = \mathbf{0}$, and hence

$$u_t(t^*, x^*) = u_x(t^*, x^*) = 0. \quad (8.64)$$

Moreover, the scalar function $h(x) = u(t^*, x)$ has a maximum at $x = x^*$, and hence, by the second derivative test,

$$h''(x^*) = u_{xx}(t^*, x^*) \leq 0. \quad (8.65)$$

But the requirements (8.64–65) are incompatible with the initial inequality (8.63). We conclude that, under our assumption, 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 at a point $(t^*, x^*) = (c, x^*)$ on the edge of the rectangle where $t = c$ and $a < x^* < b$. If such were to occur, then $g(t) = u(t, x^*)$ would necessarily be non-decreasing at $t = c$, and hence $g'(t) = u_t(c, x^*) \geq 0$ there. The preceding argument also implies that $u_{xx}(c, x^*) \leq 0$, and again these two requirements are incompatible with (8.63). 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 when $F(t, x) \leq 0$ — which includes the heat equation — requires a little trick. We set

$$v(t, x) = u(t, x) + \varepsilon x^2, \quad \text{where} \quad \varepsilon > 0.$$

Then,

$$\frac{\partial v}{\partial t} = \frac{\partial u}{\partial t} = \gamma \frac{\partial^2 u}{\partial x^2} + F(t, x) = \gamma \frac{\partial^2 v}{\partial x^2} - 2\gamma\varepsilon + F(t, x) = \gamma \frac{\partial^2 v}{\partial x^2} + \tilde{F}(t, x),$$

where

$$\tilde{F}(t, x) = F(t, x) - 2\gamma\varepsilon < 0$$

everywhere in R . Thus, by the previous argument, any local maximum of v can only occur when $t = 0$ or $x = a$ or $x = b$. Now we let $\varepsilon \rightarrow 0$ and conclude the same for u . More precisely, suppose that $u(t, x) \leq M$ when either $t = 0$ or $x = a$ or $x = b$. Then

$$v(t, x) \leq M + \varepsilon \max\{a^2, b^2\}$$

on the three sides of the rectangle, and hence, by the previous arguments,

$$u(t, x) \leq v(t, x) \leq M + \varepsilon \max\{a^2, b^2\} \quad \text{for all} \quad (t, x) \in R.$$

Now, letting $\varepsilon \rightarrow 0$ proves that $u(t, x) \leq M$ everywhere in R , which completes the proof. *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 8.6. Suppose $u(t, x)$ solves the heat equation $u_t = \gamma u_{xx}$, with $\gamma > 0$, for $a \leq x \leq b$, $0 \leq t \leq c$. Let m and M be, respectively, the minimum and maximum values for the initial and boundary temperatures, meaning that $m \leq u(t, x) \leq M$ for $t = 0$, $a \leq x \leq b$, and also $x = a$ or b , $0 \leq t \leq c$. 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 8.5. To establish the lower bound, we note that $\tilde{u}(t, x) = -u(t, x)$ also satisfies the heat equation, and satisfies $\tilde{u}(t, x) \leq -m$ on the three edges of the rectangle, and hence, by the Maximum Principle, everywhere. But this implies $u(t, x) = -\tilde{u}(t, x) \geq m$. *Q.E.D.*

One immediate application of the Maximum Principle is to prove uniqueness of solutions to the heat equation.

Theorem 8.7. There is at most one solution to the Dirichlet initial-boundary value problem for the forced heat equation.

Proof: Suppose u and \tilde{u} are any two solutions. Then their difference $v = u - \tilde{u}$ solves the homogeneous initial-boundary value problem for the unforced heat equation, with minimum and maximum boundary values $m = 0 \leq v(t, x) \leq 0 = M$ for $t = 0$, $a \leq x \leq b$, and also $x = a$ or b , $0 \leq t \leq c$. But then Corollary 8.6 implies that $0 \leq v(t, x) \leq 0$ everywhere, which implies that $u \equiv \tilde{u}$, thereby establishing uniqueness. *Q.E.D.*

Remark: Existence of the solution follows from the convergence of our Fourier series — assuming the initial and boundary data and the forcing function are sufficiently nice.

8.4. Nonlinear Diffusion.

First order partial differential equations, beginning with elementary scalar transport equations that we studied in Chapter 2, and progressing on to the equations of gas dynamics, the full-blown Euler equations of fluid mechanics, and yet more complicated systems for plasmas, magneto-hydrodynamics and other physical processes, serve to model conservative wave motion. Such systems fail to account for frictional or 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 or viscous 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 non-uniformly to a discontinuous shock wave, leading to an alternative mechanism — called the method of viscosity solutions — for analyzing conservative dynamical processes.

Burgers' Equation

The simplest nonlinear diffusion equation is known as[†] *Burgers' equation*

$$u_t + uu_x = \gamma u_{xx}, \quad (8.66)$$

and is obtained by combining a linear diffusion term coming from the heat equation with the nonlinear transport equation (2.27). As with the heat equation, the diffusion coefficient $\gamma \geq 0$ must be non-negative in order that the initial value problem be well-posed in forwards time. In fluid and gas dynamics, one can interpret 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 Navier–Stokes equations (1.4), [132]. When the viscosity coefficient vanishes, $\gamma = 0$, Burgers' equation reduces to the nonlinear transport equation (2.27), which, as a consequence, is often referred to as the *inviscid Burgers' equation*.

Since Burgers' equation is first order in t , we expect that its solutions will be uniquely prescribed by their initial values, say,

$$u(0, x) = f(x), \quad -\infty < x < \infty. \quad (8.67)$$

(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 the nonlinear waves that we analyzed in Section 2.3, perhaps steepening into some sort of shock. But, as we will see, 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 (8.66–67) remains smooth and well-defined for all subsequent times, [132].

The simplest explicit solutions are the *traveling waves*, for which

$$u(t, x) = v(\xi) = v(x - ct), \quad \text{where} \quad \xi = x - ct, \quad (8.68)$$

indicates a fixed profile, moving to the right with constant speed c . By the chain rule,

$$\frac{\partial u}{\partial t} = -cv'(\xi), \quad \frac{\partial u}{\partial x} = v'(\xi), \quad \frac{\partial^2 u}{\partial x^2} = v''(\xi).$$

Substituting these expressions into Burgers' equation (8.66), we conclude that $v(\xi)$ must satisfy the nonlinear second order ordinary differential equation

$$-cv' + vv' = \gamma v''.$$

[†] The equation is named after the applied mathematician J.M. Burgers, [29], and so the apostrophe goes after the “s”. Burgers' equation was apparently first studied as a physical model by Bateman, [12], in the early twentieth century.

