1
Introduction to Partial Differential Equations

Partial differential equations (PDEs) is one of the basic areas of applied analysis, and it is difficult to imagine any area of applications where its impact is not felt. In recent decades there has been tremendous emphasis on understanding and modeling nonlinear processes; such processes are often governed by nonlinear PDEs, and the subject has become one of the most active areas in applied mathematics and central in modern-day mathematical research. Part of the impetus for this surge has been the advent of high-speed, powerful computers, where computational advances have been a major driving force.

This initial chapter focuses on developing a perspective on understanding problems involving PDEs and how the subject interrelates with physical phenomena. It also provides a transition from an elementary course, emphasizing eigenfunction expansions and linear problems, to a more sophisticated way of thinking about problems that is suggestive of and consistent with the methods in nonlinear analysis.

Section 1.1 summarizes some of the basic terminology of elementary PDEs, including ideas of classification. In Section 1.2 we begin the study of the origins of PDEs in physical problems. This interdependence is developed from the basic, one-dimensional conservation law. In Section 1.3 we show how constitutive relations can be appended to the conservation law to obtain equations that model the fundamental processes of diffusion, advection or transport, and reaction. Some of the common equations, such as the diffusion equation, Burgers' equation, Fisher's equation, and the porous media equation, are obtained

An Introduction to Nonlinear Partial Differential Equations, Second Edition.
By J. David Logan

as models of these processes. In Section 1.4 we introduce initial and boundary value problems to see how auxiliary data specialize the problems. Finally, in Section 1.5 we discuss wave propagation in order to fix the notion of how evolution equations carry boundary and initial signals into the domain of interest. We also introduce some common techniques for determining solutions of a certain form (e.g., traveling wave solutions). The ideas presented in this chapter are intended to build an understanding of evolutionary processes so that the fundamental concepts of hyperbolic problems and characteristics, as well as diffusion problems, can be examined in later chapters with a firmer base.

1.1 Partial Differential Equations

1.1.1 Equations and Solutions

A *partial differential equation* is an equation involving an unknown function of several variables and its partial derivatives. To fix the notion, a *second-order PDE in two independent variables* is an equation of the form

$$G(x, t, u, u_x, u_t, u_{xx}, u_{tt}, u_{xt}) = 0, \quad (x, t) \in D, \tag{1.1.1}$$

where, as indicated, the independent variables x and t lie in some given domain D in $\mathbb{R}^2$. By a *solution* to (1.1.1) we mean a twice continuously differentiable function $u = u(x, t)$ defined on D that, when substituted into (1.1.1), reduces it to an identity on D. The function $u(x, t)$ is assumed to be twice continuously differentiable, so that it makes sense to calculate its first and second derivatives and substitute them into the equation; a smooth solution like this is called a *classical solution* or *genuine solution*. Later we extend the notion of solution to include functions that may have discontinuities, or discontinuities in their derivatives; such functions are called *weak solutions*. The xt domain D where the problem is defined is referred to as a *spacetime domain*, and PDEs that include time t as one of the independent variables are called *evolution* equations. When the two independent variables are both spatial variables, say, x and y rather than x and t, the PDE is an *equilibrium* or *steady-state* equation. Evolution equations govern time-dependent processes, and equilibrium equations often govern physical processes after the transients caused by initial or boundary conditions die away.

Graphically, a solution $u = u(x, t)$ of (1.1.1) is a smooth surface in three-dimensional xtu space lying over the domain D in the xt plane, as shown in Figure 1.1. An alternative representation is a plot in the xu-plane of the function $u = u(x, t_0)$ for some fixed time $t = t_0$ (see Figures 1.1 and 1.2). Such

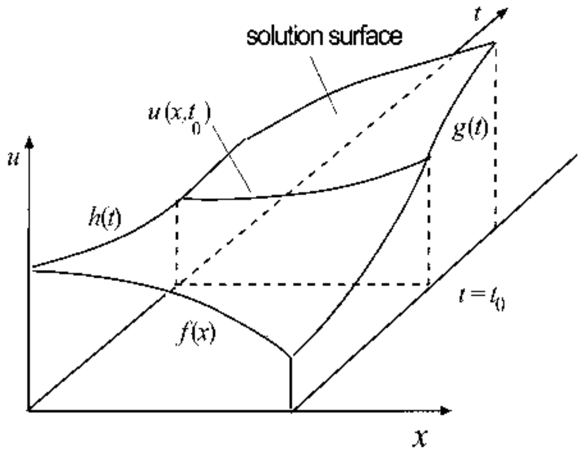

Figure 1.1 Solution surface $u = u(x,t)$ in xtu space, also showing a time snapshot or wave profile $u(x,t_0)$ at time t_0. The functions f, g, and h represent values of u on the boundary of the domain, which are often prescribed as initial and boundary conditions.

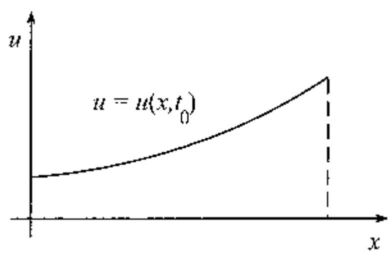

Figure 1.2 Time snapshot $u(x,t_0)$ at $t = t_0$ graphed in xu space. Often several snapshots for different times t are graphed on the same set of xu coordinates to indicate how the wave profiles are evolving in time.

representations are called *time snapshots* or *wave profiles* of the solution; time snapshots are profiles in space of the solution $u = u(x,t)$ frozen at a fixed time t_0, or, stated differently, slices of the solution surface at a fixed time t_0. Occasionally, several time snapshots are plotted simultaneously on the same set of xu axes to indicate how profiles change. It is also helpful on occasion to think of a solution in abstract terms. For example, suppose that $u = u(x,t)$ is a solution of a PDE for $x \in \mathbb{R}$ and $0 \leq t \leq T$. Then for each t, $u(x,t)$ is a function of x (a profile), and it generally belongs to some space of functions $\mathbf{X}$. To fix the idea, suppose that $\mathbf{X}$ is the set of all twice continuously differentiable functions on $\mathbb{R}$ that approach zero at infinity. Then the solution can be regarded as a mapping from the time interval $[0, T]$ into the function space $\mathbf{X}$; that is, to each t in $[0, T]$ we associate a function $u(\cdot, t)$, which is the wave profile at time t.

A PDE has infinitely many solutions, depending on arbitrary functions. For example, the *wave equation*

$$u_{tt} - c^2 u_{xx} = 0 \tag{1.1.2}$$

has a general solution that is the superposition (sum) of a right traveling wave $F(x - ct)$ of speed c and a left traveling wave $G(x + ct)$ of speed c; that is,

$$u(x,t) = F(x - ct) + G(x + ct) \tag{1.1.3}$$

for any twice continuously differentiable functions F and G. (See the Exercises at the end of this section.) We contrast the situation in ordinary differential equations, where solutions depend on arbitrary constants; there, initial or boundary conditions fix the arbitrary constants and select a unique solution. For PDEs this occurs as well; initial and boundary conditions are usually imposed and select one of the infinitude of solutions. These auxiliary or subsidiary conditions are suggested by the underlying physical problem from which the PDE arises, or by the type of PDE. A condition on u or its derivatives given at $t = 0$ along some segment of the x axis is called an *initial condition*, while a condition along any other curve in the xt plane is called a *boundary condition*. PDEs with auxiliary conditions are called *initial value problems*, *boundary value problems*, or *initial–boundary value problems*, depending on the type of subsidiary conditions that are specified.

Example. The initial value problem for the wave equation is

$$u_{tt} - c^2 u_{xx} = 0, \quad x \in \mathbb{R}, \quad t > 0, \tag{1.1.4}$$

$$u(x,0) = f(x), \quad u_t(x,0) = g(x), \quad x \in \mathbb{R}, \tag{1.1.5}$$

where f and g are given twice continuously differentiable functions on $\mathbb{R}$. The unique solution is given by (see Exercise 2)

$$u(x,t) = \frac{1}{2}[f(x - ct) + f(x + ct)] + \frac{1}{2c}\int_{x-ct}^{x+ct} g(s)\, ds, \tag{1.1.6}$$

which is *D' Alembert's formula*. So, in this example we think of the auxiliary data (1.1.5) as selecting one of the infinitude of solutions given by (1.1.3). Note that the solution at (x, t) depends only on the initial data (1.1.5) in the interval $[x - ct, x + ct]$. $\square$

Statements regarding the single second-order PDE (1.1.1) can be generalized in various directions. Higher-order equations (as well as first-order equations), several independent variables, and several unknown functions (governed by systems of PDEs) are all possibilities.

