

Chapter 22

Nonlinear Partial Differential Equations

The ultimate topic to be touched on in this book is the vast and active field of nonlinear partial differential equations. Leaving aside quantum mechanics, which remains to date an inherently linear theory, most real-world physical systems, including gas dynamics, fluid mechanics, elasticity, relativity, ecology, neurology, thermodynamics, and many more, are modeled by *nonlinear* partial differential equations. Attempts to survey, in such a small space, even a tiny fraction of such an all-encompassing range of phenomena, methods, results, and mathematical developments, are doomed to failure. So we will be content to introduce a handful of prototypical, seminal examples that arise in the study of nonlinear waves and that serve to highlight some of the most significant physical and mathematical phenomena not encountered in simpler linear systems. We will only have space to look at simple one-dimensional models; the far more complicated nonlinear systems that govern our three-dimensional dynamical universe quickly lead one to the cutting edge of contemporary research.

Historically, comparatively little was known about the extraordinary range of behavior exhibited by the solutions to nonlinear partial differential equations. Many of the most fundamental phenomena that now drive modern-day research, including solitons, chaos, stability, blow-up and singularity formation, asymptotic properties, etc., remained undetected or at best dimly perceived in the pre-computer era. The last sixty years has witnessed a remarkable blossoming in our understanding, due in large part to the insight offered by the availability of high performance computers coupled with great advances in the understanding and development of suitable numerical approximation schemes. New analytical methods, new mathematical theories, coupled with new computational algorithms have precipitated this revolution in our understanding and study of nonlinear systems, an activity that continues to grow in intensity and breadth. Each leap in computing power coupled with theoretical advances has led to yet deeper understanding of nonlinear phenomena, while simultaneously demonstrating how far we have yet to go. To make sense of this bewildering variety of methods, equations, and results, it is essential build upon a firm foundation on, first of all, linear systems theory, and secondly, nonlinear algebraic equations and nonlinear ordinary differential equations.

Our presentation is arranged according to the order of the underlying differential equation. First order nonlinear partial differential equations model nonlinear waves and arise in gas dynamics, water waves, elastodynamics, chemical reactions, transport of pollutants, flood waves in rivers, chromatography, traffic flow, and a wide range of biological and ecological systems. One of the most important nonlinear phenomena, with no linear counterpart, is the break down of solutions in finite time, resulting in the formation of

discontinuous shock waves. A striking example is the supersonic boom produced by an airplane that breaks the sound barrier. As in the linear wave equation, the signals propagate along the characteristics, but in the nonlinear case the characteristics can cross each other, precipitating the onset of a shock. The characterization of the shock dynamics requires additional physical information, in the form of a conservation law, that supplements the original partial differential equation.

Parabolic second order partial differential equations govern nonlinear diffusion processes, including thermodynamics, chemical reactions, dispersion of pollutants, and population dynamics. The simplest and most well understood is Burgers' equation, which can, surprisingly, be linearized by transforming it to the heat equation. This accident provides an essential glimpse into the world of nonlinear diffusion processes. In the limit, as the diffusion or viscosity tends to zero, the solutions to Burgers' equation tend to the shock wave solutions to the limiting first order dispersionless equation, and thus provides an alternate mechanism for unraveling shock dynamics.

Third order partial differential equations arise in the study of dispersive wave motion, including water waves, plasma waves, waves in elastic media, and elsewhere. We first treat the basic linear dispersive model, comparing and contrasting it with the hyperbolic models we encountered earlier in this text. The distinction between group and wave velocity — observed when, for instance, surface waves propagate over water — is developed. Finally, we introduce the remarkable Korteweg–deVries equation, which serves as a model for waves in shallow water, waves in plasmas, and elsewhere. Despite its intrinsic nonlinearity, it supports stable localized traveling wave solutions, known as *solitons*, that, remarkably, maintain their shape even under collision. The Korteweg–deVries equation is the prototypical example of an integrable system, and this discovery in the mid 1960's inaugurated intense and ongoing research in the remarkable physical models that exhibit integrability, a development that has had many ramifications in both pure and applied mathematics.

22.1. Nonlinear Waves and Shocks.

Before attempting to tackle any nonlinear partial differential equations, we must carefully review the solution to the simplest linear first order partial differential equation.

Linear Transport and Characteristics

The *transport equation*

$$u_t + cu_x = 0, \tag{22.1}$$

is so named because it models the transport of, say, a pollutant in a uniform fluid flow. Let us begin by assuming that the *wave speed* c is constant. According to Proposition 14.8, every solution is constant along the characteristic lines of slope

$$\frac{dx}{dt} = c, \quad \text{namely} \quad x - ct = \text{constant}. \tag{22.2}$$

As a consequence, the solutions are *traveling waves* of the form

$$u(t, x) = p(x - ct), \tag{22.3}$$

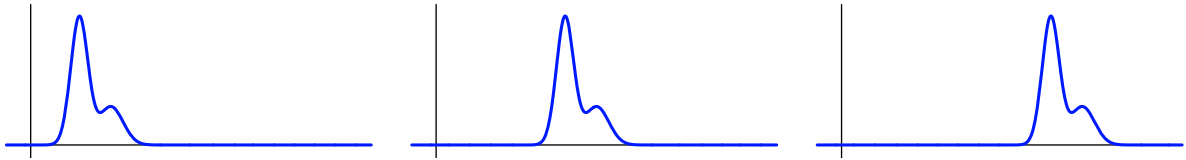


Figure 22.1. Traveling Wave.

where $p(\xi)$ is an arbitrary function of the *characteristic variable* $\xi = x - ct$. To a stationary observer, the solution (22.3) appears as a wave of unchanging form moving at velocity c . When $c > 0$, the wave translates to the right, as illustrated in Figure 22.1. When $c < 0$, the wave moves to the left, while $c = 0$ corresponds to a permanent wave form that remains fixed at its original location.

Slightly more complicated, but still linear, is the non-uniform transport equation

$$u_t + c(x)u_x = 0, \quad (22.4)$$

where the wave velocity $c(x)$ depends upon the spatial position. This equation models unidirectional waves propagating through a non-uniform, but static medium. Generalizing the construction (22.2), we define a *characteristic curve* to be a solution to the autonomous ordinary differential equation

$$\frac{dx}{dt} = c(x). \quad (22.5)$$

Thus, unlike the constant velocity version, the characteristics are no longer necessarily straight lines. Nevertheless, the preceding observation remains valid:

Proposition 22.1. *Solutions to the linear transport equation (22.4) are constant on its characteristic curves.*

Proof: Let $x(t)$ be a characteristic curve, i.e., a solution to (22.5), parametrized by the time t . Let $h(t) = u(t, x(t))$ be the value of the solution at the point $(t, x(t))$ on the given characteristic curve. Our goal is to prove that $h(t)$ is a constant function of t , and, as usual, this is done by proving that its derivative is identically zero. To differentiate $h(t)$, we invoke the chain rule:

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

We replaced dx/dt by $c(x)$ since we are assuming that $x(t)$ is a characteristic curve, and hence satisfies (22.5). The final combination of derivatives is zero whenever u solves the transport equation (22.4). *Q.E.D.*

Since the characteristic curve differential equation (22.5) is autonomous, it can be immediately solved:

$$b(x) \equiv \int \frac{dx}{c(x)} = t + k, \quad (22.6)$$

where k is the constant of integration. Thus, the characteristic curves are “parallel”, each being a translate of the graph of $t = b(x)$ in the direction of the t axis. The characteristic

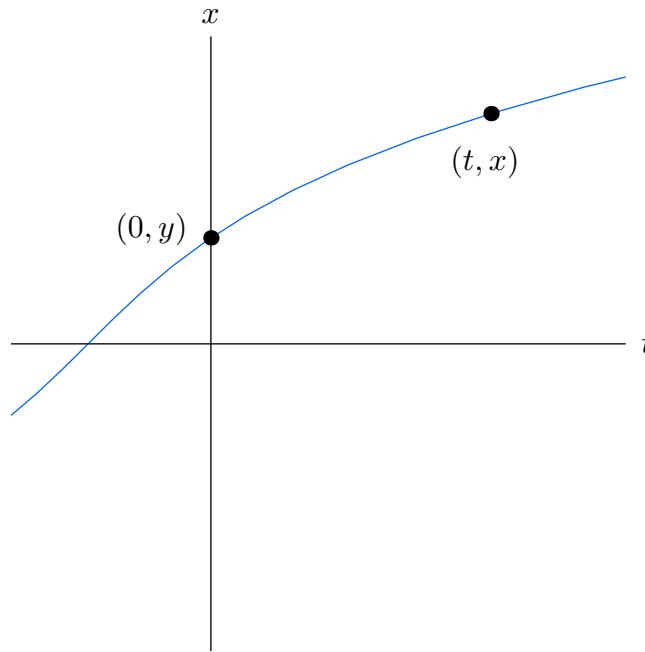


Figure 22.2. Characteristic Curve.

curves are therefore defined by the formula $x = g(t + k)$, where $g = b^{-1}$ is the inverse function. (See Section 20.1 for full details.)

Observe that the characteristic curves are the level sets of the *characteristic variable* $\xi = b(x) - t$. As a consequence, any function which is constant along the characteristic curves depends only on the value of the characteristic variable at each point, and hence takes the form

$$u(t, x) = p(b(x) - t) \quad (22.7)$$

for some function $p(\xi)$. In other words, the characteristic curves are the common level curves of *all* solutions to the transport equation. It is easy to check directly that, provided $b(x)$ is defined by (22.6), $u(t, x)$ solves the partial differential equation (22.4) for *any* choice of function $p(\xi)$.

To find the solution that satisfies the prescribed initial conditions

$$u(0, x) = f(x) \quad (22.8)$$

we merely substitute the general solution formula (22.7). This leads to the equation

$$p(b(x)) = f(x), \quad \text{and, therefore,} \quad p(\xi) = f \circ b^{-1}(\xi) = f(g(\xi)).$$

The resulting solution formula has a simple graphical interpretation: to find its value $u(t, x)$ at a given point, we look at the characteristic curve passing through (t, x) . If this curve intersects the x axis at the point $(0, y)$, then $u(t, x) = u(0, y) = f(y)$. The construction is illustrated in Figure 22.2. Incidentally, if the characteristic curve through (t, x) doesn't intersect the x axis, the solution value $u(t, x)$ is not prescribed by the initial data.