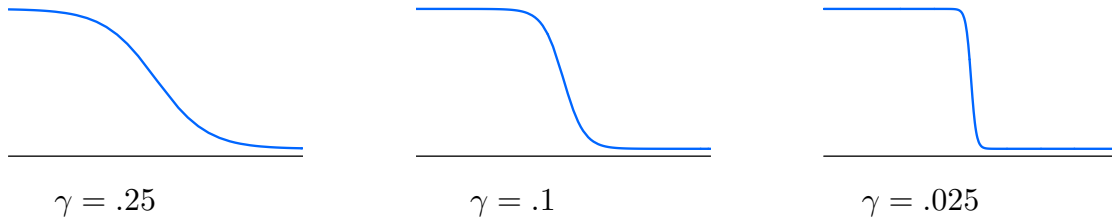


Figure 8.4. Traveling Wave Solutions to Burgers' Equation.

This equation can be solved by first integrating both sides with respect to ξ , and so

$$\gamma v' = k - cv + \frac{1}{2}v^2,$$

where k is a constant of integration. As in the analysis following Proposition 2.3, as $\xi \rightarrow \pm\infty$, the bounded solutions to such an autonomous first order ordinary differential equation tend to one of the equilibrium points or roots of the right hand side. Therefore, to obtain a *bounded* traveling wave solution $v(\xi)$, the quadratic polynomial on the right hand side must have two real roots, which requires $k < \frac{1}{2}c^2$. Assuming this holds, we rewrite the equation in the form

$$2\gamma \frac{dv}{d\xi} = (v-a)(v-b), \quad \text{where} \quad c = \frac{1}{2}(a+b), \quad k = \frac{1}{2}ab. \quad (8.69)$$

For bounded solutions, we restrict to the case when $a < v < b$. Integrating (8.69) 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

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

and recalling (8.69), we conclude that the bounded traveling wave solutions to Burgers' equation all have the explicit form

$$u(t, x) = \frac{ae^{(b-a)(x-ct-\delta)/(2\gamma)} + b}{e^{(b-a)(x-ct-\delta)/(2\gamma)} + 1},$$

where $a < b$ and δ are arbitrary constants. Observe that

$$\lim_{x \rightarrow -\infty} u(t, x) = b, \quad \lim_{x \rightarrow \infty} u(t, x) = a,$$

and hence our solution is a monotonically decreasing function of x , going from b to a . 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 8.4 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 $\gamma \rightarrow 0$, the solutions converge to the step shock wave solution (2.45) to the nonlinear transport equation.

Indeed, a profound fact is that, in the *inviscid limit* as the diffusion becomes vanishingly small, $\gamma \rightarrow 0$, the solutions to Burgers' equation (8.66) converge to the shock wave solution to (2.27) constructed by the Equal Area Rule. More generally, this method allows one to monitor solutions as they evolve into regimes where multiple shocks interact and merge. This is in accordance with our physical intuition, that most physical systems retain a very small dissipative component, which serves to slightly smooth out discontinuities that might appear in a theoretical model that fails to take dissipation/viscosity/etc. 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 [80, 122, 132] 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, [50; vol. 6, p. 102]. Its rediscovery by the applied mathematicians Eberhard Hopf, [67], and Julian Cole, [33], 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 changes of variables will do the trick! However, the resulting nonlinear equation, while evidently linearizable by inverting the change of variables, is rarely of any 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 regimes.

In the present context, our starting point is the linear heat equation

$$v_t = \gamma v_{xx}. \tag{8.70}$$

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

$$v(t, x) = e^{\alpha \varphi(t, x)}, \quad \text{so} \quad \varphi(t, x) = \frac{1}{\alpha} \log v(t, x), \tag{8.71}$$

where α is a nonzero constant. The function $\varphi(t, x)$ is real provided $v(t, x) > 0$ 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 the resulting solution $v(t, x)$ is positive for all $t > 0$. This follows from the Maximum Principle for the heat equation, cf. Corollary 8.6.

To determine the differential equation satisfied by the function φ , we invoke the chain and product rules to differentiate (8.71):

$$v_t = \alpha \varphi_t e^{\alpha \varphi}, \quad v_x = \alpha \varphi_x e^{\alpha \varphi}, \quad v_{xx} = (\alpha \varphi_{xx} + \alpha^2 \varphi_x^2) e^{\alpha \varphi}.$$

Substituting the first and last formulae into the heat equation (8.70) and canceling a common exponential factor, we conclude that $\varphi(t, x)$ satisfies the nonlinear partial differential equation

$$\varphi_t = \gamma \varphi_{xx} + \gamma \alpha \varphi_x^2, \quad (8.72)$$

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

$$\varphi_{tx} = \gamma \varphi_{xxx} + 2\gamma \alpha \varphi_x \varphi_{xx}. \quad (8.73)$$

If we now set

$$\frac{\partial \varphi}{\partial x} = u, \quad (8.74)$$

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 (8.66) when $\alpha = -1/(2\gamma)$. In this manner, we have arrived at the famous *Hopf–Cole transformation*.

Theorem 8.8. *If $v(t, x) > 0$ is any positive solution to the linear heat equation $v_t = \gamma v_{xx}$, then*

$$u(t, x) = \frac{\partial}{\partial x} \left(-2\gamma \log v(t, x) \right) = -2\gamma \frac{v_x}{v} \quad (8.75)$$

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 $\tilde{\varphi}(t, x)$ that satisfies (8.74); for example

$$\tilde{\varphi}(t, x) = \int_0^x u(t, y) dy.$$

If $u(t, x)$ is any solution to Burgers' equation, then $\tilde{\varphi}(t, x)$ satisfies (8.73). Integrating both sides of the latter equation with respect to x , we conclude that

$$\tilde{\varphi}_t = \gamma \tilde{\varphi}_{xx} + \gamma \alpha \tilde{\varphi}_x^2 + g(t),$$

for some integration “constant” $g(t)$. Thus, unless $g(t) \equiv 0$, our potential function $\tilde{\varphi}$ doesn't satisfy the potential Burgers' equation (8.72), but that's because we chose the “wrong” potential. Indeed, if we define

$$\varphi(t, x) = \tilde{\varphi}(t, x) - \eta(t), \quad \text{where} \quad \eta'(t) = g(t),$$

then

$$\varphi_t = \tilde{\varphi}_t - g(t) = \gamma \tilde{\varphi}_{xx} + \gamma \alpha \tilde{\varphi}_x^2 = \gamma \varphi_{xx} + \gamma \alpha \varphi_x^2,$$