1.1.2 Classification

PDEs are classified into different types, depending on either the type of physical phenomena from which they arise or a mathematical basis. As the reader has learned from previous experience, there are three fundamental types of equations: those that govern diffusion processes, those that govern wave propagation, and those that govern equilibrium phenomena. Equations of mixed type also occur. We consider a single, second order PDE of the for

$$a(x,t)u_{xx} + 2b(x,t)u_{xt} + c(x,t)u_{tt} = d(x,t,u,u_x,u_t), \quad (x,t) \in D, \quad (1.1.7)$$

where a, b, and c are continuous functions on D, and not all of a, b, and c vanish simultaneously at some point of D. The function d on the right side is assumed to be continuous as well. Classification is based on the combination of the second-order derivatives in the equation. If we define the *discriminant* Δ by $\Delta = b^2 - ac$, then (1.1.7) is *hyperbolic* if $\Delta > 0$, *parabolic* if $\Delta = 0$, and *elliptic* if $\Delta < 0$.

Hyperbolic and parabolic equations are evolution equations that govern wave propagation and diffusion processes, respectively, and elliptic equations are associated with equilibrium or steady-state processes. In the latter case, we use x and y as independent variables rather than x and t. There is also a close relationship between the classification and the kinds of initial and boundary conditions that may be imposed on a PDE to obtain a well-posed mathematical problem, or one that is physically relevant. Because classification is based on the highest-order derivatives in (1.1.7), or the *principal part* of the equation, and because Δ depends on x and t, equations may change type as x and t vary throughout the domain.

Now we demonstrate that equation (1.1.7) can be transformed to certain simpler, or *canonical*, forms, depending on the classification, by a change of independent variables

$$\xi = \xi(x,t), \quad \eta = \eta(x,t). \quad (1.1.8)$$

We now perform this calculation, with the view of actually trying to determine (1.1.8) such that (1.1.7) reduces to a simpler form in the $\xi\eta$ coordinate system. The transformation (1.1.8) is assumed to be invertible, which requires that the Jacobian $J = \xi_x\eta_t - \xi_t\eta_x$ be nonzero in any region where the transformation is applied. A straightforward application of the chain rule, which the reader can verify, shows that the left side of (1.1.7) becomes, under the change of independent variables (1.1.8)

$$au_{xx} + 2bu_{xt} + cu_{tt} + \cdots = Au_{\xi\xi} + 2Bu_{\xi\eta} + Cu_{\eta\eta} + \cdots, \tag{1.1.9}$$

where the three dots denote terms with lower-order derivatives, and where

$$\begin{aligned} A &= a\xi_x^2 + 2b\xi_x\xi_t + c\xi_t^2, \\ B &= a\xi_x\eta_x + b(\xi_x\eta_t + \xi_t\eta_x) + c\xi_t\eta_t, \\ C &= a\eta_x^2 + 2b\eta_x\eta_t + c\eta_t^2. \end{aligned}$$

Notice that the expressions for A and C have the same form, namely

$$a\phi_x^2 + 2b\phi_x\phi_t + c\phi_t^2,$$

and are independent.

In the *hyperbolic case* we can choose ξ and η such that $A = C = 0$. To this end, set

$$a\phi_x^2 + 2b\phi_x\phi_t + c\phi_t^2 = 0. \tag{1.1.10}$$

Because the discriminant Δ is positive, we can write (1.1.10) as (assume that a is not zero)

$$\frac{\phi_x}{\phi_t} = -\frac{b \pm \sqrt{b^2 - ac}}{a}.$$

To determine ϕ, we regard it as defining loci (curves) in the xt plane via the equation $\phi(x,t) = \text{const}$. The differentials dx and dt along one of these curves satisfy the relation $\phi_x dx + \phi_t dt = 0$ or $dt/dx = -\phi_x/\phi_t$. Therefore

$$\frac{dt}{dx} = \frac{b \pm \sqrt{b^2 - ac}}{a} \tag{1.1.11}$$

is a differential equation whose solutions determine the curves $\phi(x,t) = \text{const}$. On choosing the $+$ and $-$ signs in (1.1.11), respectively, we obtain $\xi(x,t)$ and $\eta(x,t)$ as integral curves of (1.1.11), making $A = C = 0$. Consequently, if (1.1.7) is hyperbolic, it can be reduced to the *canonical hyperbolic form*

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

where the three dots denote terms involving lower-order derivatives (we leave it as an exercise to show that B is nonzero in this case).

The differential equations (1.1.11) are called the *characteristic equations* associated with (1.1.7), and the two sets of solution curves $\xi(x,t) = \text{const}$ and $\eta(x,t) = \text{const}$ are called the *characteristic curves*, or just the *characteristics*; ξ and η are called *characteristic coordinates*. In summary, in the hyperbolic case there are two real families of characteristics that provide a coordinate system where the equation reduces to a simpler form. Characteristics are the fundamental concept in the analysis of hyperbolic problems because characteristic coordinates form a natural curvilinear coordinate system in which to examine these problems. In some cases, PDEs simplify to ODEs along the characteristic curves.

In the *parabolic case* ($b^2 - ac = 0$) there is just one family of characteristic curves, defined by

$$\frac{dt}{dx} = \frac{b}{a}.$$

Thus we may choose $\xi = \xi(x,t)$ as an integral curve of this equation to make $A = 0$. Then, if $\eta = \eta(x,t)$ is chosen as any smooth function independent of ξ (i.e., so that the Jacobian is nonzero), one can easily determine that $B = 0$ automatically, giving the *parabolic canonical form*

$$u_{\xi\xi} + \cdots = 0.$$

Characteristics rarely play a role in parabolic problems.

In the *elliptic case* ($b^2 - ac < 0$) there are no real characteristics and, as in the parabolic case, characteristics play no role in elliptic problems. However, it is still possible to eliminate the mixed derivative term in (1.1.7) to obtain an elliptic canonical form. The procedure is to determine complex characteristics by solving (1.1.11), and then take real and imaginary parts to determine a transformation (1.1.8) that makes $A = C$ and $B = 0$ in (1.1.9). We leave it as an exercise to show that the transformation is given by

$$\alpha = \frac{1}{2}(\xi + \eta), \quad \beta = \frac{1}{2i}(\xi - \eta).$$

Then the *elliptic canonical form* is

$$u_{\alpha\alpha} + u_{\beta\beta} + \cdots = 0,$$

where the Laplacian operator becomes the principal part.

Example. It is easy to see that the characteristic curves for the wave equation (1.1.2), which is hyperbolic, are the straight lines $x - ct = \text{const}$ and $x + ct = \text{const}$. These are shown in Figure 1.3. In this case the characteristic coordinates are given by $\xi = x - ct$ and $\eta = x + ct$. In these coordinates the wave equation transforms to $u_{\xi\eta} = 0$. We regard characteristics as curves in spacetime moving with speeds c and $-c$, and from the general solution (1.1.3) we observe that signals are propagated along these curves. In hyperbolic problems, in general, the characteristics are curves in spacetime along which signals are transmitted. □

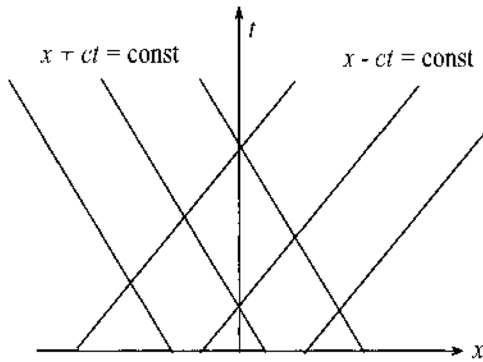

Figure 1.3 Characteristic diagram for the wave equation showing the forward and backward characteristics $x - ct = \text{const}$ and $x + ct = \text{const}$.