Example 22.2. Let us solve the particular transport equation

$$\frac{\partial u}{\partial t} + \frac{1}{x^2 + 1} \frac{\partial u}{\partial x} = 0 \quad (22.9)$$

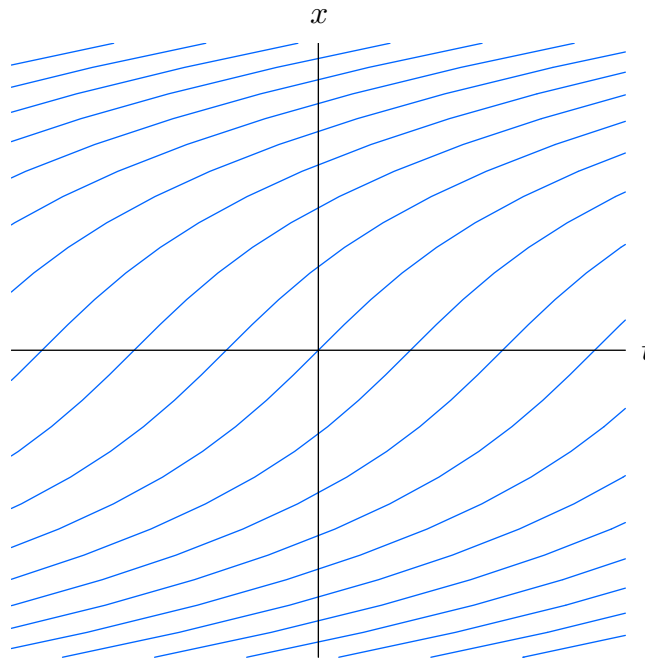


Figure 22.3. Characteristic Curves for $u_t + \frac{1}{x^2 + 1} u_x = 0$.

by the method of characteristics. According to (22.5), the characteristic curves satisfy the first order ordinary differential equation

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

Separating variables and integrating, we find

$$\int (x^2 + 1) dx = \frac{1}{3} x^3 + x = t + k,$$

where k is the integration constant. Some of the resulting characteristic curves are plotted in Figure 22.3.

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

$$u = p\left(\frac{1}{3} x^3 + x - t\right),$$

where $p(\xi)$ is an arbitrary function. A typical solution, corresponding to initial data

$$u(t, 0) = \frac{1}{1 + (x + 2.75)^2},$$

is plotted at times $t = 0, 2, 5, 10, 25, 50$ in Figure 22.4. The fact that the characteristic curves are not straight means that, although the solution remains constant along each individual curve, a stationary observer will witness a dynamically changing profile as the wave moves along through the non-uniform medium. The wave speeds up as it approaches the origin, and then slows back down once it passes and moves off to the right. As a result, we observe the wave spreading out as it approaches the origin, and then contracting as it moves off to the right.

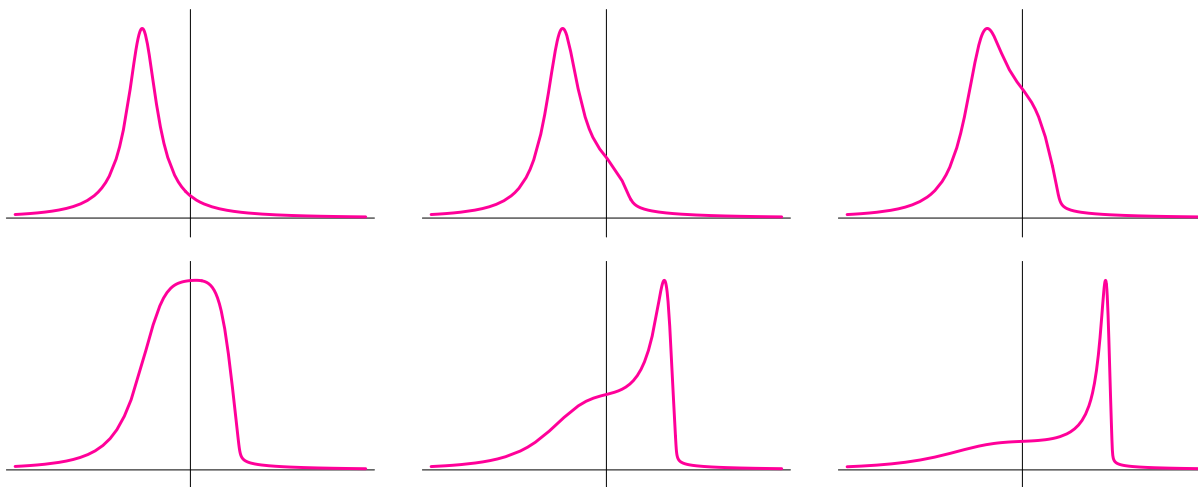


Figure 22.4. Solution to $u_t + \frac{1}{x^2 + 1} u_x = 0$.

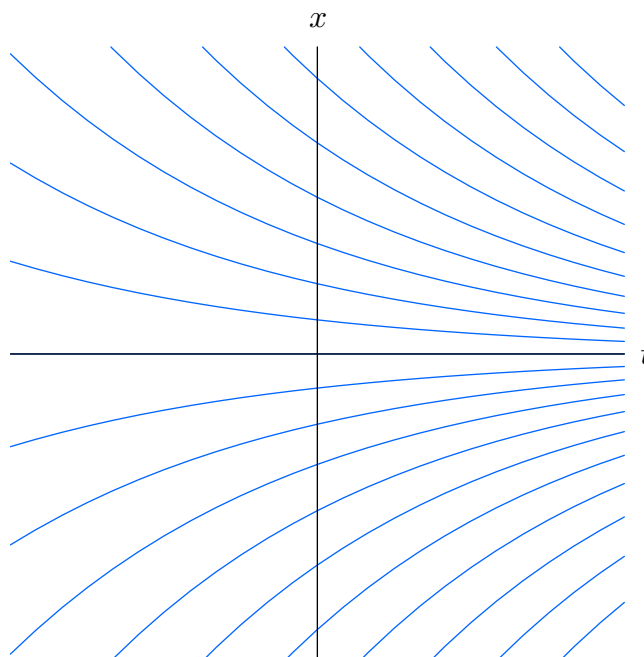


Figure 22.5. Characteristic Curves for $u_t - x u_x = 0$.

Example 22.3. Consider the equation

$$u_t - x u_x = 0. \quad (22.10)$$

In this case, the characteristic curves are the solutions to

$$\frac{dx}{dt} = -x, \quad \text{and so} \quad x e^t = k, \quad (22.11)$$

where k is the constant of integration; see Figure 22.5. It is easier to adopt $\xi = x e^t$ as the characteristic variable here, noting that its level sets are the characteristic curves. The

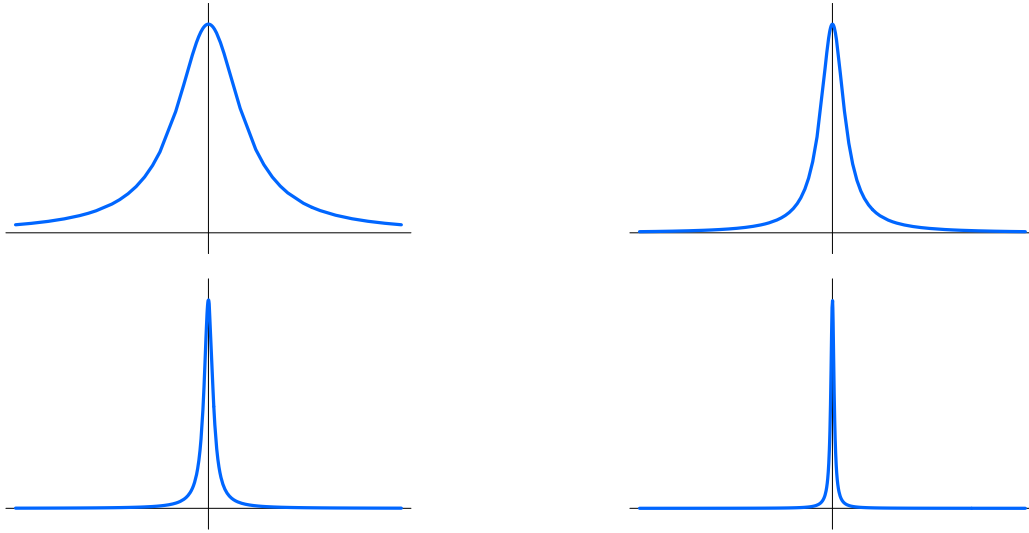


Figure 22.6. Solution to $u_t - x u_x = 0$.

solution therefore takes the form

$$u = p(xe^t), \quad (22.12)$$

where $p(\xi)$ is an arbitrary function of $\xi = xe^t$. Given the initial data

$$u(0, x) = f(x), \quad \text{the resulting solution is} \quad u = f(xe^t).$$

For example, the solution

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

corresponding to initial data $u(t, 0) = f(x) = (x^2 + 1)^{-1}$ is plotted in Figure 22.6 at times $t = 0, 1, 2, 3$. Note that since the characteristic curves all converge on the t axis, the solution becomes more and more concentrated at the origin. In the limit, it converges to the function that is zero everywhere except for the value $u(t, 0) \equiv 1$ at the origin. *Warning:* The limit is *not* a delta function, since its value at $x = 0$ remains bounded.

A Nonlinear Transport Equation

Perhaps the simplest possible nonlinear partial differential equations is the *nonlinear transport equation*

$$u_t + u u_x = 0. \quad (22.13)$$

first systematically studied by Poisson and Riemann in the early nineteenth century. Since it appears in so many applications, this equation appears in the literature under a variety of names, including the Riemann equation, the inviscid Burgers' equation, and the dispersionless Korteweg–deVries equation. It and its multi-dimensional and multi-component generalizations play a crucial role in the modeling of gas dynamics, traffic flow, flood waves in rivers, chromatography, chemical reactions, and other areas; see [188].

The first order partial differential equation (22.13) has the form of a transport equation, whose wave velocity $c = u$ depends, not on the position x , but rather on the size of the disturbance. Larger waves move faster, and overtake smaller waves. Waves of elevation, where $u > 0$, move to the right, while waves of depression, where $u < 0$, move to the left.

Fortunately, the method of characteristics that was developed for linear wave equations also works in the present nonlinear situation and leads to a complete solution. Mimicking our previous construction, (22.5), let us define a *characteristic curve* of the nonlinear wave equation (22.13) to be a solution to the ordinary differential equation

$$\frac{dx}{dt} = u(t, x). \quad (22.14)$$

As such, the characteristics depend upon the solution u , which, in turn, is based on the characteristic variable. So we appear to be trapped in a circular argument. The resolution of the apparent conundrum is to observe that, as in the linear case, the solution $u(t, x)$ remains constant along its characteristics, and this fact will allow us to simultaneously specify both.

To prove this claim, suppose that $x = x(t)$ parametrizes a characteristic curve. We need to show that the function $h(t) = u(t, x(t))$, which is obtained by evaluating the solution along the curve, is constant. As usual, constancy is proved by checking that its derivative is identically zero. Invoking the chain rule, and then (22.14), we deduce that

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

The final expression vanishes because u is assumed to solve the wave equation (22.13) at all values of (t, x) , including those on the curve $(t, x(t))$. This verifies our claim that $h(t)$ is constant, and so the solution u is constant on the characteristic curve. This has the implication that the right hand side of equation (22.14) is a constant whenever $x = x(t)$ defines a characteristic curve, and so the derivative dx/dt is a constant — namely the value of u on the curve. In this manner, we arrive at the key deduction that the characteristic curve must be a *straight line*

$$x = ut + k, \quad (22.15)$$

whose *characteristic slope* u equals the value assumed by the solution u on it. The larger u is, the steeper the characteristic line, and the faster that part of the wave travels.