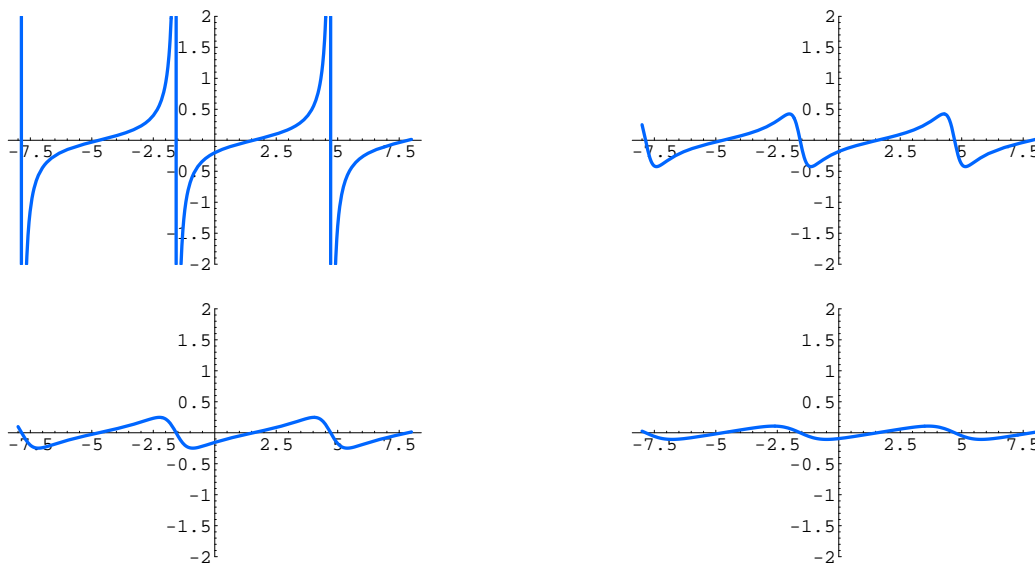


Figure 8.5. A Solution to Burgers' Equation.

and hence the modified potential $\varphi(t, x)$ is a solution to the potential Burgers' equation (8.72). From this it easily follows that

$$v(t, x) = e^{-\varphi(t, x)/(2\gamma)} \quad (8.76)$$

is a positive solution to the heat equation, from which $u(t, x)$ can be recovered via the Hopf–Cole transformation (8.75). Thus, we have proved that every solution to Burgers' equation comes from a positive solution to the heat equation via the Hopf–Cole transformation.

Example 8.9. 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

$$u(t, x) = \frac{2\gamma b \omega e^{-\gamma\omega^2 t} \sin \omega x}{a + b e^{-\gamma\omega^2 t} \cos \omega x} \quad (8.77)$$

to Burgers' equation. A representative example is plotted in Figure 8.5. 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 8.5. This family of solutions is primarily affected by the viscosity term, and rapidly decays to zero.

To solve the initial value problem (8.66–67) for Burgers' equation, we note that, under the Hopf–Cole transformation (8.75),

$$v(0, x) = \exp\left(-\frac{\varphi(0, x)}{2\gamma}\right) = \exp\left(-\frac{1}{2\gamma} \int_0^x f(y) dy\right) \equiv h(x). \quad (8.78)$$

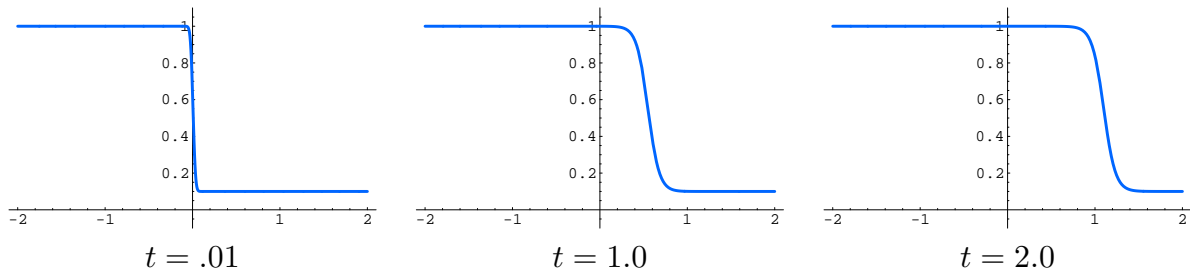


Figure 8.6. Shock Wave Solution to Burgers' Equation.

Remark: The lower limit of the integral can be changed from 0 to any other convenient value without affecting the final form of $u(t, x)$ in (8.75). The only effect is to multiply $v(t, x)$ by an overall constant, which does not change $u(t, x)$.

According to formula (8.15) (adapted to general diffusivity, as in Exercise ■), the solution to the initial value problem (8.70, 78) 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, the solution to the Burgers' initial value problem (8.66–67), valid for $t > 0$, is

$$u(t, x) = \frac{2\gamma}{v(t, x)} \frac{\partial v}{\partial x}, \quad \text{where} \quad v(t, x) = \int_{-\infty}^{\infty} e^{-H(t, x; \xi)} d\xi, \quad (8.79)$$

$$H(t, x; \xi) = \frac{1}{2\gamma} \int_0^\xi f(\eta) d\eta + \frac{(x - \xi)^2}{4\gamma t}.$$

Example 8.10. To demonstrate the smoothing effect of the diffusion terms, let us see what happens to the initial data

$$u(0, x) = \begin{cases} a, & x < 0, \\ b, & x > 0, \end{cases} \quad (8.80)$$

in the form of a step function. We assume that $a > b$, which would correspond to a shock wave in the inviscid limit $\gamma = 0$. (In Exercise ■, the reader is asked to analyze the case $a < b$ which corresponds to a rarefaction wave.) In this case,

$$H(t, x; \xi) = \frac{(x - \xi)^2}{4\gamma t} + \begin{cases} \frac{a\xi}{2\gamma}, & \xi < 0, \\ \frac{b\xi}{2\gamma}, & \xi > 0. \end{cases} \quad (8.81)$$

After some algebraic manipulations, the solution (8.79) is found to have the explicit form