If the coefficients a, b, and c of the second-order derivatives in equation (1.1.7) depend on x, t, and u, then (1.1.7) is called a *quasilinear* equation. In this case we make the same classification as above, depending on the sign of the discriminant Δ; now the type of the equation depends not only on the spacetime domain but also on the solution u itself. The canonical forms listed above are no longer valid in this case, and the characteristics defined by (1.1.11) cannot be determined a priori since a, b, and c depend on u, the unknown solution itself. Therefore, there is a significant increase in difficulty when the principal part of the equation is nonlinear.

There are other ways to approach the classification problem. In the preceding discussion the focus was on determining transformations under which a simplification occurs. In Section 6.1 we take a different perspective and ask whether it is possible to determine the solution u near a curve where the values of u and its first derivatives are known. That discussion is accessible to the reader at the present juncture, if desired. Yet another view of classification is presented in Chapter 4, where hyperbolic systems are discussed. Finally, from a physical perspective, we observe later in this chapter that hyperbolic problems are associated with wave propagation; parabolic problems, with diffusion; and elliptic problems, with equilibria.

1.1.3 Linear versus Nonlinear

The most important classification criterion is to distinguish PDEs as *linear* or *nonlinear*. Roughly, a homogeneous PDE is linear if the sum of two solutions is a solution, and a constant multiple of a solutions is a solution. Otherwise, it is nonlinear. The division of PDEs into these two categories is a significant one. The mathematical methods devised to deal with these two classes of equations are often entirely different, and the behavior of solutions differs substantially. One underlying cause is the fact that the solution space to a linear, homogeneous PDE is a vector space, and the linear structure of that space can be used with advantage in constructing solutions with desired properties that can meet diverse boundary and initial conditions. Such is not the case for nonlinear equations.

It is easy to find examples where nonlinear PDEs exhibit behavior with no linear counterpart. One is the breakdown of solutions and the formation of singularities, such as shock waves. A second is the existence of solitons, which are solutions to nonlinear dispersion equations. These solitary wave solutions maintain their shapes through collisions, in much the same was as linear equations do, even though the interactions are not linear. Nonlinear equations have come to the forefront because, basically, the world is nonlinear!

More formally, linearity and nonlinearity are usually defined in terms of the properties of the operator that defines the PDE itself. Let us assume that the PDE (1.1.1) can be written in the form

$$Lu = F, \tag{1.1.12}$$

where $F = F(x, t)$ and L is an operator that contains all the operations (differentiation, multiplication, composition, etc.) that act on $u = u(x, t)$. For example, the wave equation $u_{tt} - u_{xx} = 0$ can be written $Lu = 0$, where L is the partial differential operator $\partial_t^2 - \partial_x^2$. In (1.1.12) we reiterate that all terms involving the unknown function u are on the left side of the equation and are contained in the expression Lu; the right side of (1.1.12) contains in F only expressions involving the independent variables x and t. If $F = 0$, then (1.1.12) is said to be *homogeneous*; otherwise, it is *nonhomogeneous*. We say that an operator L is *linear* if it is additive and if constants factor out of the operator, that is, (1) $L(u + v) = Lu + Lv$, and (2) $L(cu) = cLu$, where u and v are functions (in the domain of the operator) and c is any constant. The PDE (1.1.12) is *linear* if L is a linear operator; otherwise, the PDE is *nonlinear.*

Example. The equation $Lu = u_t + uu_x = 0$ is nonlinear because, for example, $L(cu) = cu_t + c^2uu_x$, which does not equal $cLu = c(u_t + uu_x)$. □

Conditions (1) and (2) stated above imply that a linear homogeneous equation $Lu = 0$ has the property that if $u_1, u_2, \ldots, u_n$ are n solutions, the linear combination

$$u = c_1u_1 + c_2u_2 + \cdots + c_nu_n$$

is also a solution for any choice of the constants $c_1, c_2, \ldots, c_n$. This fact is called the *superposition principle* for linear equations. For nonlinear equations we cannot superimpose solutions in this manner. The superposition principle can often be extended to infinite sums for linear problems, provided that convergence requirements are met. Superposition for linear equations allows one to construct, from a given set of solutions, another solution that meets initial or boundary requirements by choosing the constants $c_1, c_2, \ldots$ judiciously. This observation is the basis for the Fourier method, or eigenfunction expansion method, for linear, homogeneous boundary value problems, and we review this procedure at the end of the section. Moreover, superposition can often be extended to a family of solutions depending on a continuum of values of a parameter. More precisely, if $u = u(x, t; k)$ is a family of solutions of a linear homogeneous PDE for all values of k in some interval of real numbers I, one can superimpose these solutions formally using integration by defining

$$u(x,t) = \int_I c(k)u(x,t;k)\,dk,$$

where $c = c(k)$ is a function of the parameter k. Under certain conditions that must be established, the superposition $u(x,t)$ may again be a solution. As in the finite case, there is flexibility in selecting $c(k)$ to meet boundary or initial conditions. In fact, this procedure is the vehicle for transform methods for solving linear PDEs (Laplace transforms, Fourier transforms, etc.). We review this technique below. Finally, for a homogeneous, linear PDE the real and imaginary parts of a complex solution are both solutions. This is easily seen from the calculation

$$L(v + iw) = Lv + iLw = 0 + 0 = 0,$$

where the real-valued functions v and w satisfy $Lv = 0$ and $Lw = 0$. None of these methods based on superposition are applicable to nonlinear problems, and other methods must be sought. In summary, there is a profound difference between properties and solution methods for linear and nonlinear problems.

If most solution methods for linear problems are inapplicable to nonlinear equations, what methods can be developed? We mention a few.

1. *Perturbation Methods.* Perturbation methods are applicable to problems where a small or large parameter can be identified. In this case an approximate solution is sought as a series expansion in the parameter.

2. *Similarity Methods.* The similarity method is based on the PDE and its auxiliary conditions being invariant under a family of transformations depending on a small parameter. The invariance transformation allows one to identify a canonical change of variables that reduces the PDE to an ordinary differential equation (ODE), or reduces the order of the PDE.

3. *Characteristic Methods.* Nonlinear hyperbolic equations, which are associated with wave propagation, can be analyzed with success in characteristic coordinates (i.e., coordinates in spacetime along which the waves or signals propagate).

4. *Transformations.* Sometimes it is possible to identify transformations that change a given nonlinear equation into a simpler equation that can be solved.

5. *Numerical Methods.* Fast, large-scale computers have given tremendous impetus to the development and analysis of numerical algorithms to solve nonlinear problems and, in fact, have been a stimulus to to the analysis of nonlinear equations.

6. *Traveling Wave Solutions.* Seeking solutions with special properties is a key technique. For example, traveling waves are solutions to evolution problems that represent fixed waveforms moving in time. The assumption of a traveling wave profile to a PDE sometimes reduces it to an ODE, often facilitating the analysis and solution. Traveling wave solutions form one type of similarity solution.

7. *Steady State Solutions and Their Stability.* Many PDEs have steady-state, or time-independent, solutions. Studying these equilibrium solutions and their stability is an important activity in many areas of application.

8. *Ad Hoc Methods.* The mathematical and applied science literature is replete with articles illustrating special methods that analyze a certain type of nonlinear PDE, or restricted classes of nonlinear PDEs.

These methods are primarily solution methods, which represent one aspect of the subject of nonlinear PDEs. Other basic issues are questions of existence and uniqueness of solutions, the regularity (smoothness) of solutions, and the investigation of stability properties of solutions. These and other theoretical questions have spawned investigations based on modern topological and algebraic concepts, and the subject of nonlinear PDEs has evolved into one of the most diverse, active areas of applied analysis.

1.1.4 Linear Equations

In this subsection we review, through examples, two techniques from elementary PDEs that illustrate the use of the superposition principles mentioned above. These calculations arise later in analyzing the local stability of equilibrium solutions to nonlinear problems.