The corresponding characteristic variable $\xi = x - tu$ depends upon the solution, which can now be written in implicit form

$$u = f(x - tu), \quad (22.16)$$

where $f(\xi)$ is an arbitrary function of the characteristic variable. The solution $u(t, x)$ can be found by solving the algebraic equation (22.16). For example, if

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

is an affine function, with α, β constant, then

$$u = \alpha(x - tu) + \beta, \quad \text{and hence} \quad u(t, x) = \frac{\alpha x + \beta}{1 + \alpha t} \quad (22.17)$$

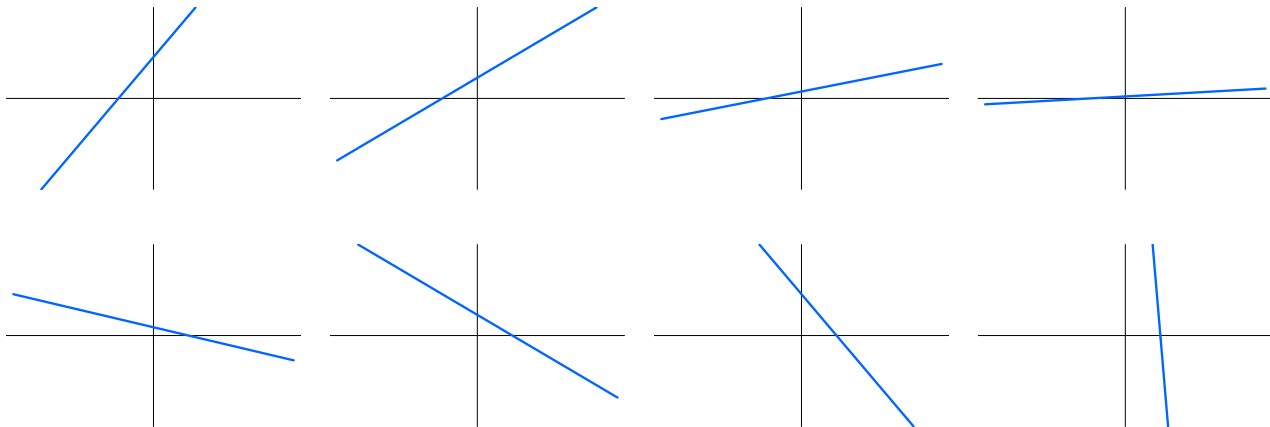


Figure 22.7. Two Solutions to $u_t + uu_x = 0$.

is the corresponding solution to the nonlinear transport equation. At each fixed t , the graph of the solution is a straight line. If $\alpha > 0$, the solution flattens out as $t \rightarrow \infty$. On the other hand, if $\alpha < 0$, the straight line rapidly steepens to vertical as t approaches the critical time $t_* = -1/\alpha$, at which point the solution ceases to exist — it is said to “blow up”. In Figure 22.7, we graph the solution with $\alpha = 1$, $\beta = .5$, when $t = 0, 1, 5, 20$ on the top row, and $\alpha = -.2$, $\beta = .1$, at times $t = 0, 3, 4, 4.9$ on the bottom row. In the second case, the solution becomes vertical as $t \rightarrow 5$ and then ceases to exist.

In general, to construct the solution $u(t, x)$ to the initial value problem

$$u(0, x) = f(x), \quad (22.18)$$

we note that, at $t = 0$, the implicit solution formula (22.16) reduces to $u(0, x) = f(x)$. Thus, the function f coincides with the initial data. However, because (22.16) is an implicit equation for $u(t, x)$, it is not immediately evident

- (a) whether it can be solved to give a well-defined value for $u(t, x)$, and,
- (b) even granted this, how to describe the solution’s qualitative features and dynamical behavior.

A more instructive and revealing strategy is based on the following geometrical construction, inspired by the linear version appearing in Figure 22.2. Through each point $(0, y)$ on the x axis, draw the characteristic line

$$x = t f(y) + y \quad (22.19)$$

whose slope, namely $f(y) = u(0, y)$, equals the value of the initial data at that point. According to the preceding argument, the solution will have the same value on the entire characteristic line (22.19), and so

$$u(t, t f(y) + y) = f(y). \quad (22.20)$$

For example, if $f(y) = y$, then $u(t, x) = y$ whenever $x = ty + y$; eliminating y , we recover $u(t, x) = x/(t + 1)$, which agrees with one of our straight line solutions (22.17).

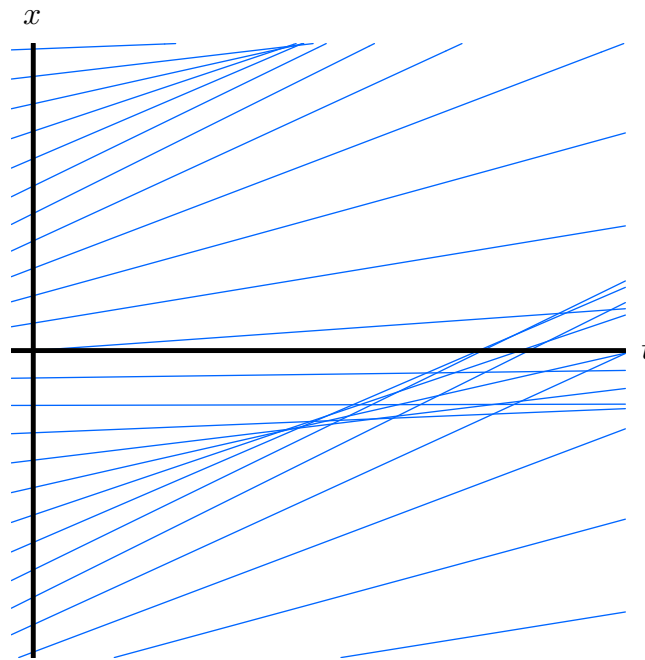


Figure 22.8. Characteristic Lines for $f(x) = \frac{1}{4} \sin(1.8x - .8)$.

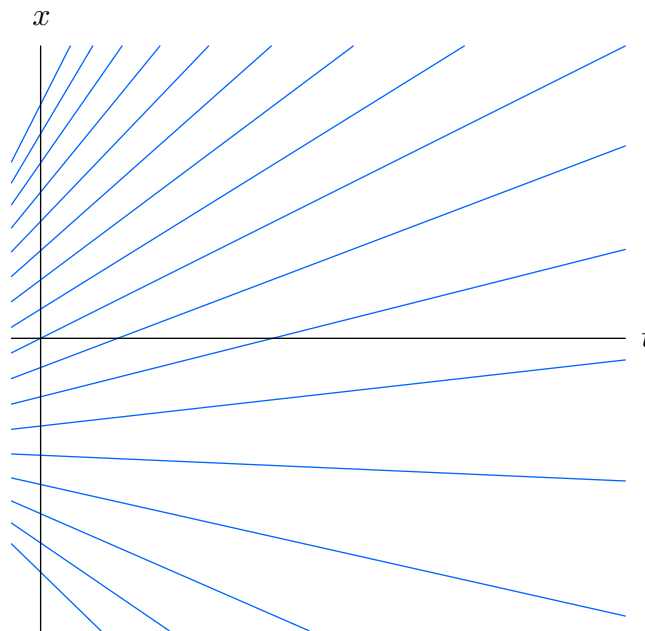


Figure 22.9. Characteristic Lines for a Rarefaction Wave.

Now, the trouble with our construction is immediately apparent from the illustrative Figure 22.8. Any two characteristic lines that are not parallel must cross each other somewhere. The value of the solution must equal to the slope of the characteristic line, and so, at the crossing point, the solution is required to assume two *different* values, one corresponding to each line. Something is clearly amiss, and we need to study the resulting solutions in more depth.

It turns out that there are three basic scenarios. The first, trivial case is when all the

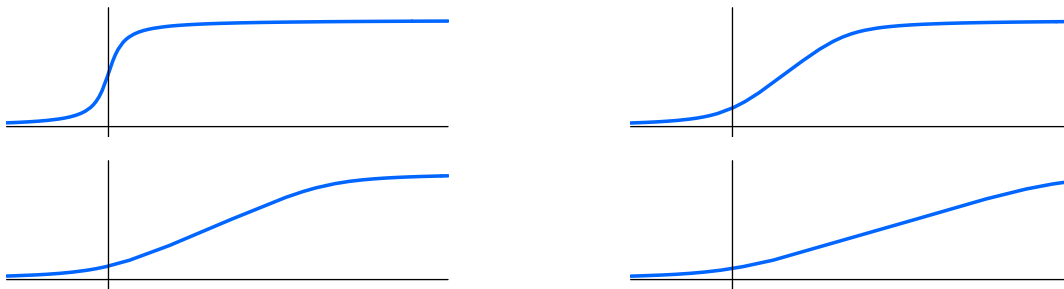


Figure 22.10. Rarefaction Wave.

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

The next simplest case occurs when the initial data $f(x)$ is everywhere increasing, so $f(x) \leq f(y)$ whenever $x \leq y$, which is assured if its derivative is never negative: $f'(x) \geq 0$. In this case, as in sketched in Figure 22.9, the characteristic lines emanating from the x axis fan out into the right half plane, and so never cross each other when $t \geq 0$. Each point (t, x) for $t \geq 0$ lies on a unique characteristic line, and the value of the solution at (t, x) is equal to the slope of the line. Consequently, the solution is well-defined at all future times. Physically, such solutions represent *rarefaction waves*, which gradually spread out as time progresses. A typical example, corresponding to initial data

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

is plotted in Figure 22.10 at successive times $t = 0, 1, 2, 3$. Note how the slope of the solution gradually diminishes as the rarefaction wave spreads out.

The more interesting case is when $f'(x) < 0$. Now some of the characteristic lines starting at $t = 0$ will cross at some point in the future. If (t, x) lies on two or more distinct characteristic lines, the value of the solution $u(t, x)$, which should equal the characteristic slope, is no longer uniquely determined. Although one might be tempted to deal with such multiply-valued solutions in a purely mathematical framework, from a physical standpoint this is unacceptable. The solution $u(t, x)$ is supposed to represent a physical quantity, e.g., density, velocity, pressure, etc., and must therefore assume a unique value at each point. The mathematical model has broken down, and fails to agree with the physical reality.