$$u(t, x) = a + (b - a)w(t, x), \quad (8.82)$$

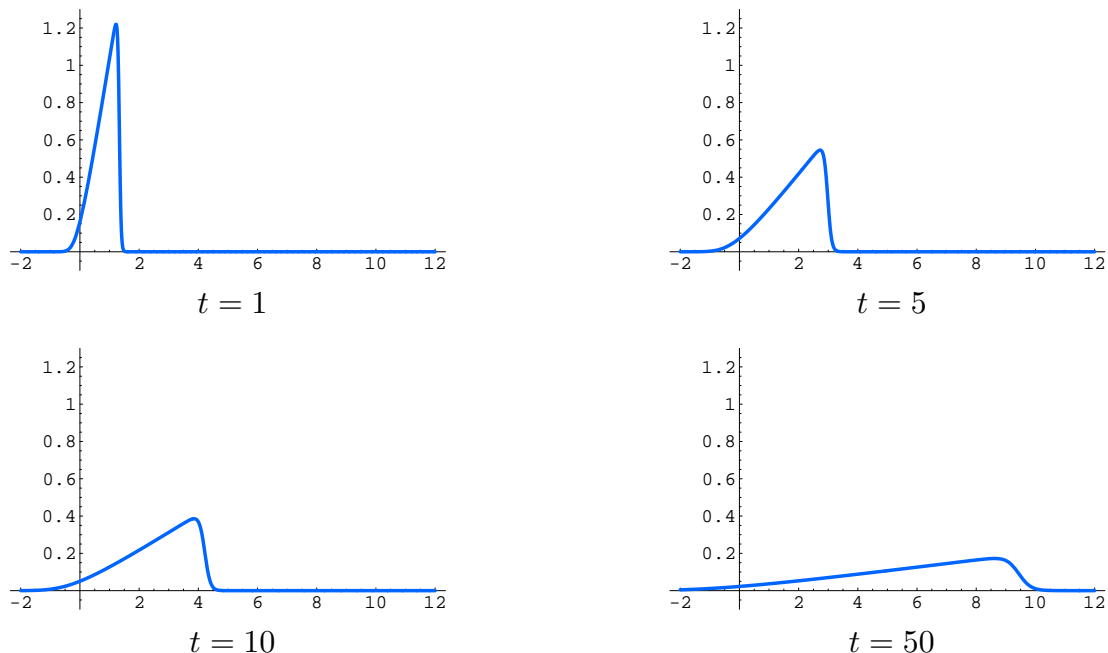


Figure 8.7. Triangular Wave Solution to Burgers' Equation.

where

$$w(t, x) = \frac{1 + \operatorname{erf}\left(\frac{x - bt}{2\sqrt{\gamma t}}\right)}{1 + \operatorname{erf}\left(\frac{x - bt}{2\sqrt{\gamma t}}\right) + \exp\left(\frac{b - a}{2\gamma}(x - ct)\right) \left[1 - \operatorname{erf}\left(\frac{x - at}{2\sqrt{\gamma t}}\right)\right]}, \quad (8.83)$$

with $c = \frac{1}{2}(a + b)$, and $\operatorname{erf} z$ denoting the error function (2.81). The solution, for $a = 1$, $b = .1$, and $\gamma = .03$, is plotted at various times in Figure 8.6. Observe that, as with the heat equation, the sharp transition region is immediately smoothed, and the solution soon assumes the form of a smoothly varying transition region between the two step heights. The larger the diffusion coefficient in relation to the jump discontinuity, the more pronounced the smoothing effect. Moreover, as $\gamma \rightarrow 0$, the solution $u(t, x)$ converges to the shock wave solution (2.45) 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, recalling the limiting behavior of the error function in (2.82), if $x < ct$, then the exponential term in (8.83) tends to ∞ , and hence $w(t, x) \rightarrow 0$ as $\gamma \rightarrow 0$. On the other hand, if $x > ct$, then the exponential term tends to 0, and so $w(t, x) \rightarrow 1$, proving convergence of $u(t, x)$ to the shock wave solution.

Example 8.11. Consider the case when the initial data $u(0, x) = \delta(x)$ is a concentrated delta function impulse at the origin. In the solution formula (8.79), starting the integral for $H(t, x; \xi)$ at 0 is problematic, but as noted earlier, we are free to select any

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

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

We then evaluate

$$v(t, x) = \int_{-\infty}^{\infty} e^{-H(t, x; \xi)} d\xi = \sqrt{\pi\gamma t} \left[(1 + e^{-1/(2\gamma)}) + (1 - e^{-1/(2\gamma)}) \operatorname{erf} \left(\frac{x}{2\sqrt{\gamma t}} \right) \right].$$

Therefore, the solution to the initial value problem is

$$u(t, x) = \frac{2\gamma}{v(t, x)} \frac{\partial v}{\partial x} = 2 \sqrt{\frac{\gamma}{\pi t}} \frac{e^{-x^2/(4\gamma t)}}{\coth \left(\frac{1}{4\gamma} \right) - \operatorname{erf} \left(\frac{x}{2\sqrt{\gamma t}} \right)}, \quad (8.84)$$

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 8.7. As you can see, the initial concentration diffuses out, but, unlike the heat equation, the wave does not remain symmetric owing to the additional nonlinear advection term. The effect is that the wave steepens in front and spreads out in back as it propagates. Eventually, as the diffusion progresses, the triangular wave becomes vanishingly small.

8.5. Dispersion and Solitons.

In this section, we venture beyond the usual realm of second order equations. While considerably less common than those of first and second order, higher order partial differential equations do arise in certain applications, particularly in models for wave motion, [132], and plate mechanics, [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 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 serves to induce the 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 evolve, dispersive waves will decay even while they conserve energy. Waves on the surface of a liquid are good examples of dispersive waves — an initially concentrated disturbance, caused by throwing a rock in a pond, will spread out as its different Fourier components move off at different speeds.

Our second example is a remarkable nonlinear third order evolution equation known as the Korteweg–deVries equation, that combines dispersive effects with nonlinear transport. As with Burgers' equation (but for very different mathematical reasons), the dispersive term stabilizes 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, with taller waves moving faster. Even more remarkable are the interaction properties of these solitary waves. One ordinarily expects a nonlinearity to induce very complicated and unpredictable behavior. However, when two solitary wave solutions to the Korteweg–deVries 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 by the American mathematicians Martin Kruskal and Norman Zabusky in the mid 1960's, who distinguished such solutions with the neologism *soliton*. It was then found that solitons occur 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, theoretical physics, and has even contributed to the solution of long outstanding problems in complex function theory. Further development of the amazing properties of integrable soliton equations can be found in [1, 43].

Linear Dispersion

The simplest nontrivial third order partial differential equation is the linear equation

$$u_t + u_{xxx} = 0, \tag{8.85}$$

which is a simple model for the unidirectional[†] propagation of linear dispersive waves. To avoid complications engendered by boundary conditions, we shall only look at solutions on the entire line, so $-\infty < x < \infty$. Since the equation only involves first order time derivatives, one expects its solutions to be uniquely specified by a single initial condition

$$u(0, x) = f(x), \quad -\infty < x < \infty. \tag{8.86}$$

In wave mechanics, $u(t, x)$ represents the height of the fluid at time t and position x , and the initial condition (8.86) 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 the solution $u(t, \cdot) \in L^2(\mathbb{R})$ remains square integrable at all times t (a fact that can be justified a priori), let

$$\hat{u}(t, k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u(t, x) e^{-ikx} dx$$

be its spatial Fourier transform. Owing to the effect of the Fourier transform on derivatives, as formulated in Corollary 7.7, the Fourier transform converts the partial differential

[†] 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 relies on an (approximate) factorization of the underlying differential operator.