Example. (*Separation of Variables*) Consider the following problem for $u = u(x,t)$ on the bounded interval $I : 0 \le x \le 1$ with $t > 0$, that is

$$u_t = Au, \quad 0 < x < 1, \; t > 0, \tag{1.1.13}$$

$$u(0,t) = u(1,t) = 0, \quad t > 0, \tag{1.1.14}$$

$$u(x,0) = f(x), \quad 0 \le x \le 1, \tag{1.1.15}$$

where A is a linear, spatial differential operator of the form

$$Au = -(pu_x)_x + qu.$$

The functions $p = p(x)$ and $q = q(x)$ are given, with p of one sign on I, and p, p', and q continuous on I. Problems of this type are solved by Fourier's method, or the method of eigenfunction expansions. The idea is to construct infinitely many solutions that satisfy the PDE and the boundary conditions, equations (1.1.13) and (1.1.14), and then superimpose them, rigging up the constants so that the initial condition (1.1.15) is satisfied. This technique is called *separation of variables*, based on an assumption that the solution has the form $u(x,t) = g(t)y(x)$, where g and y are to be determined. When we substitute this form into the PDE and rearrange terms we obtain

$$\frac{g'}{g} = \frac{Ay}{y},$$

where the left side depends only on t and the right side depends only on x. A function of t can equal a function of x for all x and t only if both are equal to a constant, say, $-\lambda$, called the *separation constant*. Therefore

$$\frac{g'}{g} = \frac{Ay}{y} = -\lambda,$$

and we obtain two ODEs, one for g and one for y:

$$g' = -\lambda g, \quad -Ay = \lambda y.$$

We say that the equation separates. If we substitute the assumed form of u into the boundary conditions (1.1.14), then we obtain

$$y(0) = y(1) = 0.$$

The temporal equation is easily solved to get $g(t) = ce^{-\lambda t}$, where c is an arbitrary constant. The spatial equation along with its homogeneous (zero) boundary conditions give a boundary value problem (BVP) for y:

$$-Ay = \lambda y, \quad 0 < x < 1, \tag{1.1.16}$$

$$y(0) = y(1) = 0. \tag{1.1.17}$$

This BVP for y, which is differential eigenvalue problem called a *Sturm Liouville problem*, has the property there are infinitely many real, discrete values of the separation constant λ, say, $\lambda = \lambda_n$, $n = 1, 2, \ldots$, for which there are corresponding solutions $y = y_n(x)$, $n = 1, 2, \ldots$. The λ_n are called the *eigenvalues* for the problem and the corresponding solutions $y = y_n(x)$ are called the *eigenfunctions*. The eigenvalues have the property that they are ordered and $|\lambda_n| \to \infty$ as $n \to \infty$. Therefore we have obtained a countably infinite number of solutions to the PDE that satisfy the boundary conditions:

$$u_n(x,t) = c_n e^{-\lambda_n t} y_n(x), \quad n = 1, 2, \ldots.$$

Now, here is where superposition is used. We add up these solutions and pick the constants c_n so that the initial condition (1.1.15) is satisfied, thus obtaining the solution to the problem; that is, we form

$$u(x,t) = \sum_{n=1}^{\infty} c_n e^{-\lambda_n t} y_n(x).$$

Formally applying the initial condition gives

$$u(x,0) = f(x) = \sum_{n=1}^{\infty} c_n y_n(x). \tag{1.1.18}$$

The right side is an expansion of the initial condition f in terms of the eigenfunctions y_n, and we can use it to determine the coefficients c_n. This calculation is enabled by a very important property of the eigenfunctions, namely, orthogonality. If we define the inner product of two functions ϕ and ψ by

$$(\phi, \psi) = \int_0^1 \phi(x)\psi(x)dx,$$

then we say ϕ and ψ are *orthogonal* if $(\phi, \psi) = 0$. The set of eigenfunctions y_n of the Sturm–Liouville problem (1.1.16)–(1.1.17) are mutually orthogonal, or

$$(y_n, y_m) \equiv \int_0^1 y_n(x) y_m(x) dx = 0, \quad n \neq m.$$

Therefore, if we multiply (1.1.18) by a fixed but arbitrary y_m and formally integrate over the interval I, we then obtain

$$(f, y_m) = \sum_{n=1}^{\infty} c_n (y_n, y_m).$$

Because of orthogonality, the infinite series on the right side collapses to the single term $c_m(y_m, y_m)$. Therefore the coefficient c_m is given by

$$c_m = \frac{(f, y_m)}{(y_m, y_m)}.$$

This relation is true for any m, and so the coefficients c_n are

$$c_n = \frac{(f, y_n)}{(y_n, y_n)}, \quad n = 1, 2, ... \tag{1.1.19}$$

Therefore, we have obtained the solution of (1.1.13)–(1.1.15) in the form of a series representation, or eigenfunction expansion,

$$u(x,t) = \sum_{n=1}^{\infty} \frac{(f, y_n)}{(y_n, y_n)} e^{-\lambda_n t} y_n(x).$$

The preceding calculation took a lot for granted, but it can be shown rigorously that the steps are valid. □

An expansion of a function $f(x)$ in terms of the eigenfunctions $y_n(x)$, as in (1.1.18), is called the generalized *Fourier series* for f, and the coefficients c_n, given by (1.1.19), are the *Fourier coefficients*. It can be shown that that the series converges in the mean-square sense:

$$\int_0^1 \left(f(x) - \sum_{n=1}^{N} c_n y_n(x) \right)^2 dx \to 0 \quad \text{as} \quad N \to \infty.$$

Pointwise and uniform convergence theorems require suitable smoothness conditions on the function f.

The method of separation of variables is successful under general boundary conditions of the form

$$\alpha u(0,t) + \beta u_x(0,t) = 0, \quad \gamma u(1,t) + \delta u_x(1,t) = 0,$$

where α, β, γ, and δ are given constants. Of course, the interval over which the problem is defined may be any *bounded* interval $a \le x \le b$; we chose $a = 0$ and $b = 1$ for simplicity of illustration. The method may be extended to problems over higher-dimensional, bounded, spatial domains, as well as to nonhomogeneous problems. For example, if the PDE in (1.1.13)–(1.1.15) is replaced by the nonhomogeneous equation

$$u_t = Au + F(x,t), \quad 0 < x < 1, \; t > 0,$$

we can expand the nonhomogeneous term F as a Fourier series of the eigenfunctions for the homogenous problem, or

$$F(x,t) = \sum_{n=1}^{\infty} \gamma_n(t) y_n(x),$$

where the $\gamma_n(t)$ are the known Fourier coefficients (t is a parameter in the expansion) that can be computed from orthogonality property of the eigenfunctions. Then we assume the solution takes the form

$$u(x,t) = \sum_{n=1}^{\infty} c_n(t) y_n(x).$$

Substituting these forms into the PDE and the initial condition determines the $c_n(t)$ and therefore the solution to the nonhomogeneous problem.

Problems that are defined over infinite spatial domains require different techniques based on transform methods.

Example. (*Transform Method*) Consider the following problem on an infinite spatial domain:

$$\begin{aligned} u_t &= u_{xx}, \quad x > 0, \quad t > 0, \\ u(0,t) &= 0, \quad t > 0, \\ u(x,0) &= f(x), \quad x > 0. \end{aligned}$$

Because there are two derivatives with respect to x, we expect to impose another boundary condition at infinity. Therefore, we demand that u be bounded as $x \to \infty$. Further, we assume that f is piecewise continuous and absolutely integrable over $x > 0$. We can proceed as in the preceding example and try a solution of the form $u(x,t) = g(t)y(x)$, where g and y are to be determined. Substituting into the differential equation leads to

$$g' = -\lambda g, \quad -y'' = \lambda y,$$

where λ is the separation constant. As before, $g(t) = ce^{-\lambda t}$, where c is an arbitrary constant. The boundary conditions imply that y is bounded and $y(0) = 0$. Thus we have the boundary value problem

$$\begin{aligned} -y'' &= \lambda y, \quad x > 0, \\ y(0) &= 0, \quad y \text{ bounded.} \end{aligned}$$