Before confronting this difficulty, let us first, from a theoretical standpoint, try to understand what happens if we were to continue the solution as a multiply-valued function. To be specific, consider the initial data

$$u(0, x) = \frac{\pi}{6} - \frac{1}{3} \tan^{-1} x, \quad (22.21)$$

appearing in the first plot in Figure 22.12. The corresponding characteristic lines are sketched in Figure 22.11. Initially, they do not cross, and the solution remains a well-defined, single-valued function. However, eventually one reaches a critical time, $t = t_\star > 0$, when the first two characteristic lines cross each other. Subsequently, a wedge-shaped region appears in the (t, x) plane, consisting of points which lie on the intersection of three

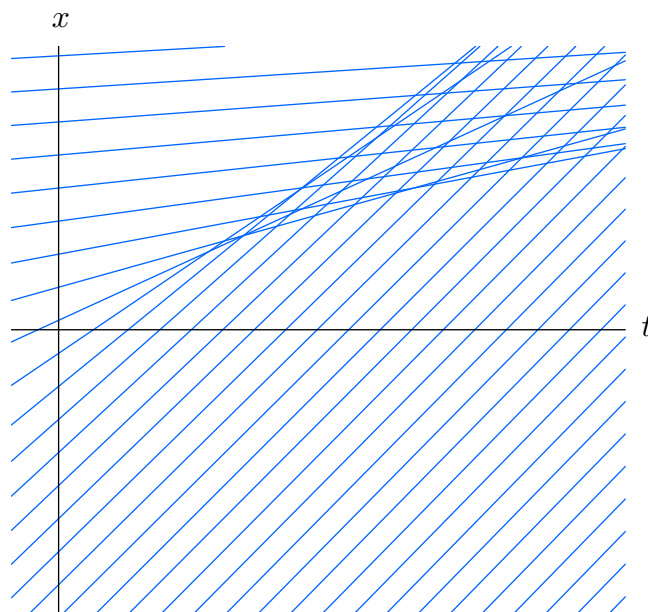


Figure 22.11. Characteristics for a Shock Wave.

different characteristic lines with different slopes; at such points, the solution achieves three distinct values. Outside the wedge, the points only belong to a single characteristic line, and the solution remains single-valued. (The boundary of the wedge consists of points where only two characteristic lines cross.)

To fully appreciate what is going on, look now at the sequence of pictures of the multiply-valued solution at successive times in Figure 22.12. Since the initial data is positive, $f(x) > 0$, all the characteristic slopes are positive. As a consequence, all the points on the solution curve will move to the right, at a speed equal to their height. Since the initial data is a decreasing function, points lying to the left will move faster than those on the right, and eventually overtake them. Thus, as time passes, the solution steepens. At the critical time t_* when the first two characteristic lines cross, say at x_* , the tangent to the solution curve has become vertical:

$$\frac{\partial u}{\partial x}(t, x_*) \longrightarrow \infty \quad \text{as} \quad t \longrightarrow t_*.$$

Afterwards, the solution graph no longer represents a single-valued function; its overlapping lobes lie over points (t, x) in the aforementioned wedge.

The critical time t_* can be determined from the implicit solution formula (22.16). Indeed, if we differentiate with respect to x , we find

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

is the characteristic variable, which is constant along the characteristic lines. Solving,

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



Figure 22.12. Multiply-Valued Solution.

Therefore, the slope blows up,

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

In other words, if the initial data has negative slope at position x , so $f'(x) < 0$, then the solution along the characteristic line emanating from the point $(0, x)$ will break down at the time $-1/f'(x)$. As a consequence, the earliest critical time is

$$t_{\star} = \min \left\{ -\frac{1}{f'(x)} \mid f'(x) < 0 \right\}. \quad (22.22)$$

For instance, for the particular initial configuration (22.21) represented by the pictures,

$$f'(x) = -\frac{1}{3(1+x^2)}, \quad \text{and so the critical time is} \quad t_{\star} = \min(3(1+x^2)) = 3.$$

Now, while mathematically plausible, such a multiply-valued solution is physically untenable. So what happens after the critical time t_{\star} ? One needs to choose which of the possible solution values at each point (t, x) contained in the wedge is physically appropriate. Indeed, the mathematics by itself is incapable of specifying how to continue the solution past the critical time at which the characteristics begin to cross. We therefore must return to the underlying physics, and ask what sort of phenomenon are we trying to model. The most instructive is to view the differential equation as a simple model of compressible fluid flow in a single space variable, e.g., motion of gas in a long pipe. If we push a piston down the end of a long pipe then the gas will move ahead of the piston and thereby be compressed. However, if the piston moves too rapidly, the gas piles up on top of itself, and a shock wave forms and propagates down the pipe. Mathematically, the shock is represented by a discontinuity where the solution abruptly changes value.

One way to resolve our mathematical dilemma relies on the fact that the partial differential equation takes the form of a conservation law, in accordance with the following definition[†].

Definition 22.4. A *conservation law* is an equation of the form

$$\frac{\partial T}{\partial t} + \frac{\partial X}{\partial x} = 0. \quad (22.23)$$

The functions T and X are known, respectively, as the *conserved density* and associated *flux*.

In the simplest situations, the conserved density $T(t, x, u)$ and flux $X(t, x, u)$ depend on the time t , the position x , and the solution $u(t, x)$ to the physical system. (Higher order conservation laws, which also depend upon derivatives of u , will appear in the final section.) We can clearly rewrite the nonlinear transport equation (22.13) in the following conservation law form:

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} \left(\frac{1}{2} u^2 \right) = 0, \quad (22.24)$$

where the conserved density and flux are, respectively,

$$T = u, \quad X = \frac{1}{2} u^2.$$

The reason for calling (22.23) a conservation law comes from the following observation.

Proposition 22.5. Given a conservation law (22.23),

$$\frac{d}{dt} \int_a^b T \, dx = - \left. X \right|_{x=a}^b. \quad (22.25)$$

The proof of (22.25) is immediate — assuming sufficient smoothness that allows one to bring the derivative inside the integral sign, and then invoking the Fundamental Theorem of Calculus:

$$\frac{d}{dt} \int_a^b T \, dx = \int_a^b \frac{\partial T}{\partial t} \, dx = - \int_a^b \frac{\partial X}{\partial x} \, dx = - \left. X \right|_{x=a}^b.$$

Formula (22.25) says that the rate of change of the integrated density over an interval depends only on the flux through its endpoints. In particular, if there is no net flux into or out of the interval, then the integrated density is *conserved*, meaning that it remains constant over time. All physical conservation laws — mass, momentum, energy, and so on — for systems governed by partial differential equations are of this form. (For ordinary differential equations, conservation laws coincide with first integrals, as discussed in Section 20.3.)

[†] Here we describe the one-dimensional situation. See Exercise ■ for conservation laws for n -dimensional dynamics.

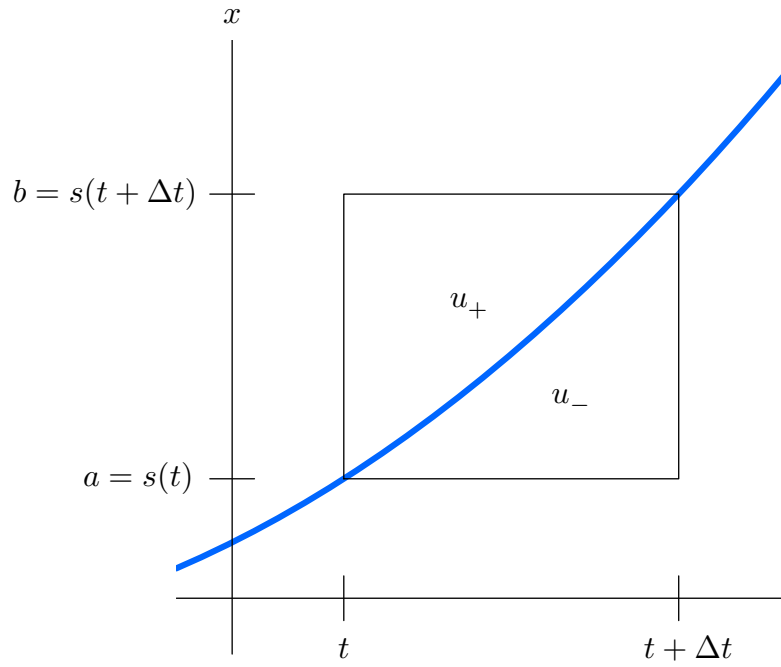


Figure 22.13. Conservation of Mass Near a Shock.

For the transport equation (22.24), the integrated conservation law (22.25) takes the specific form

$$\frac{d}{dt} \int_a^b u(t, x) dx = \frac{1}{2} [u(t, a)^2 - u(t, b)^2]. \quad (22.26)$$

Viewing the equation as a model for, say, compressible fluid flow in a pipe, the integral on the left hand side represents the total mass of the fluid contained in the interval $[a, b]$. The right hand side represents the mass flux *into* the interval through its two endpoints, and thus the conservation equation (22.26) is the mathematical formalization of basic mass conservation — mass is neither created nor destroyed, but can only enter a region as a flux through its boundary. In particular, if there is zero mass flux, then we deduce conservation of the total mass.

With this in hand, let us return to the physical context of the nonlinear transport equation. We will assume that mass conservation continues to hold even within a shock, which, from a purely molecular standpoint, makes eminent physical sense. By definition, a *shock* is a jump discontinuity in the solution $u(t, x)$. Suppose that, at time t , a shock occurs at position $x = s(t)$. We require[†] that both the left and right hand limits

$$u_-(t) = u(t, s(t)^-) = \lim_{x \rightarrow s(t)^-} u(t, x), \quad u_+(t) = u(t, s(t)^+) = \lim_{x \rightarrow s(t)^+} u(t, x),$$

of the solution on either side of the shock discontinuity are well defined. Let us further assume that, in time, the shock $x = s(t)$ follows a smooth — meaning C^1 — path. Now, referring to Figure 22.13, Consider a small time interval, from t to $t + \Delta t$. During this

[†] With more analytical work, [188], the listed assumptions can all be rigorously justified.

time, the shock moves from position $a = s(t)$ to position $b = s(t + \Delta t)$. The total mass contained in the interval $[a, b]$ at time t , before the shock has passed through, is

$$m(t) = \int_a^b u(t, x) dx \approx \bar{u}_+(t) (b - a) = \bar{u}_+(t) [s(t + \Delta t) - s(t)],$$

where $\bar{u}_+(t)$ is the average value of $u(t, x)$ over the interval. After the shock has passed, the total mass has become

$$m(t + \Delta t) = \int_a^b u(t + \Delta t, x) dx \approx \bar{u}_-(t) (b - a) = \bar{u}_-(t) [s(t + \Delta t) - s(t)],$$

where $\bar{u}_-(t)$ refers to the average value of $u(t + \Delta t, x)$ over the same interval. In the limit as $\Delta t \rightarrow 0$, the point $b = s(t + \Delta t) \rightarrow s(t) = a$, and hence the averages

$$\lim_{\Delta t \rightarrow 0} \bar{u}_+(t) = u_+(t), \quad \lim_{\Delta t \rightarrow 0} \bar{u}_-(t) = u_-(t),$$

tend to the limiting solution values on the right and left hand sides of the shock discontinuity. Thus, the limiting rate of change in mass across the shock at time t is

$$\begin{aligned} \frac{dm}{dt} &= \lim_{\Delta t \rightarrow 0} \frac{m(t + \Delta t) - m(t)}{\Delta t} \\ &= \lim_{\Delta t \rightarrow 0} [\bar{u}_-(t) - \bar{u}_+(t)] \frac{s(t + \Delta t) - s(t)}{\Delta t} = [u_-(t) - u_+(t)] \frac{ds}{dt}, \end{aligned}$$

which is the product of the shock speed times minus the jump magnitude at the shock discontinuity. On the other hand, at any $t < \tau < t + \Delta t$, the mass flux into the interval $[a, b]$ is, according to the right hand side of (22.26),

$$\frac{1}{2} [u(\tau, a)^2 - u(\tau, b)^2] \rightarrow \frac{1}{2} [u_-(t)^2 - u_+(t)^2] \quad \text{as} \quad \Delta t \rightarrow 0.$$