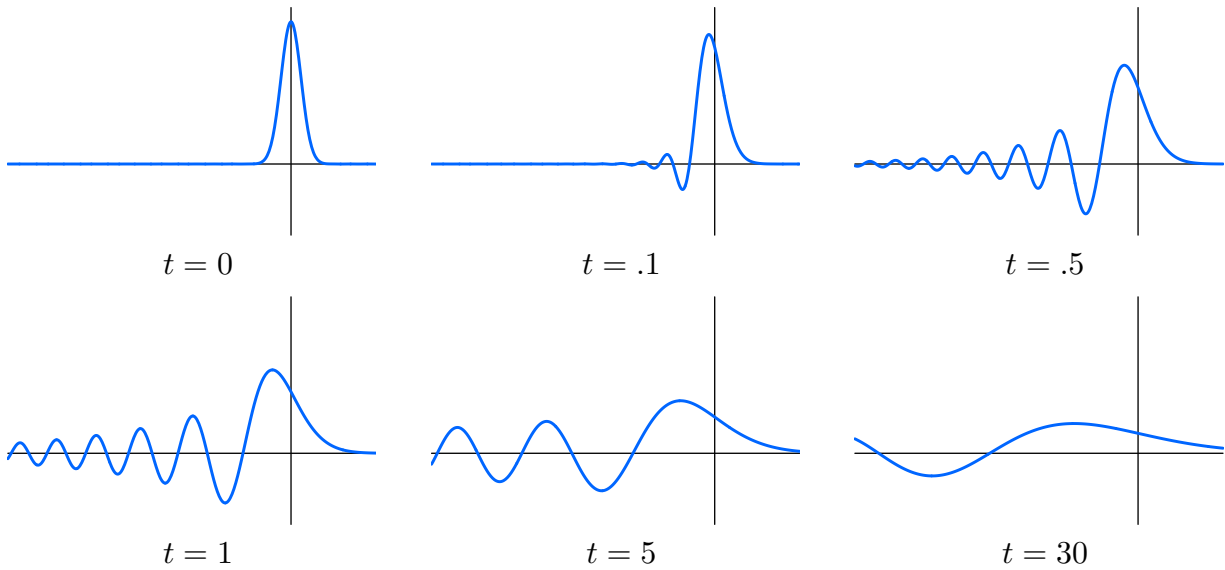


Figure 8.8. Gaussian Solution to the Dispersive Wave Equation.

equation (8.85) into a first order, linear ordinary differential equation

$$\frac{\partial \hat{u}}{\partial t} + (ik)^3 \hat{u} = \frac{\partial \hat{u}}{\partial t} - ik^3 \hat{u} = 0, \quad (8.87)$$

in which the frequency variable k appears as a parameter. The corresponding initial conditions

$$\hat{u}(0, k) = \hat{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ikx} dx \quad (8.88)$$

are provided by the Fourier transform of (8.86). The solution to the initial value problem (8.87–88) is

$$\hat{u}(t, k) = \hat{f}(k) e^{ik^3 t}.$$

Inverting the Fourier transform yields the explicit formula for the solution

$$u(t, x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(k) e^{i(kx+k^3 t)} dk \quad (8.89)$$

to the initial value problem for the dispersive wave equation (8.85–86).

Example 8.12. Suppose that the initial profile

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

is a Gaussian. According to our Table of Fourier transforms,

$$\hat{f}(k) = \frac{e^{-k^2/4}}{\sqrt{2}},$$

and hence the corresponding solution to the dispersive wave equation (8.85) is

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

the imaginary part vanishes thanks to the oddness of the integrand. (Indeed, the solution must be real.) A plot of the solution at various times appears in Figure 8.8. 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 $u(t, x) \rightarrow 0$ uniformly as $t \rightarrow \infty$, even though, according to Exercise ■, 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 (constant).

Example 8.13. The *fundamental solution* to the dispersive wave equation corresponds to 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 (8.89) is

$$u(t, x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i(kx + k^3 t)} dk = \frac{1}{\pi} \int_0^{\infty} \cos(kx + k^3 t) dk, \quad (8.90)$$

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 does not converge because the integrand does not go to zero as $|k| \rightarrow \infty$. However, the increasingly rapid oscillations induced by the cubic term tend to cancel each other out and produce a convergent integral. To prove this, we perform a (non-obvious) integration by parts:

$$\begin{aligned} \int_0^l \cos(kx + k^3 t) dk &= \int_0^l \frac{1}{x + 3k^2 t} \frac{d}{dk} \sin(kx + k^3 t) dk \\ &= \frac{\sin(kx + k^3 t)}{x + 3k^2 t} \Big|_{k=0}^l - \int_0^l \frac{d}{dk} \left(\frac{1}{x + 3k^2 t} \right) \sin(kx + k^3 t) dk \\ &= \frac{\sin(lx + l^3 t)}{x + 3l^2 t} + \int_0^l \frac{6kt \sin(kx + k^3 t)}{(x + 3k^2 t)^2} dk. \end{aligned} \quad (8.91)$$

As $l \rightarrow \infty$, as long as $t \neq 0$, the first term on the right goes to zero, while the final integral converges absolutely due to the rapid decay of the integrand.

While the integral in the solution formula (8.90) cannot be evaluated in terms of elementary functions, it is related to the integral defining the *Airy function*:

$$\text{Ai}(z) = \frac{1}{\pi} \int_0^{\infty} \cos \left(sz + \frac{1}{3} s^3 \right) ds, \quad (8.92)$$

an important special function that was first employed by the nineteenth century British applied mathematician George Airy in his studies of optical caustics (e.g., a lens or magnifying glass) and rainbows. Indeed, applying the change of variables

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

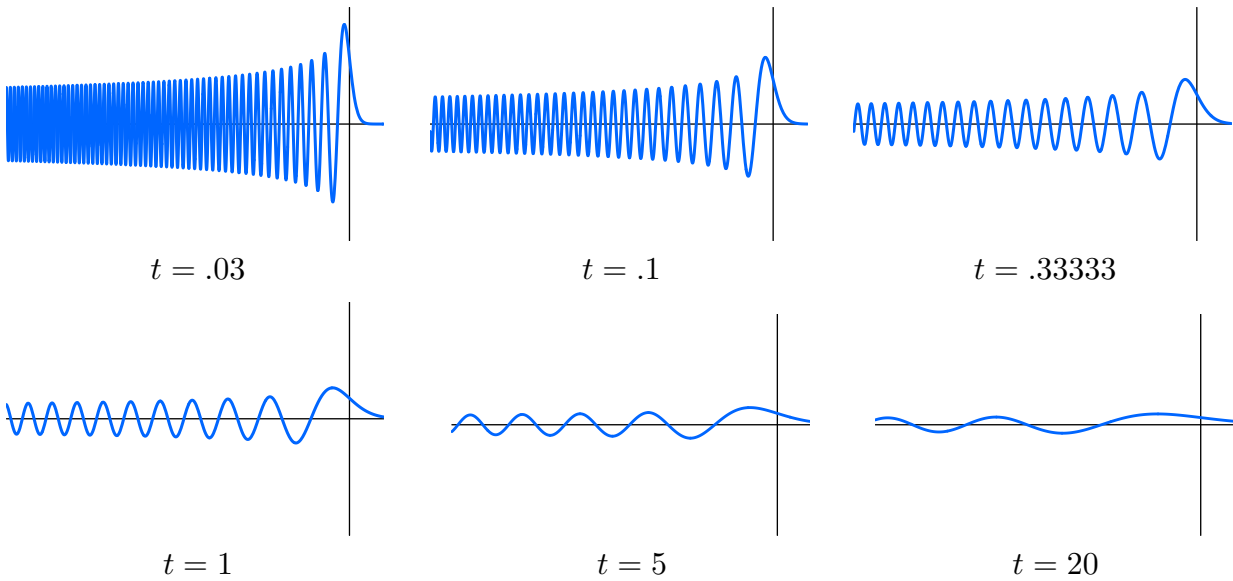


Figure 8.9. Fundamental Solution to the Dispersive Wave Equation.