This is a boundary value problem on the semi-infinite domain $x > 0$. If $\lambda \le 0$ there are no nontrivial, bounded solutions (check this), and therefore $\lambda > 0$. Let us write $\lambda = k^2$; then the general solution to the boundary value problem is

$$y(x) = a \sin kx,$$

where a is an arbitrary constant and $k > 0$. Consequently we have found a family of solutions depending upon a parameter k:

$$u(x,t,k) = ae^{-k^2 t} \sin kx.$$

In contrast to the last example, on a bounded interval, where the eigenvalues were discrete, the eigenvalues in the present case form a continuum. We superimpose these solutions over all k and write

$$u(x,t) = \int_0^\infty a(k)e^{-k^2 t}\sin kx\,dk,$$

where $a(k)$ is a function of k. We can determine $a(k)$ from the initial condition. Putting $t=0$ in the last equation gives

$$f(x) = \int_0^\infty a(k)\sin kx\,dk. \tag{1.1.20}$$

This is an integral equation from which we can recover $a(k)$ using a special case of the Fourier integral theorem: If f is piecewise continuous and absolutely integrable on $x>0$, and if

$$f(x) = \int_0^\infty a(k)\sin kx\,dk,$$

then

$$a(k) = \frac{2}{\pi}\int_0^\infty f(\xi)\sin k\xi\,d\xi.$$

The function a is the Fourier sine transform of f, and f is the inverse sine transform of a. Putting everything together gives an integral representation of the solution to the problem, namely

$$\begin{aligned} u(x,t) &= \frac{2}{\pi}\int_0^\infty \left(\int_0^\infty f(\xi)\sin k\xi\,d\xi\right) e^{-k^2 t}\sin kx\,dk \\ &= \frac{2}{\pi}\int_0^\infty f(\xi)\left(\int_0^\infty e^{-k^2 t}\sin k\xi\,\sin kx\,dk\right) d\xi, \end{aligned}$$

where we have changed the order of integration in the last step. Actually, the interior integral can be calculated analytically, or looked up in a table, which we leave as an exercise. □

The reader should notice the great similarity of the solution forms in these two examples—finite domain versus infinite domain, discrete eigenvalues versus a continuum of eigenvalues, and expansions in terms of sums versus integrals. The role of superposition is critical in linear problems, but it does not carry over to nonlinear problems.

Terminology. We introduce some notation and terminology for some function spaces that commonly occur in analysis. Let D be an open domain (a set that does not contain any of its boundary) in either one or several dimensions. Because D is an open domain we do not have to deal with the question of existence of derivatives at boundary points. Using an overbar, $\overline{D}$, we denote the closure of D, which consists of D and its boundary ∂D; that is, $\overline{D} = D \cup \partial D$. By $C^n(D)$ we denote the set of all continuous functions on D that have n continuous derivatives in D (partial derivatives if D is of dimension greater than 1). The space of continuous functions on D is denoted by $C(D)$, and $C^\infty(D)$ denotes the space of continuous functions on D that have derivatives of all orders. When a function u belongs to one of these sets, e.g., $C^n(D)$, we sometimes say that u is of class C^n on D.

EXERCISES

1. Show that the wave equation $u_{tt} - c^2 u_{xx} = 0$ reduces to the canonical form $u_{\xi\eta} = 0$ under the change of variables $\xi = x - ct, \eta = x + ct$, and use this information to show that the general solution of the wave equation is $u(x,t) = F(x-ct) + G(x+ct)$, where F and G are arbitrary $C^2(\mathbb{R})$ functions.

2. Derive D'Alembert's formula for the initial value problem for the wave equation using Exercise 1 and determining F and G from the initial conditions.

3. Let u be of class C^3. Show that $u = u(x,t)$ is a solution of the wave equation

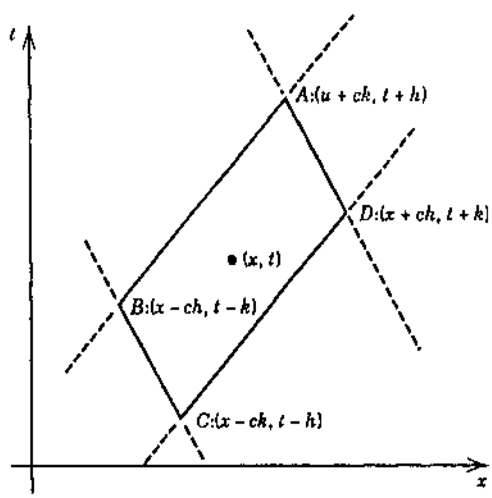

Figure 1.4 Characteristic parallelogram $ABCD$ whose sides are characteristic straight lines $x - ct = \text{const}$ and $x + ct = \text{const}$. The numbers h and k are positive constants that define the size of $ABCD$.

$u_{tt} - c^2 u_{xx} = 0$ if, and only if, u satisfies the difference equation

$$\begin{aligned} &u(x - ck, t - h) + u(x + ck, t + h) \\ &= u(x - ch, t - k) + u(x + ch, t + k) \end{aligned}$$

for all constants $h, k > 0$. Interpret this result geometrically in the xtplane by observing that the difference equation relates the value of u at the vertices of a characteristic parallelogram whose sides are the characteristic straight lines $x + ct = \text{const}$ and $x - ct = \text{const}$ (see Figure 1.4).

4. Assuming that the initial data f and g for the initial value problem for the wave equation in Exercise 3 have compact support (i.e., f and g vanish for $|x|$ sufficiently large), prove that the solution $u(x,t)$ has compact support in x for each fixed time t.

5. Find the general solution of the PDE

$$x^2 u_{xx} + 2xt u_{xt} + t^2 u_{tt} = 0$$

by transforming the equation to canonical form using characteristic coordinates $\xi = t/x,\ \eta = t$.

6. Let $u = u(x,t)$ be a solution of the nonlinear equation

$$a(u_x, u_t)u_{xx} + 2b(u_x, u_t)u_{xt} + c(u_x, u_t)u_{tt} = 0.$$

Introduce new independent variables via $\xi = \xi(x,t)$ and $\eta = \eta(x,t)$, and a new function $\phi = \phi(\xi, \eta)$ defined by $\phi = xu_x + tu_t - u$. Prove that $\phi_\xi = x$, $\phi_\eta = t$, and ϕ satisfies the *linear* PDE

$$a(\xi,\eta)\phi_{\eta\eta} - 2b(\xi,\eta)\phi_{\xi\eta} + c(\xi,\eta)\phi_{\xi\xi} = 0.$$

(This transformation, known as a *hodograph*, or a *Legendre*, transformation, transforms a nonlinear equation of the given form to a linear equation by reversing the roles of the dependent and independent variables.)

7. Find a formula for the solution of the initial-boundary value problem

$$\begin{aligned} u_{tt} &= c^2 u_{xx}, \quad x > 0, \quad t > 0, \\ u(x,0) &= f(x), \quad u_t(x,0) = g(x), \quad x > 0, \\ u(0,t) &= h(t), \quad t > 0. \end{aligned}$$

Hint: Use D'Alembert's formula for $x > ct$ and the difference equation in Exercise 3 for $0 < x < ct$. Assume sufficient differentiability.

8. Solve the outgoing signaling problem

$$u_{tt} = c^2 u_{xx}, \quad x > 0, \quad t \in \mathbb{R},$$
$$u(0,t) = s(t), \quad t \in \mathbb{R}.$$

9. Consider the PDE

$$4u_{xx} + 5u_{xt} + u_{tt} = 2 - u_t - u_x.$$

Find the characteristic coordinates and graph the characteristic curves in the xt plane. Reduce the equation to canonical form and find the general solution.

10. Classify the PDE

$$xu_{xx} - 4u_{tt} = 0.$$

In the case $x > 0$ find the characteristic coordinates and sketch the characteristics in an appropriate region in the xt plane.

11. Use the separation of variables method to find a series representation of the solution to the following problems:

(a)

$$u_t = Du_{xx}, \quad 0 < x < \pi, \quad t > 0,$$
$$u(0,t) = 0, \quad u(\pi,t) = 0, \quad t > 0,$$
$$u(x,0) = f(t), \quad 0 < x < \pi.$$

(b)

$$u_{tt} = c^2 u_{xx}, \quad 0 < x < 1, \quad t > 0,$$
$$u(0,t) = 0, \quad u(1,t) = 0, \quad t > 0,$$
$$u(x,0) = 0, \quad u_t(x,0) = g(t), \quad 0 < x < \pi.$$

(c)

$$u_{xx} + u_{yy} = 0, \quad 0 < x < 1, \ 0 < y < \pi,$$
$$u(x,0) = 0, \quad u_y(x,\pi) = 0, \quad 0 < x < 1,$$
$$u(0,y) = 0, \quad u(1,y) = g(y), \quad 0 < y < \pi.$$

12. Find the general solution of the Euler–Darboux equation

$$u_{x,y} = \frac{m}{x-y}(u_x - u_y)$$

in the case $m = 1$. *Hint*: Look at $((x-y)u)_{xy}$.