For conservation of mass to hold across the shock, the limiting value of the rate of change in mass must equal the limiting mass flux,

$$[u_-(t) - u_+(t)] \frac{ds}{dt} = \frac{1}{2} [u_-(t)^2 - u_+(t)^2],$$

from which we discover the *Rankine-Hugoniot condition*

$$\frac{ds}{dt} = \frac{1}{2} \frac{u_-(t)^2 - u_+(t)^2}{u_-(t) - u_+(t)} = \frac{u_-(t) + u_+(t)}{2}. \quad (22.27)$$

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

A shock appears when one or more characteristic lines cross. For this to occur, characteristics to the left of the shock must have larger slope (or speed), while those to the right must have smaller slope. Since the shock speed is the average of the two characteristic slopes, this means

$$u_-(t) > \frac{ds}{dt} = \frac{u_-(t) + u_+(t)}{2} > u_+(t). \quad (22.28)$$

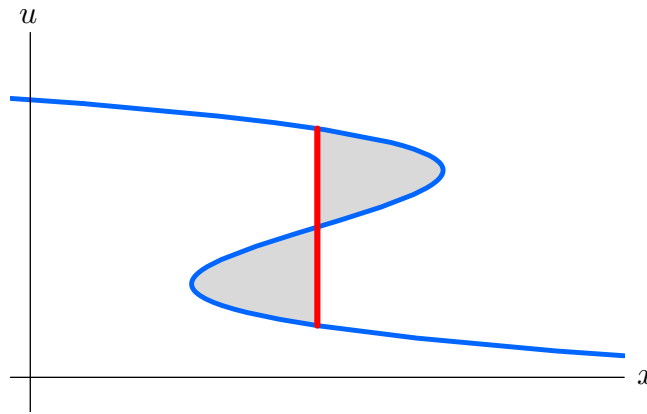


Figure 22.14. Equal Area Rule.

While it is theoretically possible to construct a shock solution to (22.13) that maintains the Rankine–Hugoniot constraint (22.27) but violates (22.28), such solutions are excluded on physical grounds, in that they violate causality, [105], which requires that characteristics are only allowed to enter shocks, not leave, and, furthermore, are not stable under small perturbations, [188]. The dynamics of shock wave solutions is then prescribed by the Rankine–Hugoniot and causality conditions (22.27, 28).

How does one determine the motion of the shock in practice? The answer is beautifully simple. Since the total mass, which at time t is the area under the curve $u(t, x)$, must be conserved, one merely draws the vertical shock line where the areas of the two lobes in the multiply-valued solution are equal, as in Figure 22.14. This *Equal Area Rule* ensures that the total mass of the shock solution matches that of the original (why?), as required by the physical conservation law.

Example 22.6. An illuminating special case is when the initial data has the form of a step function with a single jump discontinuity at the origin:

$$u(0, x) = f(x) = a + b \sigma(x) = \begin{cases} a, & x < 0, \\ b, & x > 0. \end{cases} \quad (22.29)$$

If[†] $a > b > 0$, then the initial data is already in the form of a shock wave. For $t > 0$, the mathematical solution constructed by continuing along the characteristic lines is multiply-valued in the region $bt < x < at$, where it assumes both values a and b ; see Figure 22.15. The Equal Area Rule tells us to draw the shock line halfway along, at $x = \frac{1}{2}(a + b)t$, in order that the two triangles have the same area. Therefore, the shock moves with speed $c = \frac{1}{2}(a + b)$ equal to the average of the two speeds at the jump, and so this particular shock wave solution is

$$u(t, x) = a + b \sigma(x - ct) = \begin{cases} a, & x < ct, \\ b, & x > ct, \end{cases} \quad \text{where} \quad a > c = \frac{a + b}{2} > b. \quad (22.30)$$

[†] Cases where a or b are negative are left to the exercises.

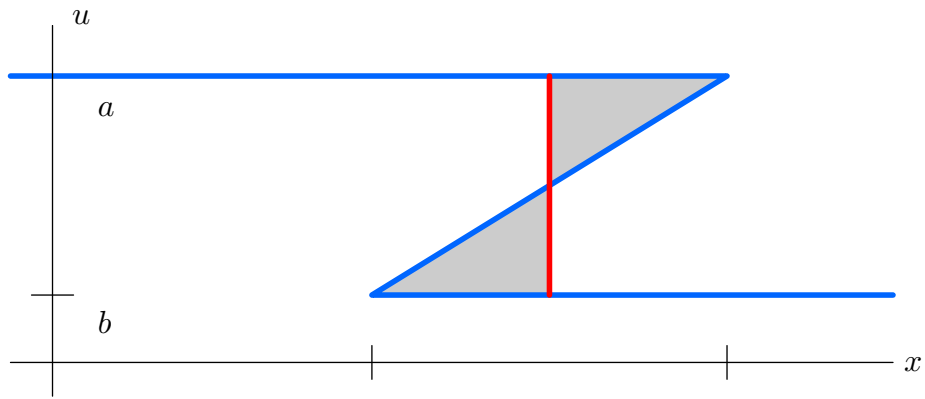


Figure 22.15. Multiply-Valued Step Wave.

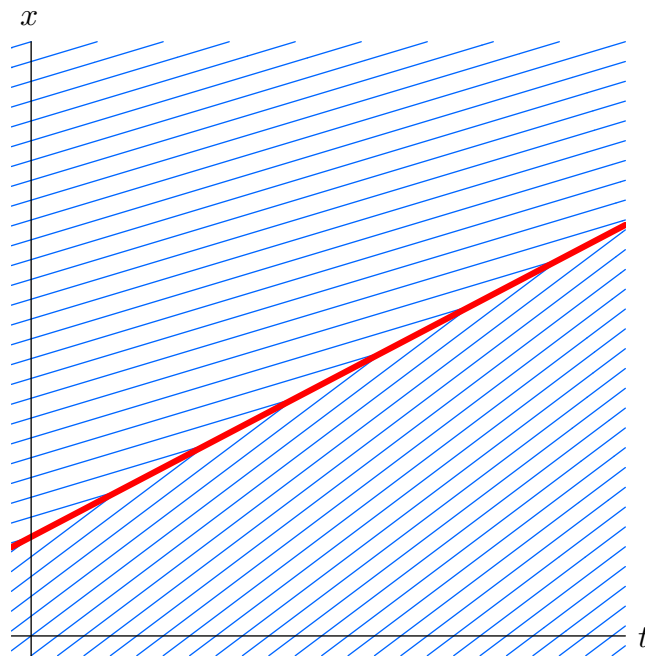


Figure 22.16. Characteristic Lines for the Step Wave Shock.

A graph of the characteristic lines appears in Figure 22.16.

By way of contrast, suppose $0 < a < b$, so the initial data has a jump upwards. In this case, the characteristic lines diverge from the initial discontinuity, and the mathematical solution is not specified at all in the wedge-shaped region $at < x < bt$. Now our task is to decide how to connect the two regions where the solution is well-defined. The simplest connection is an affine function, i.e., a straight line. Indeed, a simple modification of the rational solution (22.17) produces the function

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

which not only solves the differential equation, but also has the required values $u(t, at) = a$,

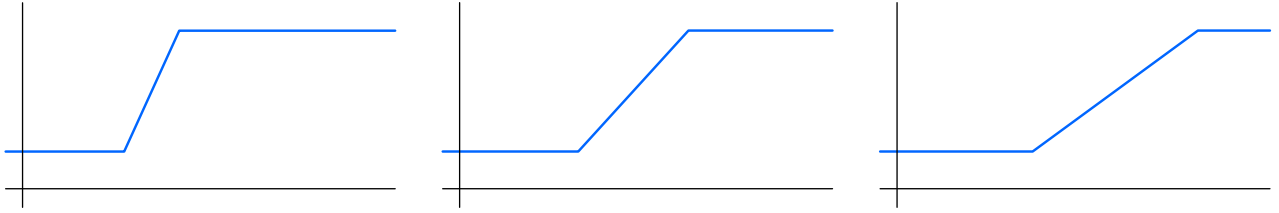


Figure 22.17. Piecewise Affine Rarefaction Wave.

and $u(t, bt) = b$ at the two edges of the wedge. The resulting solution is the piecewise affine *rarefaction wave*

$$u(t, x) = \begin{cases} a, & x \leq at, \\ x/t, & at \leq x \leq bt, \\ b, & x \geq bt, \end{cases} \quad (22.31)$$

which is graphed in Figure 22.17. In fact, it can be shown, [105], that this is the only solution that preserves the causality condition (22.28).

These prototypical solutions epitomize the basic phenomena modeled by the nonlinear transport equation: *rarefaction waves*, that emanate from regions where the initial data satisfies $f'(x) > 0$, where the solution spreads out as time progresses, and *compression waves*, emanating from regions where $f'(x) < 0$, that progressively steepen and eventually break into a shock discontinuity. Anyone caught in a traffic jam recognizes the compression waves, where the vehicles are bunched together and almost stationary, while the interspersed rarefaction waves correspond to freely moving traffic. (An intelligent driver will take advantage of the rarefaction waves moving through the jam to switch lanes!) The familiar, frustrating traffic jam phenomenon, even on accident- or construction-free stretches of highway, is an intrinsic effect of the nonlinear transport model that governs the traffic flow, [188].

Our derivation of the Rankine–Hugoniot condition (22.27) prescribing the shock speed relies on the fact that we can write the original partial differential equation in the form of a conservation law. But there are other ways to do this; for instance, multiplying the nonlinear transport equation (22.13) by u allows us write it in the alternative conservative form

$$u \frac{\partial u}{\partial t} + u^2 \frac{\partial u}{\partial x} = \frac{\partial}{\partial t} \left(\frac{1}{2} u^2 \right) + \frac{\partial}{\partial x} \left(\frac{1}{3} u^3 \right) = 0. \quad (22.32)$$

Here, the conserved density is $T = \frac{1}{2} u^2$, and the associated flux $X = \frac{1}{3} u^3$. The integral form equation (22.25) of the conservation law is

$$\frac{d}{dt} \int_a^b \frac{1}{2} u(t, x)^2 dx = \frac{1}{3} [u(t, a)^3 - u(t, b)^3]. \quad (22.33)$$

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

our previous argument, we are led to the alternative condition

$$\frac{ds}{dt} = \frac{\frac{1}{3}(u_-(t)^3 - u_+(t)^3)}{\frac{1}{2}(u_-(t)^2 - u_+(t)^2)} = \frac{2}{3} \frac{u_-(t)^2 + u_-(t)u_+(t) + u_+(t)^2}{u_-(t) + u_+(t)} \quad (22.34)$$

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

The mathematical property that characterizes the shock dynamics is known as an *entropy condition*. Entropy conditions, such as the Rankine–Hugoniot Equal Area Rule (22.27), or the alternative (22.34), allow us to follow the solution beyond the formation of a simple shock. Once a shock forms, it cannot suddenly disappear — the discontinuity remains as the solution propagates. One consequence is the irreversibility of the solutions to the nonlinear transport equation. One cannot simply run time backwards and expect shocks to spontaneously vanish. However, this irreversibility is of a different character than that of the ill-posedness in the backwards heat equation. The nonlinear transport equation can be solved for $t < 0$, but this would result, typically, in the formation of a different collection of shocks, and would not be just the time reversal of the solution.