to the Airy function integral (8.92), we deduce that the fundamental solution to the dispersive wave equation (8.85) can be written as

$$u(t, x) = \frac{1}{\sqrt[3]{3t}} \operatorname{Ai}\left(\frac{x}{\sqrt[3]{3t}}\right). \quad (8.93)$$

See Figure 8.9 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 function is to induce a highly oscillatory wave trailing off to $-\infty$. As time progresses, the dispersive effects cause the oscillations to spread out and their overall amplitude to decay in proportion to $t^{-1/3}$. On the other hand, as $t \rightarrow 0$, the solution becomes more and more oscillatory for negative x , and so converges *weakly* to the initial delta function. We also note that (8.93) is a similarity solution of the dispersive wave equation, since it is invariant under the scaling symmetry

$$(t, x, u) \mapsto (\lambda^3 t, \lambda x, \lambda u).$$

Equation (8.93) gives the response to an initial delta function concentrated at the origin. By translation invariance, we immediately deduce that

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

is the *fundamental solution* corresponding to an initial delta impulse at $x = \xi$. Therefore, we can use linear superposition to find an explicit formula for the solution to the initial value problem that bypasses the Fourier transform. Namely, writing the general initial data as a superposition of delta functions

$$u(0, x) = f(x) = \int_{-\infty}^{\infty} f(\xi) \delta(x - \xi) d\xi,$$

we conclude that the resulting solution is the self-same combination of fundamental solutions

$$u(t, x) = \frac{1}{\sqrt[3]{3t}} \int_{-\infty}^{\infty} f(\xi) \operatorname{Ai} \left(\frac{x - \xi}{\sqrt[3]{3t}} \right) d\xi. \quad (8.94)$$

The Dispersion Relation

As noted earlier, a key feature of the third order wave equation (8.85) is that waves disperse, in the sense that those of different frequencies move at different speeds. Our goal is to understand the dispersion process. To streamline the argument, we replace real trigonometric solutions by complex exponentials, and so look at a solution whose initial profile

$$u(0, x) = e^{ikx}$$

is a complex oscillatory function of *wave number* k . (We will be observing oscillations in both space and time, and will reserve the term “frequency” for the time oscillations.) Since the initial data does not decay as $|x| \rightarrow \infty$, we cannot use Fourier integral solution formula (8.89) directly. Instead, anticipating the induced wave to exhibit temporal oscillations, let us try an exponential solution ansatz

$$u(t, x) = e^{i(kx - \omega t)} \quad (8.95)$$

representing a complex oscillatory wave of temporal *frequency* ω and *wave number* (spatial frequency) k . Since

$$\frac{\partial u}{\partial t} = -i\omega e^{i(kx - \omega t)}, \quad \frac{\partial^3 u}{\partial x^3} = -ik^3 e^{i(kx - \omega t)},$$

(8.95) satisfies the partial differential equation (8.85) if and only if its frequency and wave number satisfy the *dispersion relation*

$$\omega = -k^3. \quad (8.96)$$

Therefore, the exponential solution (8.95) of wave number k takes the form

$$u(t, x) = e^{i(kx + k^3 t)}. \quad (8.97)$$

Our Fourier transform formula (8.89) for the solution can thus be viewed as a 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 (8.95). After canceling the common exponential factors, the result is an equation expressing the frequency ω as a function of the wave number k .

Any exponential solution (8.95) is automatically in the form of a traveling wave, since we can write

$$u(t, x) = e^{i(kx - \omega t)} = e^{ik(x - c_p t)}, \quad \text{where} \quad c_p = \frac{\omega}{k} \quad (8.98)$$

is the *wave speed* or, as it is more usually called, the *phase velocity*. If the dispersion relation is linear in the wave number, $\omega = ck$, as occurs in the linear transport equation

$u_t + cu_x$, then all waves move at an identical speed 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 frequency move at different speeds. In the particular case (8.85), those with wave number k move at speed $c = \omega/k = -k^2$, 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.

The general solution to the linear partial differential equation under consideration is then built up by linear superposition of the exponential solutions,

$$u(t, x) = \int_{-\infty}^{\infty} e^{i(kx - \omega t)} g(k) dk, \quad (8.99)$$

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 wave packet represented by (8.99) is less evident. To determine the speed of propagation of the packet, let us switch to a moving coordinate frame of speed c by setting $x = ct + \xi$. The solution formula (8.99) then becomes

$$u(t, ct + \xi) = \int_{-\infty}^{\infty} e^{i(ck - \omega)t} e^{ik\xi} g(k) dk. \quad (8.100)$$

For fixed ξ , the integral is of the form

$$H(t) = \int_{-\infty}^{\infty} e^{i\varphi(k)t} h(k) dk, \quad (8.101)$$

where, in our case, $\varphi(k) = ck - \omega(k)$ and $h(k) = e^{ik\xi} g(k)$. We are interested in understanding the behavior of (8.101) as $t \rightarrow \infty$. Now, if $\varphi(k) = k$, then (8.101) is just a Fourier integral, (7.9), and, as we learned in Chapter 7, $H(t) \rightarrow 0$ as $t \rightarrow \infty$, 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 $\tilde{k} = \varphi(k)$ to convert that part of the integral to a Fourier integral, and again the increasingly rapid oscillations cause the limit to vanish. In this fashion, we are led to the key insight of the *Method of Stationary Phase*, developed by the nineteenth century Scottish physicist Lord Kelvin (William Thomson) to study wave propagation. Namely, for large $t \gg 0$, the primary contribution to the highly oscillatory integral (8.101) occurs at the *stationary points* of the phase function, where $\varphi'(k) = 0$. A rigorous justification of the method, along with precise error bounds, can be found in [102].

In the present context, the Method of Stationary Phase implies that the main contribution to the integral (8.100) occurs when

$$0 = \frac{d}{dk} (\omega - ck) = \frac{d\omega}{dk} - c. \quad (8.102)$$

Thus, surprisingly, the main effect on the solution of the waves with wave number k is felt when moving, not at the phase velocity c_p , but rather at the *group velocity*

$$c_g = \frac{d\omega}{dk} . \quad (8.103)$$

Unless the dispersion relation is linear in the wave number, the group velocity (8.103) is *not the same* as the phase velocity (8.98).