1.2 Conservation Laws

Many of the fundamental equations in the natural and physical sciences are obtained from *conservation laws*, which are balance laws, equations expressing the fact that some quantity is balanced throughout a process. In thermodynamics, for example, the first law states that the change in internal energy in a system is equal to, or is balanced by, the total heat added to the system plus the work done on the system. Thus the first law of thermodynamics is really an energy balance law, or conservation law. As another example, consider a fluid flowing in some region of space that consists of chemical species undergoing chemical reaction. For a given chemical species, the time rate of change of the total amount of that species in the region must equal the rate at which the species flows into the region, minus the rate at which the species flows out, plus the rate at which the species is created, or consumed, by the chemical reactions. This is a verbal statement of a conservation law for the amount of the given chemical species. Similar balance or conservation laws occur in all branches of science. In the biosciences, for example, the rate of change of an animal population in a fixed region must equal the birth rate, minus the death rate, plus the migration rate (emigration or immigration) into or out of the region.

1.2.1 One Dimension

Mathematically, conservation laws translate into integral or differential equations, which are then regarded as the *governing equations* or *equations of motion* of the process. These equations dictate how the process evolves in time. Here we are interested in processes governed by partial differential equations. We now formulate the basic one-dimensional conservation law, out of which will evolve some of the basic models and concepts in nonlinear PDEs.

Let us consider a quantity $u = u(x, t)$ that depends on a single spatial vari-

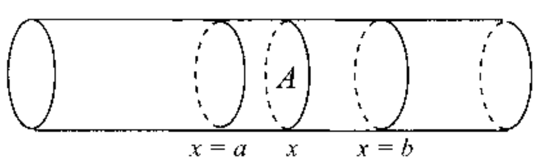

Figure 1.5 Cylindrical tube of cross-sectional area A showing a cross section at x and a finite section I: $a \le x \le b$. There is no variation of any quantity in a fixed cross section.

able $x \in \mathbb{R}$ and time $t > 0$. We assume that u is a density or concentration measured in an amount per unit volume, where the amount may refer to population, mass, energy, or any quantity. By definition, u varies in only one spatial direction, the direction denoted by x. We imagine further that the quantity is distributed in a tube of cross-sectional area A (see Figure 1.5). Again, by assumption, u is constant in any cross section of the tube, and the variation is only in the x direction. Now consider an arbitrary segment of the tube denoted by the interval $I = [a, b]$. The total amount of the quantity u inside I at time t is

$$\text{Total amount of quantity in } I = \int_a^b u(x,t) A \, dx.$$

Now assume that there is motion of the quantity in the tube in the axial direction. We define the *flux* $\phi(x,t)$ of u at x at time t as the amount of the quantity u flowing through the cross section at x at time t, per unit area, per unit time. Thus the dimensions of ϕ are $[\phi] = \text{amount}/(\text{area} \cdot \text{time})$, where the bracket notation denotes *dimensions of*. By convention, we take ϕ to be positive if the flow at x is in the positive x direction, and to be negative at x if the flow is in the negative x direction. Therefore, at time t the net rate that the quantity is flowing into the interval I is the rate that it is flowing in at $x = a$ minus the rate that it is flowing out at $x = b$:

$$\text{Net rate that the quantity flows into } I = A\phi(a,t) - A\phi(b,t).$$

Finally, the quantity u may be created or destroyed inside I by an external or internal source (e.g., by a chemical reaction if u were a species concentration, or by birth or death if u were a population density). We denote this *source function*, which is a local function acting at each x, by $f(x,t,u)$ and its dimensions are given by $[f] = \text{amount}/(\text{volume} \cdot \text{time})$. (The source f could also depend on derivatives of u.) Consequently, f is the rate that u is created (or destroyed) at x at time t, per unit volume. Note that the source function f may depend on u itself, as well as space and time. If f is positive, we say that it is a *source*, and if f is negative, we say that it is a *sink*. Now, given f, we may calculate the total rate that u is created in I by integration. We have

$$\text{Rate that quantity is produced in} I \text{ by sources} = \int_a^b f(x,t,u(x,t)) A \, dx.$$

The fundamental conservation law may now be formulated for the quantity u: For any interval I, we have

$$\begin{aligned} &\text{Time rate of change of the total amount in } I \\ &\quad = \text{net rate that the quantity flows into } I \\ &\quad\quad + \text{rate that the quantity is produced in } I. \end{aligned}$$

In terms of the mathematical symbols and expressions that we introduced above, after canceling the constant cross-sectional area A, we have

$$\frac{d}{dt}\int_a^b u(x,t)\,dx = \phi(a,t) - \phi(b,t) + \int_a^b f(x,t,u)\,dx. \tag{1.2.1}$$

In summary, (1.2.1) states that the rate that u changes in the interval I must equal the net rate at which u flows into I plus the rate that u is produced in I by sources. Equation (1.2.1) is called a *conservation law in integral form*, and it holds even if u, ϕ, or f is not a smooth (continuously differentiable) function. The latter remark is important when we consider physical processes giving rise to shock waves, or discontinuous solutions, in subsequent chapters.

If conditions are placed on the triad u, ϕ, and f, then (1.2.1) may be transformed into a single PDE. Two results from elementary integration theory are required to make this transformation: (1) the fundamental theorem of calculus, and (2) the result on differentiating an integral with respect to a parameter in the integrand. Precisely:

1. $\int_a^b \phi_x(x,t)dx = \phi(b,t) - \phi(a,t)$.
2. $d/dt \int_a^b u(x,t)dx = \int_a^b u_t(x,t)dx$.

These results are valid if ϕ and u are continuously differentiable functions on $\mathbb{R}^2$. Of course, (1) and (2) remain correct under less stringent conditions, but the assumption of smoothness is all that is required in the subsequent discussion. Therefore, assuming smoothness of u and ϕ, as well as continuity of f, equations (1) and (2) imply that the conservation law (1.2.1) may be written

$$\int_a^b \left[u_t(x,t) + \phi_x(x,t) - f(x,t,u)\right] dx = 0 \quad \text{for all intervals } I = [a,b]. \tag{1.2.2}$$

Because the integrand is a continuous function of x, and because (1.2.2) holds for all intervals of integration I, it follows that the integrand must vanish identically:

$$u_t + \phi_x = f(x,t,u), \quad x \in \mathbb{R}, \quad t > 0. \tag{1.2.3}$$

Equation (1.2.3) is a PDE relating the density $u = u(x,t)$ and the flux $\phi = \phi(x,t)$. Both are regarded as unknowns, whereas the form of the source function f is assumed to be given. Equation (1.2.3) is called a *local conservation law*, in contrast to the integral form (1.2.1). The ϕ_x term is called the *flux term* because it arises from the movement, or transport, of u through the cross section at x. The source term f is called a *reaction term* (especially in chemical contexts) or a *growth* or *interaction* term (in biological contexts). Finally, we have defined the flux ϕ as a function of x and t; this dependence on space and time may occur through dependence on u or its derivatives. For example, a physical assumption may require us to posit $\phi(x,t) = \phi(x,t,u(x,t))$, where the flux is dependent on u itself.

It is important to observe that (1.2.3) was derived under assumptions of smoothness. If smoothness of the density or flux is not guaranteed, as occurs in the study of discontinuous solutions, then (1.2.3) must be abandoned in favor of the integral form (1.2.1) of the conservation law, which is always valid. This issue becomes the focus of study in Chapter 3.

Finally, we make a general observation about deriving conservation laws. In the preceding discussion the balance law was applied to an entire interval I. Assuming smoothness, we can derive the conservation law directly using a small interval $[x, x + \Delta x]$ and then take the limit as $\Delta x \to 0$. Applying the balance law to the small box $[x, x + \Delta x]$, we obtain

$$\frac{\partial}{\partial t}\left(u(\xi,t)A\,\Delta x\right) = A\phi(x,t) - A\phi(x+\Delta x,t) + f(t,\eta,u(\eta,t))A\,\Delta x,$$

where ξ and η are points in $[x, x+\Delta x]$, guaranteed by the mean value theorem. Now, dividing by $A\,\Delta x$ gives

$$\frac{\partial}{\partial t}u(\xi,t) = \frac{\phi(x,t) - \phi(x+\Delta x,t)}{\Delta x} + f(t,\eta,u(\eta,t)).$$

Taking the limit as $\Delta x \to 0$ gives the conservation law (1.2.3). These two methods for obtaining the conservation law are called, appropriately, the *large-box method* and the *small-box method*.