Continuing past the initial shock formation, as other characteristic lines start to cross, additional shocks appear. The shocks themselves continue propagate, often at different velocities. When a fast moving shock catches up with a slow moving shock, one must then decide how to merge the shocks together to retain a physically consistent solution. The selected entropy condition continues to resolve the ambiguities. However, at this point, the mathematical details have become too complicated for us to pursue in any more detail, and we refer the interested reader to Whitham’s book, [188], which includes a wide range of applications to equations of gas dynamics, flood waves in rivers, motion of glaciers, chromatography, traffic flow, and many other physical systems.

22.2. Nonlinear Diffusion.

First order partial differential equations, beginning with elementary scalar transport equations, and progressing on to the equations of gas dynamics, the full-blown Euler equations of fluid mechanics, and yet more complicated systems for plasmas and other complicated physical processes, are used to model conservative wave motion. Such systems fail to account for frictional and/or viscous effects, which are typically modeled by a parabolic diffusion equation such as the heat equation. 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 viscous term helps smooth out abrupt shock discontinuities, and the result is a well-determined and smooth dynamical process. Moreover, in the inviscid limit, as the diffusion term becomes vanishingly small, the smooth viscous solutions converge non-uniformly to the appropriate discontinuous shock wave, leading to an alternative mechanism for analyzing conservative nonlinear dynamical processes.

Burgers' Equation

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

$$u_t + uu_x = \gamma u_{xx}, \quad (22.35)$$

and is obtained by appending a linear diffusion term to the nonlinear transport equation (22.13). In fluids and gases, 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 mechanics, [188]. As with the heat equation, the diffusion coefficient $\gamma > 0$ must be positive in order that initial value problem be well-posed in forwards time.

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

$$u(0, x) = f(x), \quad -\infty < x < \infty. \quad (22.36)$$

(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 22.1, 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. Indeed, it can be proved that, under rather mild assumptions on the initial data, the solution to the initial value problem (22.35, 36) remains smooth and well-defined for all subsequent times, [188].

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

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

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 (22.35), we conclude that $v(\xi)$ must satisfy the nonlinear second order ordinary differential equation

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

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

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

[†] The equation is named after the applied mathematician J.M. Burgers, [36], and so the apostrophe goes after the “s”. Burgers' equation was apparently first studied as a physical model by Bateman, [14], although its solution already appears as an exercise in a nineteenth century ordinary differential equations text, [72; vol. 6, p. 102].

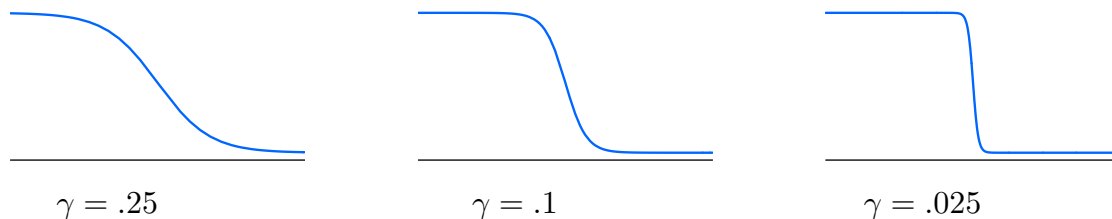


Figure 22.18. Traveling Wave Solutions to Burgers' Equation.

where k is a constant of integration. As in Section 20.1, the non-constant solutions to such an autonomous first order ordinary differential equation tend to either $\pm\infty$ or to one of the equilibrium points, i.e., the roots of the right hand side, as $t \rightarrow \pm\infty$. Thus, 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

$$2\gamma \frac{dv}{d\xi} = (v-a)(v-b), \quad \text{where} \quad c = \frac{1}{2}(a+b). \quad (22.37)$$

To obtain the bounded solutions, we concentrate on the case when $a < v < b$. Integrating (22.37) by the usual method, we find

$$\int \frac{2\gamma dv}{(v-a)(v-b)} = \frac{2\gamma}{b-a} \log \left(\frac{b-v}{v-a} \right) = \xi - \delta,$$

for δ a constant of integration, and hence

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

Thus, the bounded traveling wave solutions all have the explicit form

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

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 going from b to a . The wave travels to the right, unchanged in form, with speed equal to the average of its asymptotic values. In Figure 22.18 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 (22.30) to the nonlinear transport equation, which, as a result, is often referred to as the *inviscid Burgers' equation*.

Indeed, the profound fact is that, in the *inviscid limit* as the diffusion becomes vanishingly small, $\gamma \rightarrow 0$, the solutions to Burgers' equation (22.35) converge to the shock

wave solution to (22.13) constructed by the Equal Area Rule. This observation is in accordance with our physical intuition, that all physical systems retain a very small dissipative component, that serves to smooth out discontinuities that might appear in a theoretical model that fails to take the dissipation/viscosity/damping/etc. into account. In modern theory, this so-called *viscosity solution method* has been successfully used to characterize the discontinuous solutions to a broad range inviscid nonlinear wave equations is as the limit, as the viscosity goes to zero, of classical solutions to a diffusive version. Thus, the viscosity solutions to the nonlinear transport equation resulting from Burgers' equation are consistent with the Equal Area Rule for drawing the shock discontinuities. More generally, this method allows one to monitor the solutions as they evolve into regimes where multiple shocks merge and interact. We refer the interested reader to [119, 188].

The Hopf–Cole Transformation

By a remarkable stroke of luck, the nonlinear Burgers' equation can be converted into the linear heat equation and thereby explicitly solved. The *linearization* of Burgers' equation first appeared in an obscure exercise in a nineteenth century differential equations textbook, [72; vol. 6, p. 102]. Its modern rediscovery by Eberhard Hopf, [104], and Julian Cole, [43], was a milestone in the modern era of nonlinear partial differential equations, and is named the Hopf–Cole transformation in their honor.

Finding a way to covert a 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 through the inverse change of variables, is rarely of any independent interest. Sometimes there is a lucky accident, and such “accidental” linearizations can have a profound impact on our understanding of more complicated nonlinear systems.

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

$$v_t = \gamma v_{xx}. \quad (22.38)$$

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

$$v(t, x) = e^{\alpha \varphi(t, x)}, \quad \text{so} \quad \varphi(t, x) = \frac{1}{\alpha} \log v(t, x), \quad (22.39)$$

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. Theorem 14.3.

To determine the differential equation satisfied by the function φ , we invoke the chain rule to differentiate (22.39):

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

Substituting the first and last formulae into the heat equation (22.38) and canceling a common exponential factor, we conclude that $\varphi(t, x)$ satisfies the nonlinear partial differential equation

$$\varphi_t = \gamma \varphi_{xx} + \gamma \alpha \varphi_x^2, \quad (22.40)$$

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

$$\varphi_{tx} = \gamma \varphi_{xxx} + 2\gamma \alpha \varphi_x \varphi_{xx}. \quad (22.41)$$

If we now set

$$\frac{\partial \varphi}{\partial x} = u, \quad (22.42)$$

so that φ has 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 (22.35) with $\alpha = -1/(2\gamma)$. In this manner, we have arrived at the famous *Hopf–Cole transformation*.

Theorem 22.7. *If $v(t, x) > 0$ is any positive solution to the linear heat equation $v_t = \gamma v_{xx}$, then*

$$u(t, x) = \frac{\partial}{\partial x} \left(-2\gamma \log v(t, x) \right) = -2\gamma \frac{v_x}{v} \quad (22.43)$$

solves Burgers' equation $u_t + u u_x = \gamma u_{xx}$.

Do all solutions to Burgers' equation arise in this way? In order to decide, we run the argument in reverse. First, choose a potential function $\tilde{\varphi}(t, x)$ that satisfies (22.42); 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 (22.41). Integrating both sides of the latter equation with respect to x , we conclude that

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

for some integration “constant” $h(t)$. Thus, unless $h(t) \equiv 0$, our potential function $\tilde{\varphi}$ doesn't satisfy the potential Burgers' equation (22.40), but that's because we chose the “wrong” potential. Indeed, if we define

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

then

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

and hence the modified potential $\varphi(t, x)$ is a solution to the potential Burgers' equation (22.40). From this it easily follows that

$$v(t, x) = e^{-\varphi(t, x)/(2\gamma)} \quad (22.44)$$

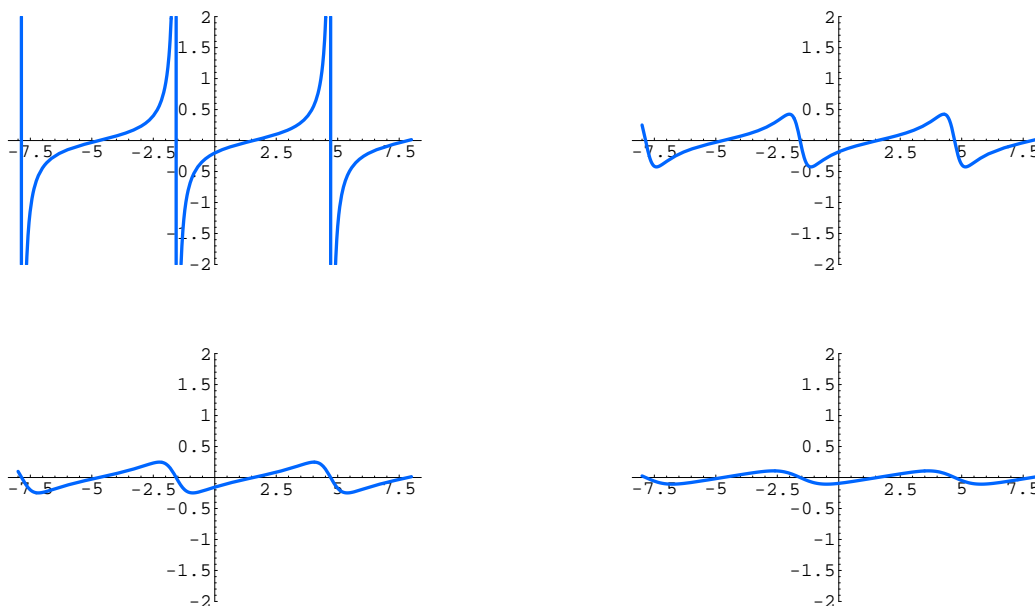


Figure 22.19. A Solution to Burgers' Equation.

is a positive solution to the heat equation, from which $u(t, x)$ can be recovered via the Hopf–Cole transformation (22.43). 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 22.8. 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 solution

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

to Burgers' equation; a typical example is plotted in Figure 22.19. 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 — see the first graph in Figure 22.19. This particular solution primarily feels the effects of the diffusivity, and rapidly goes to zero.