The phase velocity $c_p = \omega(k)/k$ governs the speed of propagation of an individual wave of wave number k . On the other hand, the more subtle group velocity $c_g = \omega'(k)$ determines the speed of propagation of the energy associated with wave number k , and, in a dispersive environment, these are different! For example, in the case of the dispersive wave equation (8.85), $\omega = -k^3$ and so $c_g = -3k^2$, 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 Figures 8.8, 8.9: while the bulk of the disturbance is spreading out rather rapidly to the left, the individual wave crests are moving much slower. The distinction between group velocity and phase velocity is well understood by surfers. In shallow water, the wave energy outpaces the individual crests, and so waves that are large far out at sea will be diminished in size by the time they reach the shore, as their energy amplifies the waves that are breaking ahead of them.

On the other hand, the dispersion relation associated with deep water waves is (ignoring physical constants) $\omega = \sqrt{k}$, [132]. 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 — the opposite of the shallow water dynamics. For an experimental verification, just throw a stone in a still pond and follow each individual wave crest. It starts out very small far behind the disturbance, steadily grows as it moves through the disturbance, but eventually subsiding and disappearing into the still water ahead of the expanding wave packet triggered by the stone.

The Korteweg–de Vries Equation

The simplest wave model that combines dispersion with nonlinearity is the celebrated *Korteweg–de Vries equation*

$$u_t + u_{xxx} + uu_x = 0. \quad (8.104)$$

The equation was first derived by the French applied mathematician Boussinesq, [21; eq. (30)], [22; eqs. (283, 291)], in 1872 as a model for surface water waves. Two decades later, it was rediscovered by the Dutch applied mathematician Diederik Korteweg and his student Gustav de Vries, [81], and, despite Boussinesq's priority, is named after them. In the early 1960's, the American mathematical physicists Martin Kruskal and Norman Zabusky, [136], used the Korteweg–de Vries equation as a continuum model for a system of nonlinear mass-spring chains studied by Fermi, Pasta and Ulam, [47]. Numerical experimentation indicated its many remarkable properties, which were soon rigorously established. Their work sparked the rapid development one of the most remarkable and far-ranging discoveries of the modern era: integrable nonlinear partial differential equations, [1, 43].

The most important special solutions to the Korteweg–deVries equation are the *traveling waves*. We seek solutions

$$u = v(\xi) = v(x - ct), \quad \text{where} \quad \xi = x - ct,$$

that have a fixed profile while translating to the right with speed c . By the chain rule,

$$\frac{\partial u}{\partial t} = -cv'(\xi), \quad \frac{\partial u}{\partial x} = v'(\xi), \quad \frac{\partial^3 u}{\partial x^3} = v'''(\xi).$$

Substituting these expressions into the Korteweg–deVries equation (8.104), we conclude that $v(\xi)$ must satisfy the nonlinear third order ordinary differential equation

$$v''' + vv' - cv' = 0. \quad (8.105)$$

Let us further assume that the traveling wave is *localized*, meaning that the solution and its derivatives are vanishingly small at large distances:

$$\lim_{x \rightarrow \pm\infty} u(t, x) = \lim_{x \rightarrow \pm\infty} \frac{\partial u}{\partial x}(t, x) = \lim_{x \rightarrow \pm\infty} \frac{\partial^2 u}{\partial x^2}(t, x) = 0. \quad (8.106)$$

This implies that we should impose the boundary conditions

$$\lim_{\xi \rightarrow \pm\infty} v(\xi) = \lim_{\xi \rightarrow \pm\infty} v'(\xi) = \lim_{\xi \rightarrow \pm\infty} v''(\xi) = 0, \quad (8.107)$$

on the traveling wave ordinary differential equation.

The ordinary differential equation (8.105) can, in fact be solved in closed form. First, note that

$$\frac{d}{d\xi} (v'' + \frac{1}{2}v^2 - cv) = 0, \quad \text{and hence} \quad v'' + \frac{1}{2}v^2 - cv = a,$$

is a first integral, with a indicating the constant of integration. However, the localizing boundary conditions (8.107) imply that $a = 0$. Multiplying the latter equation by v' allows us to integrate a second time

$$v'(v'' + \frac{1}{2}v^2 - cv) = \frac{d}{d\xi} [\frac{1}{2}(v')^2 + \frac{1}{6}v^3 - \frac{1}{2}cv^2] = 0.$$

Thus,

$$\frac{1}{2}(v')^2 + \frac{1}{6}v^3 - \frac{1}{2}cv^2 = b,$$

where b is a second constant of integration, which, again by the boundary conditions (8.107), is also zero: $b = 0$. We conclude that $v(\xi)$ satisfies the first order ordinary differential equation

$$\frac{dv}{d\xi} = v \sqrt{c - \frac{1}{3}v}.$$

We integrate this autonomous equation by the usual separation of variables method:

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

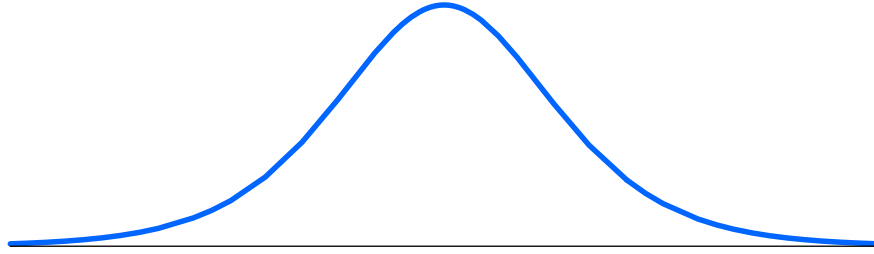


Figure 8.10. Solitary Wave.

Consulting a table of integrals, and then solving for v , we conclude that the solution has the form

$$v(\xi) = 3c \operatorname{sech}^2\left(\frac{1}{2}\sqrt{c}\xi + \delta\right),$$

where

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

is the *hyperbolic secant function*. The solution has the form graphed in Figure 8.10. It is a symmetric, monotone, exponentially decreasing function on either side of its maximum height of $3c$. (Despite looking similar, it is *not* a Gaussian.) The resulting localized traveling wave solutions to the Korteweg–deVries equation are thus

$$u(t, x) = 3c \operatorname{sech}^2\left[\frac{1}{2}\sqrt{c}(x - ct) + \delta\right], \quad (8.108)$$

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.

The solution (8.108) is known as a *solitary wave solution* since it represents a localized wave that travels unchanged in shape. Solitary waves were first observed in nature by the British engineer J. Scott Russell, [118], who recounts how such a wave was generated by the sudden motion of a barge along an Edinburgh canal. Scott Russell ended up chasing the propagating wave on horseback for several miles. Russell’s observations were dismissed by his contemporary Airy, who, relying on his linearly dispersive model for surface waves, claimed that such localized disturbances could not exist. Much later, Boussinesq established the proper nonlinear surface wave model (8.104), valid for long waves in shallow water, and then derived the solitary wave solution (8.108), thereby fully exonerating Scott Russell’s insight.