1.2.2 Higher Dimensions

It is straightforward to formulate conservation laws in higher dimensions. In this section we limit the discussion to three-dimensional Euclidean space $\mathbb{R}^3$. Let $x = (x_1, x_2, x_3)$ denote a point in $\mathbb{R}^3$, and assume that $u = u(x,t)$ is a scalar density function representing the amount per unit volume of some quantity of interest distributed throughout a domain $D \subset \mathbb{R}^3$. In this domain let V be an arbitrary region with a smooth boundary denoted by ∂V. By an argument similar to the one-dimensional case, the total amount of the quantity in V is given by the volume integral

$$\text{Total amount in } V = \int_V u(x,t)\,dx,$$

where $dx = dx_1 dx_2 dx_3$ represents a volume element in $\mathbb{R}^3$. We prefer to write the volume integral over V with a single integral sign rather than the usual triple integral. Now, we know that the time rate of change of the total amount in V must be balanced by the rate that the quantity is produced in V by

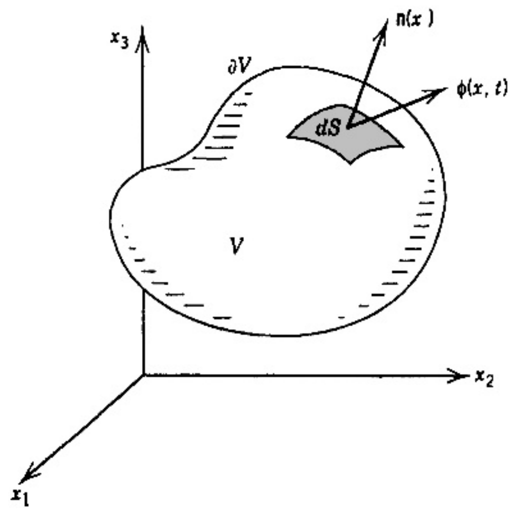

Figure 1.6 Volume V with boundary ∂V showing a surface element dS with outward normal $\mathbf{n}$ and flux vector ϕ. The outward normal determines the orientation of the surface element.

sources, plus the net rate that the quantity flows through the boundary of V. We let $f(x, t, u)$ denote the source term, so that the rate that the quantity is produced in V is given by

$$\text{Rate that } u \text{ is produced by sources } = \int_V f(x, t, u)\, dx.$$

In three dimensions the flow can be in any direction, and therefore the flux is given by a vector $\phi(x, t)$. If $\mathbf{n}(x)$ denotes the outward unit normal vector to the region V (see Figure 1.6), then the net outward flux of the quantity u through the boundary ∂V is given by the surface integral

$$\text{Net outward flux through } \partial V = \int_{\partial V} \phi(x, t) \cdot \mathbf{n}(x)\, dS,$$

where dS denotes a surface element on ∂V. Hence, the balance law for u is given by

$$\frac{d}{dt} \int_V u\, dx = - \int_{\partial V} \phi \cdot \mathbf{n}\, dS + \int_V f\, dx. \tag{1.2.4}$$

The minus sign on the flux term occurs because outward flux decreases the rate that u changes in V.

The integral form of the conservation law (1.2.4) can be reformulated as a local condition, that is, a PDE, provided that u and ϕ are sufficiently smooth functions. In this case the surface integral can be written as a volume integral over V using the *divergence theorem* (the divergence theorem is the fundamental theorem of calculus in three dimensions)

$$\int_V \operatorname{div} \boldsymbol{\phi}\, dx = \int_{\partial V} \boldsymbol{\phi} \cdot \mathbf{n}\, dS, \tag{1.2.5}$$

where div is the divergence operator. Using (1.2.5) and bringing the derivative under the integral on the left side of (1.2.4) yields

$$\int_V u_t\, dx = -\int_V \operatorname{div} \boldsymbol{\phi}\, dx + \int_V f\, dx.$$

The arbitrariness of V then implies the differential form of the balance law:

$$u_t + \operatorname{div} \boldsymbol{\phi} = f(x, t, u), \quad x \in D, \quad t > 0. \tag{1.2.6}$$

Equation (1.2.6) is the three-dimensional version of equation (1.2.3).

EXERCISES

1. The derivation of the fundamental conservation law (1.2.3) assumed the cross-sectional area A of the tube to be constant. Derive integral and differential forms of the conservation law assuming that the area is a slowly varying function of x, that is, $A = A(x)$ and $A'(x)$ is small. [Note that $A(x)$ cannot change significantly over small changes in x; otherwise, the one-dimensional assumption of the state functions u and ϕ being constant in any cross section would be violated.]

2. Assuming that there are no sources and $\phi = \phi(u)$, show that the conservation law (1.2.1) is equivalent to

$$\int_a^b u(x, t_2)\, dx = \int_a^b u(x, t_1)\, dx + \int_{t_1}^{t_2} \phi(u(a, t))\, dt - \int_{t_1}^{t_2} \phi(u(b, t))\, dt$$

for all t_1 and t_2.

1.3 Constitutive Relations

Because equation (1.2.3), or (1.2.6), is a single PDE for two unknown quantities (the density u and the flux ϕ), intuition indicates that another equation is required to have a well-determined system. This additional equation is a relation that is usually based on an assumption about the physical properties of the

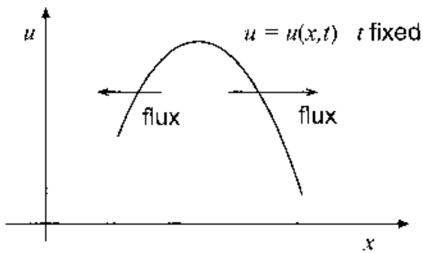

Figure 1.7 Time snapshot of the density distribution $u(x, t)$ illustrating Fick's law. The arrows indicate the direction of the flow, from higher concentrations to lower concentrations. The flow is said to be *down the gradient*.

medium, or the processes involved, which in turn is based on empirical reasoning. Equations expressing these assumptions are called *constitutive relations* or *equations of state*. Thus, a constitutive equation is on a different level from the basic conservation law; the latter is a fundamental law of nature connecting the density u to the flux ϕ, whereas a constitutive relation is an approximate equation whose origin is in empirics. We present several key examples.

Example. (*Diffusion Equation*) At the outset assume that no sources are present ($f = 0$) and the process is governed by the basic conservation law in one dimension

$$u_t + \phi_x = 0, \quad x \in \mathbb{R}, \quad t > 0. \tag{1.3.1}$$

In many physical processes it is observed that the amount of the substance, represented by its density u, flowing through a cross section at x at time t is proportional to the density gradient u_x; that is, $\phi(x, t) \propto u_x(x, t)$. If $u_x > 0$, then $\phi < 0$ and the substance flows to the left, and if $u_x < 0$, then $\phi > 0$ and the substance flows to the right. Figure 1.7 illustrates the situation. We say that u flows *down the gradient*. For example, the second law of thermodynamics states that heat behaves in this manner; heat flows from hotter regions to colder regions, and the steeper the temperature distribution curve, the more rapid the flow of heat. As another example, if u represents a concentration of insects, one may observe that insects move from high concentrations to low concentrations with a rate proportional to the concentration gradient. Therefore, we assume the basic constitutive law

$$\phi(x, t) = -Du_x(x, t), \tag{1.3.2}$$

which is known as *Fick's law*. Processes described by this law are called *linear diffusion processes*. The positive proportionality constant D is called the *diffusion constant*, and it has dimensions $[D] = \text{length}^2/\text{time}$. Fick's law accurately describes the behavior of many physical and biological systems, and in Chapter 5 we give a supporting argument for it on the basis of a probability model and random walk.

Equations (1.3.1) and (1.3.2) give a pair of PDEs for the two unknowns u and ϕ. They combine easily to form a single second-order linear PDE for the unknown density $u = u(x,t)$ given by

$$u_t - Du_{xx} = 0. \tag{1.3.3}$$

Equation (1.3.3), called the *diffusion equation*, governs conservative processes when the flux is specified by Fick's law. It may not be clear at this time why (1.3.3) should be termed the diffusion equation; suffice it to note for the moment that Fick's law implies that a substance moves into adjacent regions because of concentration gradients. We refer to the Du_{xx} term in (1.3.3) as the *diffusion term*. □

The diffusion constant D defines a crude characteristic time (or time scale) T for the process. If L is a length scale (e.g., the length of the container), the quantity

$$T = \frac{L^2}{D} \tag{1.3.4}$$

is the only constant in the process with dimensions of time, and T gives a measure of the time required for discernible changes in concentration to occur over the length L.