To solve the initial value problem (22.35–36) for Burgers' equation, we note that, under the Hopf–Cole transformation,

$$v(0, x) = h(x) = \exp \left(-\frac{\varphi(0, x)}{2\gamma} \right) = \exp \left(-\frac{1}{2\gamma} \int_0^x f(y) dy \right), \quad (22.45)$$

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 (22.43). The only effect is to multiply $v(t, x)$ by an overall constant, which does not change $u(t, x)$.

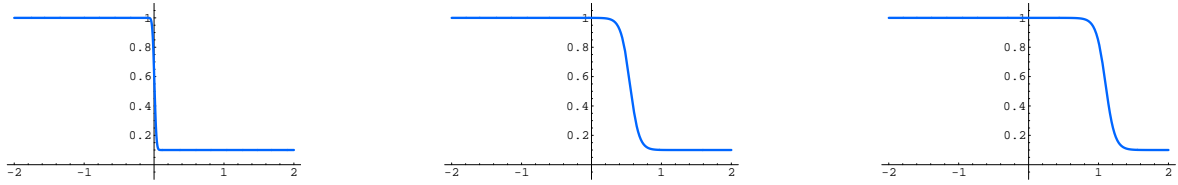


Figure 22.20. Shock Wave Solution to Burgers' Equation.

According to formula (14.61) (adapted to general diffusivity, as in Exercise ■), the solution to the initial value problem (22.38, 45) for the heat equation can be expressed as a convolution integral with the fundamental solution:

$$v(t, x) = \frac{1}{2\sqrt{\pi\gamma t}} \int_{-\infty}^{\infty} e^{-(x-y)^2/(4\gamma t)} h(y) dy.$$

Therefore, the solution to the Burgers' initial value problem (22.35, 36) is

$$u(t, x) = \frac{\int_{-\infty}^{\infty} \frac{x-y}{t} e^{F(t,x,y)} dy}{\int_{-\infty}^{\infty} e^{F(t,x,y)} dy} \quad \text{where} \quad F(t, x, y) = -\frac{1}{2\gamma} \int_0^y f(z) dz - \frac{(x-y)^2}{4\gamma t}. \quad (22.46)$$

Example 22.9. To demonstrate the smoothing effect of the diffusion terms, let us see what happens to the initial data

$$u(0, x) = \begin{cases} a, & x < 0, \\ b, & x > 0, \end{cases} \quad (22.47)$$

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,

$$F(t, x, y) = -\frac{(x-y)^2}{4\gamma t} - \begin{cases} -\frac{ay}{2\gamma}, & y < 0, \\ -\frac{by}{2\gamma}, & y > 0. \end{cases}$$

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

$$u(t, x) = a + \frac{b-a}{1 + h(t, x) \exp \frac{b-a}{2\gamma} (x - ct)} \quad (22.48)$$

where

$$c = \frac{a+b}{2}, \quad h(t, x) = \frac{1 - \operatorname{erf} \left(\frac{x-bt}{\sqrt{4\gamma t}} \right)}{1 - \operatorname{erf} \left(\frac{x-at}{\sqrt{4\gamma t}} \right)}, \quad (22.49)$$

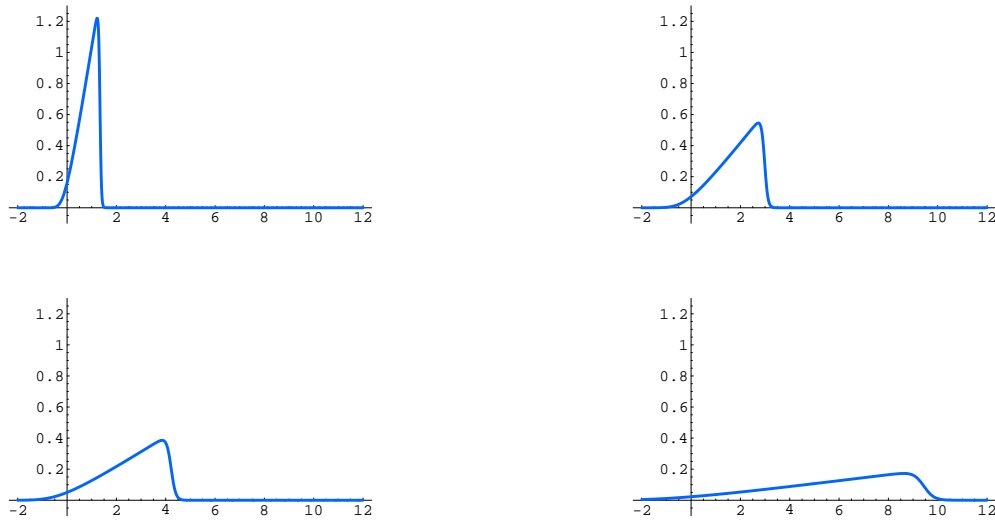


Figure 22.21. Triangular Wave Solution to Burgers' Equation.

and $\operatorname{erf} z$ denotes the error function (14.63). The solution, with $a = 1$, $b = .1$ and $\gamma = .03$ is plotted in Figure 22.20 at times $t = .01, 1.0, 2.0$. Note that the sharp transition region for the shock is immediately smoothed, and the solution rapidly settles into the form of a continuously varying transition layer between the two step heights. The larger the diffusion coefficient in relation to the initial solution heights a, b , the more significant the smoothing effect. Observe that, as $\gamma \rightarrow 0$, the function $h(t, x) \rightarrow 1$, and hence the solution converges to the shock wave solution (22.30) to the transport equation, in which the speed of the shock is the average of the two initial values.

Example 22.10. 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 (22.46), starting the integral for $F(t, x, y)$ at 0 is problematic, but as noted earlier, we are free to select any other starting point, e.g., $-\infty$. Thus, we take

$$F(t, x, y) = -\frac{1}{2\gamma} \int_{-\infty}^y \delta(z) dz - \frac{(x-y)^2}{4\gamma t} = \begin{cases} -\frac{(x-y)^2}{4\gamma t}, & y < 0, \\ -\frac{1}{2\gamma} - \frac{(x-y)^2}{4\gamma t}, & y > 0. \end{cases}$$

Substituting this into (22.46), we can evaluate the upper integral in elementary terms, while the lower integral involves the error function (14.63); after a little algebra, we find

$$u(t, x) = \sqrt{\frac{4\gamma}{\pi t}} \frac{e^{-x^2/(4\gamma t)}}{\coth\left(\frac{1}{4\gamma}\right) - \operatorname{erf}\left(\frac{x}{\sqrt{4\gamma t}}\right)}, \quad (22.50)$$

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 $a = 1$, at times $t = 1, 5, 10, 50$, appears in Figure 22.21. As you can see, the initial concentration diffuses out, but, unlike the heat equation, the wave does not remain symmetric owing to the advection terms in the equation. The effect is to steepen in front as it propagates. Eventually the triangular wave spreads out as the diffusion progresses.

22.3. Dispersion and Solitons.

Finally, we study a remarkable third order evolution equation that originally arose in the modeling of surface water waves, that serves to introduce yet further phenomena, both linear and nonlinear. The third order derivative models dispersion, in which waves of different frequencies move at different speeds. Coupled with the same nonlinearity as in the inviscid and viscous Burgers' (22.13, 35), the result is one of the most remarkable equations in all of mathematics, with far-reaching implications, not only in fluid mechanics and applications, but even in complex function theory, physics, etc., etc.

Linear Dispersion

So far, in our study of partial differential equations, we have not ventured beyond second order. Higher order equations do occur in applications, particularly in models for wave motion. The simplest linear partial differential equation of a type that we have not yet considered is the third order equation

$$u_t + u_{xxx} = 0 \tag{22.51}$$

It is the third in a hierarchy of simple evolution equations that starts with the simple ordinary differential equation $u_t = u$, then proceeds to the transport equation $u_t = u_x$, and then the heat equation $u_t = u_{xx}$ modeling basic diffusion processes. The third order case (22.51) is a simple model for linear dispersive waves.

To avoid additional complications caused by boundary conditions, we shall only look at the equation on the entire line, so $x \in \mathbb{R}$. The solution to the equation is uniquely specified by initial data

$$u(0, x) = f(x), \quad -\infty < x < \infty. \tag{22.52}$$

See [1] for a proof.

Let us apply the Fourier transform to solve the initial value problem. Let

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

be the spatial Fourier transform of the solution, which is assumed to remain in L^2 at all t , a fact that can be justified a posteriori. In view of the effect of the Fourier transform on derivatives — see Corollary 13.22 — the Fourier transform converts the partial differential equation (22.51) into a first order, linear ordinary differential equation

$$\widehat{u}_t - ik^3 \widehat{u} = 0, \tag{22.53}$$

parametrized by k , with initial conditions

$$\widehat{u}(0, k) = \widehat{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ikx} dx \quad (22.54)$$

given by the Fourier transform of (22.52). Solving the initial value problem (22.53–54) by the usual technique, we find

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

Inverting the Fourier transform yields the explicit formula for the solution

$$u(t, x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ik^3 t + ikx} \widehat{f}(k) dk \quad (22.55)$$

to the initial value problem for the dispersive wave equation (22.51–52).

Actually, to find the solutions to the differential equation, one does not need the full power of the Fourier transform. Note that (22.55) represents a linear superposition of elementary exponential functions. Let us substitute an exponential ansatz

$$u(t, x) = e^{i\omega t + ikx} \quad (22.56)$$

representing a complex oscillatory wave of *frequency* ω , which indicates the time vibrations, and *wave number* k , which indicates the corresponding oscillations in space. Since

$$\frac{\partial u}{\partial t} = i\omega e^{i\omega t + ikx}, \quad \frac{\partial^3 u}{\partial x^3} = -ik^3 e^{i\omega t + ikx},$$

(22.56) satisfies the partial differential equation (22.51) if and only if its frequency and wave number are related by

$$\omega = k^3. \quad (22.57)$$

The result is known as the *dispersion relation* for the partial differential equation. In general, any linear constant coefficient dynamical partial differential equation admits a dispersion relation of the form $\omega = \omega(k)$ which is straightforwardly found by substituting the exponential ansatz (22.56) and canceling the common exponential factors in the resulting equation. In our particular case, the exponential solution of wave number k has the form

$$u_k(t, x) = e^{ik^3 t + ikx}.$$

Linear superposition permits us to combine them in integral form, and so, for any (reasonable) function $a(k)$ depending on the wave number,

$$u(t, x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ik^3 t + ikx} a(k) dk$$

is easily seen to be a solution to the partial differential equation. The Fourier transform solution (22.55) has this form.

Example 22.11. The *fundamental solution* corresponds to a concentrated initial disturbance

$$u(0, x) = \delta(x).$$

since the Fourier transform of the delta function is just $\widehat{\delta}(k) = 1/\sqrt{2\pi}$, the resulting solution (22.55) is

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

since the solution is real (or, equivalently, the imaginary part of the integrand is odd) while the real part of the integrand is even. The second integral can be converted into that defining the Airy function,

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

as in (C.40), by the change of variables

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

and we conclude that the fundamental solution to the dispersive wave equation (22.51) can be written in terms of the Airy function:

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

See Figure ee3■ for a graph. Furthermore, writing the general initial data as a superposition of delta functions

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

we conclude that the solution has the form

$$u(t, x) = \frac{1}{\sqrt[3]{3t}} \int_{-\infty}^{\infty} f(\xi) \text{Ai}\left(\frac{x - \xi}{\sqrt[3]{3t}}\right) d\xi. \quad (22.58)$$

Although energy is conserved, unlike the heat and diffusion equations, the dispersion of waves means that the solution dies out.