It took almost a century before the truly remarkable properties of these solutions was recognized — initially in numerical simulations, [136], and then rigorously. 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–deVries equation, suppose we start with initial conditions representing a taller solitary wave to the left of a shorter one. As time evolves, the taller wave will move faster, and so catch up to the shorter wave. They then experience a complicated nonlinear interaction, as expected. But, remarkably, after a while they

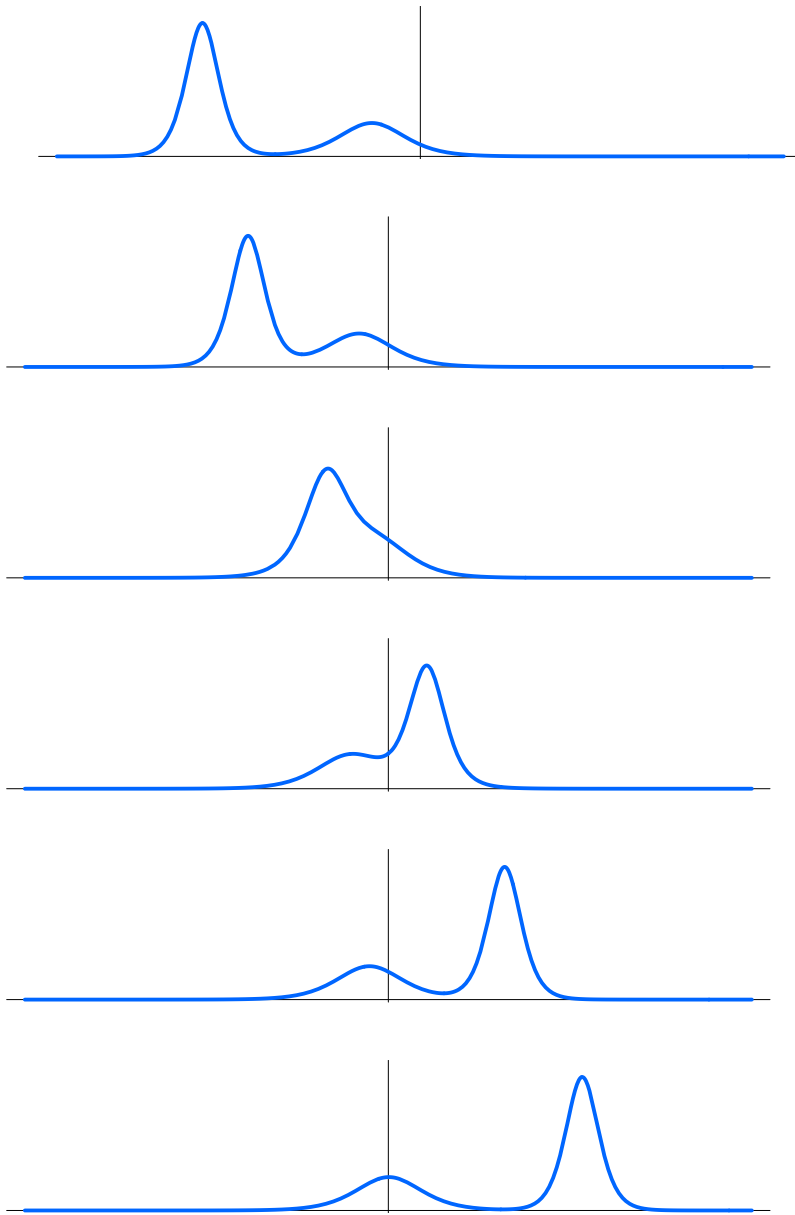


Figure 8.11. Interaction of Two Solitons.

emerge from the interaction unscathed. The smaller wave is now in back and the larger one in front. Afterwards, they proceed independently along their way, with the smaller one lagging 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 8.11 plots a typical such interaction. Owing to this “particle-like” behavior under interaction, these solutions have been endowed with a special name: *soliton*. The general formula for

a two soliton solution to the Korteweg–deVries equation is

$$u(t, x) = 12 \frac{\partial^2}{\partial x^2} \log \Delta(t, x), \quad (8.109)$$

where

$$\Delta(t, x) = \det \begin{pmatrix} 1 + \varepsilon_1(t, x) & \frac{2c_1}{c_1 + c_2} \varepsilon_2(t, x) \\ \frac{2c_2}{c_1 + c_2} \varepsilon_1(t, x) & 1 + \varepsilon_2(t, x) \end{pmatrix}, \quad (8.110)$$

and where

$$\varepsilon_j(t, x) = \exp(c_j^3 t - c_j x + \delta_j). \quad (8.111)$$

The constants c_1, c_2 represent the wave speeds of the two solitons, while δ_1, δ_2 are phase shifts. Proving that (8.109) is indeed a solution to the Korteweg–deVries equation is a straightforward, albeit tedious, exercise in differentiation. In Exercise ■, the reader is asked to investigate the asymptotic behavior, as $t \rightarrow \pm\infty$, of the solution and prove that it does, indeed, break up into two solitons.

A similar dynamic occurs with multiple soliton solutions. Faster solitons catch up slower ones lying to their right. After the various solitons finish colliding and interacting, they emerge in order, from smallest to largest, each moving at a distinct speed. As a result, each individual soliton becomes more and more separated from the others. The explicit formula for the n soliton solution can be found in Exercise ■. Furthermore, it can be shown that, starting with an *arbitrary* localized initial disturbance, the resulting solution eventually emits a finite number of solitons of different heights, moving off in increasing height order (from right to left) followed by a small dispersive tail that gradually disappears.

The source of these highly non-obvious multi-soliton solution formulas is a deep matter, lying beyond the scope of these introductory notes. Gardner, Green, Kruskal and Miura discovered a profound connection between the solutions to the Korteweg–deVries equation and the eigenvalues of the Schrödinger boundary value problem

$$-\frac{d^2\psi}{dx^2} + 6u(t, x)\psi = \lambda\psi \quad (8.112)$$

on the real line. The key result is that whenever $u(t, x)$ is a localized solution to the Korteweg–deVries equation (8.104), the eigenvalues of the Schrödinger equation (8.112) are constant, and do not vary with the time t . The multi-soliton solutions are the potentials that have only eigenvalues — no continuous spectrum or scattering states, which correspond to the dispersive tail appearing in the more general localized solutions. The solution to the inverse scattering problem, that is reconstructing the potential $u(t, x)$ from its spectrum, produces the determinantal formulas. This inverse scattering solution can be viewed as a nonlinear form of the Fourier transform, that effectively linearizes the Korteweg–deVries equation and explains its many remarkable properties. See [1, 43] for additional details.