Group velocity and wave velocity.

The Korteweg–de Vries Equation

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

$$u_t + u_{xxx} + uu_x = 0. \quad (22.59)$$

The equation was first derived by the French applied mathematician Boussinesq, [24; eq. (30)], [25; eqs. (283, 291)], in 1872 as a model for surface water waves. It was rediscovered by the Dutch mathematician Korteweg and his student de Vries, [120], over two

decades later, and, despite Boussinesq's priority, is named after them. In the early 1960's, the American mathematical physicists Martin Kruskal and Norman Zabusky, [195], derived it as a continuum limit of a model of nonlinear mass-spring chains studied by Fermi, Pasta and Ulam, [67]. Understanding the puzzling behavior of both systems coming from numerical experiments was the catalyst of one of the most remarkable and far-ranging discoveries of modern mathematics: integrable nonlinear partial differential equations.

The most important special solutions to the Korteweg-deVries equation are the *traveling waves*. We assume that the solution

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

is a wave of permanent form, 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 (22.59), we conclude that $v(\xi)$ must satisfy the nonlinear third order ordinary differential equation

$$v''' + vv' - cv' = 0. \quad (22.60)$$

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

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

To this end, we impose the boundary conditions

$$\lim_{\xi \rightarrow \pm\infty} v(\xi) = \lim_{\xi \rightarrow \pm\infty} v'(\xi) = \lim_{\xi \rightarrow \pm\infty} v''(\xi) = 0. \quad (22.61)$$

(See Exercise ■ for an analysis of the non-localized traveling wave solutions.)

The ordinary differential equation (22.60) can, in fact be solved in closed form. First, note that

$$\frac{d}{d\xi} \left[v'' + \frac{1}{2}v^2 - cv \right] = 0, \quad \text{and hence} \quad v'' + \frac{1}{2}v^2 - cv = k,$$

is a first integral, with k indicating the constant of integration. However, the localizing boundary conditions (22.61) imply that $k = 0$. Multiplying the latter equation by v' allows us to integrate a second time

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

Thus,

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

where ℓ is a second constant of integration, which, again by the boundary conditions (22.61), is also $\ell = 0$. We conclude that $v(\xi)$ satisfies the first order autonomous ordinary differential equation

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

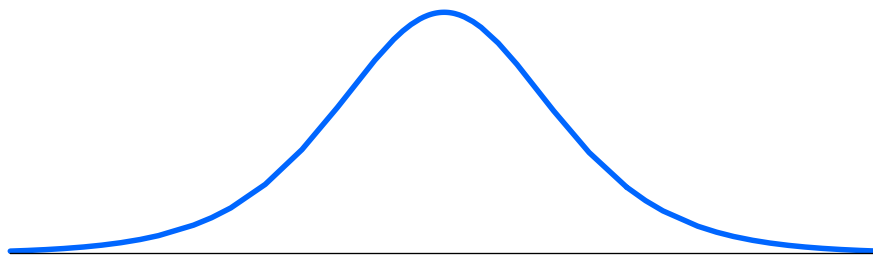


Figure 22.22. Solitary Wave.

We integrate by the usual method, cf. (20.7):

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

Using 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 22.22; it has a global maximum at $3c \operatorname{sech} 0 = 3c$ at $y = 0$, and is an even function, exponentially decay to 0 as $|\xi| \rightarrow \infty$. The resulting localized traveling wave solutions to the Korteweg–deVries equation are

$$u(t, x) = 3c \operatorname{sech}^2 \left[\frac{1}{2} \sqrt{c} (x - ct) + \delta \right], \quad (22.62)$$

where $c > 0$ and δ are arbitrary constants. The parameter c equals the speed of the wave. It is also equal to one third its amplitude, since the maximum value of $u(t, x)$ is $3c$ at the points $x = ct$, as well as the width, which is on the order of \sqrt{c} . The taller and wider the solitary wave, the faster it moves.

The solution (22.62) is known as a *solitary wave solution* since it represents a localized wave that travels unchanged in shape. Such waves were first observed by the British engineer J. Scott Russell, [165], who recounts how such a wave was generated by the sudden motion of a barge along an Edinburgh canal and then chasing it on horseback for several miles. Russell's observations were dismissed by his contemporary, the prominent mathematician George Airy, who claimed that such localized disturbances could not exist, basing his analysis upon a linearized theory. Much later, Boussinesq established the proper nonlinear surface wave model (22.59), valid for long waves in shallow water, and also derived the solitary wave solution (22.62), thereby fully exonerating Scott Russell's insight.

These nonlinear traveling wave solutions were discovered by Kruskal and Zabusky, [195], to have remarkable properties. For this reason they have been given a special new name — *soliton*. Ordinarily, combining two solutions to a nonlinear equation can be quite

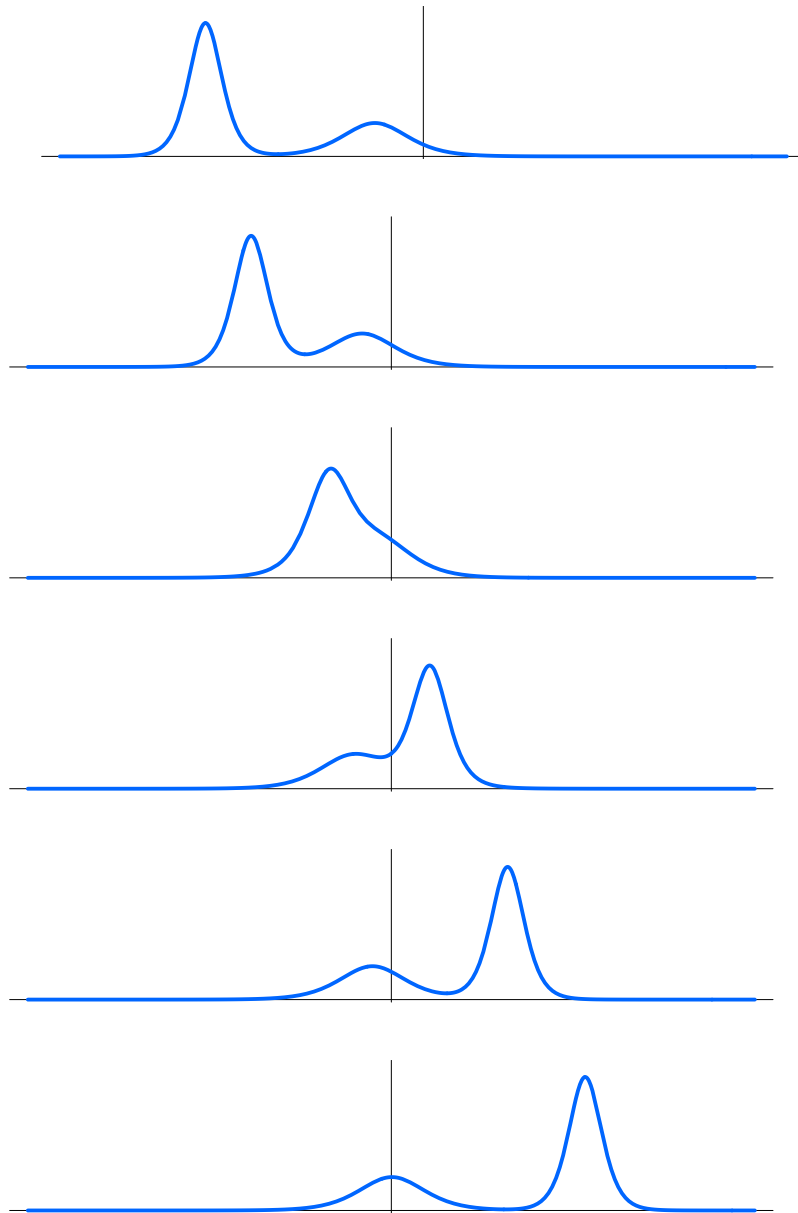


Figure 22.23. Interaction of Two Solitons.

unpredictable, and one might expect any number of scenarios to occur. If you start with initial conditions representing a taller wave to the left of a shorter wave, the solution of the Korteweg–deVries equation runs as follows. The taller wave moves faster, and so catches up the shorter wave. They then have a very complicated nonlinear interaction, as expected. But, remarkably, after a while they emerge from the interaction unscathed. The smaller wave is now in back and the larger one in front. After this, they proceed along their way, with the smaller one lagging behind the high speed tall wave. the only effect

of their encounter is a phase shift, meaning a change in the value of the phase parameter δ in each wave. See Figure 22.23. After the interaction, the position of the soliton if it had traveled unhindered by the other is shown in a dotted line. Thus, they behave like colliding particles, which is the genesis of the word “soliton”.

A similar phenomenon holds for several such soliton solutions. After some time where the various waves interact, they finally emerge with the largest soliton in front, and then in order to the smallest one in back, all progressing at their own speed, and so gradually drawing apart.

Remark: In the Korteweg–deVries equation model, one can find arbitrarily tall soliton solutions. In physical water waves, if the wave is too tall it will break. Indeed, it can be rigorously proved that the full water wave equations admit solitary wave solutions, but there is a wave of greatest height, beyond which a wave will tend to break. The solitary water waves are not genuine solitons, since there is a small, but measurable, effect when two waves collide.

Moreover, it can be proved that, starting with an *arbitrary* initial disturbance

$$u(0, x) = f(x)$$

that decays sufficiently rapidly as $|x| \rightarrow \infty$, after a sufficiently long time, the resulting solution $u(t, x)$ disintegrates into a finite number of solitons of different heights, moving off at their respective speeds to the right, and so are arranged in order from smallest to largest, plus a small dispersive tail moving to the left that rapidly disappears. Proving this remarkable result is beyond the scope of this book. It relies on the method of *inverse scattering*, that effectively linearizes the Korteweg–deVries equation with a linear eigenvalue problem of fundamental importance in one-dimensional quantum mechanics. The solitons correspond to the bound states of a quantum potential. We refer the interested reader to the introductory text [61] and the more advanced monograph [1] for details.

Like Burgers’ equation, the Korteweg–deVries equation can be linearized, but the linearization is considerably more subtle. It relies on the introduction of an auxiliary linear eigenvalue problem.

There is a remarkable transformation, known as the inverse scattering transform, which is a form of nonlinear Fourier transform, that can be used to solve the Korteweg–deVries equation. Its fascinating properties continue to be of great current research interest to this day.

22.4. Conclusion and Bon Voyage.

These are your first wee steps in a vast new realm. We are unable to discuss nonlinear partial differential equations arising in fluid mechanics, in elasticity, in relativity, in differential geometry, in computer vision, in mathematical biology. Chaos and integrability are the two great themes in modern nonlinear applied mathematics, and the student is well-advised to pursue both.

We bid you, dear reader, a fond adieu and wish you unparalleled success in your mathematical endeavors.