Example. (*Heat Equation*) If $u = u(x,t)$ is the thermal energy density (energy per volume) in a heat-conducting medium, then $u = \rho CT$, where ρ is the mass density (mass per unit volume), C is the specific heat (energy per unit mass per degree), and $T = T(x,t)$ is the temperature (degrees). In the absence of sources the conservation law is

$$(\rho CT)_t + \phi_x = 0,$$

where ϕ is the energy flux. The constitutive law, which is a heat analog of Fick's law, is

$$\phi = -KT_x(x,t),$$

where K is the thermal conductivity of the region. In heat transfer this is called *Fourier's law of heat conduction*. Thus, heat energy moves down the temperature gradient. The conservation law becomes

$$T_t - kT_{xx} = 0,$$

where $k = K/\rho C$ is the diffusivity. This is the *heat equation*, which is just the diffusion equation in the context of heat flow. □

Example. (*Reaction–Diffusion Equation*) If sources are present ($f \neq 0$), the conservation law

$$u_t + \phi_x = f(x, t, u) \tag{1.3.5}$$

and Fick's law (1.3.2) combine to give

$$u_t - Du_{xx} = f(x, t, u), \tag{1.3.6}$$

which is called a *reaction–diffusion equation.* Reaction–diffusion equations are nonlinear if the reaction term f is nonlinear in u. These equations are of great interest in nonlinear analysis and applications, particularly in combustion processes and in biological systematics. □

Example. (*Fisher's Equation*) In studies of elementary population dynamics, one proposal is that populations are governed by the logistic law, which states that the rate of change of the total population $u = u(t)$ in a fixed spatial domain is given by

$$\frac{du}{dt} = ru\left(1 - \frac{u}{K}\right), \tag{1.3.7}$$

where $r > 0$ is the *growth rate* and $K > 0$ is the *carrying capacity.* Initially, if u is small, the linear growth term ru in (1.3.7) dominates and rapid population growth results; as u becomes large, the quadratic competition term $-ru^2/K$ kicks in to inhibit the growth. For large times t, the population equilibrates toward the asymptotically stable state $u = K$, the carrying capacity. Now we add spatial effects. Suppose that the population u is a population density and depends on a spatial variable x as well as time t; that is, $u = u(x, t)$. Then, as in the preceding discussion, a conservation law may be formulated as

$$u_t + \phi_x = ru\left(1 - \frac{u}{K}\right), \tag{1.3.8}$$

where $f = f(u) = ru(1 - u/K)$ is the assumed local source term given by the logistics growth law, and ϕ is the population flux. Assuming Fick's law for the flux, we have

$$u_t - Du_{xx} = ru\left(1 - \frac{u}{K}\right). \tag{1.3.9}$$

The reaction–diffusion equation (1.3.9) is *Fisher's equation*, after R. A. Fisher, who studied the equation in the context of investigating the distribution of an advantageous gene as it diffuses through a given population. Discussion of this equation, which is one of the fundamental equations in mathematical biology, is given in subsequent chapters. □

Example. (*Burgers' Equation*) To the basic conservation law with no sources we now append the constitutive relation

$$\phi = -Du_x + Q(u). \tag{1.3.10}$$

The density u then satisfies

$$u_t - Du_{xx} + Q(u)_x = 0. \tag{1.3.11}$$

Now there are two terms contributing to the flux, a Fick's law type of term $-Du_x$, which introduces a diffusion effect, and a flux term $Q(u)$, depending only on u itself, that leads to what is interpreted later as advection. In the special case that $Q(u) = u^2/2$, equation (1.3.11) can be written

$$u_t + uu_x = Du_{xx}, \tag{1.3.12}$$

which is *Burgers' equation*. This is one of the fundamental model equations in fluid mechanics and illustrates the coupling between advection and diffusion. Later we derive (1.3.12) using a weakly nonlinear approximation of the equations of gas dynamics. When $D = 0$ (no diffusion), then

$$u_t + uu_x = 0, \tag{1.3.13}$$

which is called the *inviscid Burgers' equation* (also, Riemann's equation). It is the prototype equation for nonlinear advection and arises in gas dynamics, traffic flow, chromatography, and flood waves in rivers. Equation (1.3.13) is hyperbolic and describes wave propagation, whereas (1.3.12) is parabolic and models diffusion. □

Example. (*Advection Equation*) The simplest flux term occurs when the material forming the density is carried along by the medium having a fixed velocity, as in the case of particulants carried by, for example, wind or water. In these cases the flux is given by the simple linear relationship

$$\phi = cu, \tag{1.3.14}$$

where c is a positive constant having the dimensions of speed. The basic conservation law is

$$u_t + cu_x = 0, \tag{1.3.15}$$

which is the *advection equation*. The term *advection* refers to the horizontal movement of a physical property (e.g., a density wave); other equivalent terms are *convection* and *transport*, which have the same meaning. For example, biologists use the term advection while many engineers use the term convection. We use these terms interchangeably. This linear first-order PDE (1.3.15) is the simplest wave equation. Figure 1.8 compares how a signal at time $t = 0$ propagates under advection, diffusion, and advection-diffusion processes. □

Example. (*Diffusion in* $\mathbb{R}^3$) In Section 1.2 we obtained the basic conservation law

$$u_t + \operatorname{div} \boldsymbol{\phi} = f(x, t, u).$$

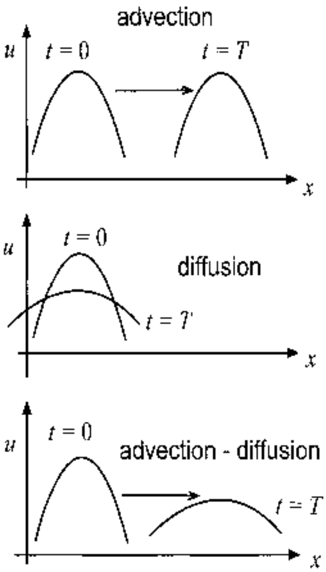

Figure 1.8 Comparison of advection, diffusion, and advection–diffusion processes.

In $\mathbb{R}^3$ Fick's law takes the form $\phi = -D \operatorname{grad} u$; that is, the flux is in the direction of the negative gradient of u. Recall the direction of maximum increase is in the direction of the gradient. Using the vector identity $\operatorname{div}(\operatorname{grad} u) = \Delta u$, where

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}$$

is the *Laplacian* operator, and assuming that the diffusion coefficient D is constant, the conservation law becomes

$$u_t - D\Delta u = f(x, t, u),$$

which is a reaction–diffusion equation in $\mathbb{R}^3$. If there are no sources ($f = 0$), we obtain the three-dimensional diffusion equation $u_t - D\Delta u = 0$. □

In summary, we introduced a few of the fundamental PDEs that have been examined extensively by practitioners of the subject. An understanding of these equations and a development of intuition regarding the fundamental processes that they, and other equations, describe is one of the goals of this treatment. We end this section with the problem of modeling flow through a porous medium.

Example. (*Porous Media*) Consider a fluid (e.g., water) seeping downward through the soil, and let $\rho = \rho(x, t)$ be the density of the fluid, with positive

x measured downward. In a given volume of soil only a fraction of the space is available to the fluid; the remaining space is reserved for the soil itself. In this sense, the soil is a porous medium through which the fluid flows. If the fraction of the volume that is available to the fluid is denoted by κ, called the *porosity*, the fluid mass balance law in a given section of unit cross-sectional area between $x = a$ and $x = b$ is given by

$$\frac{d}{dt}\int_a^b \kappa\rho(x,t)\,dx = \phi(a,t) - \phi(b,t), \tag{1.3.16}$$

where ϕ is the mass flux. We assume that the mass flux is $\phi = \rho v$, where v is the volumetric flow rate. Assuming requisite smoothness, the integral balance law (1.3.16) can be written

$$\kappa\rho_t + (\rho v)_x = 0. \tag{1.3.17}$$

Here we assumed that the porosity is constant, but in general it could depend on x or even ρ. The conservation law (1.3.17) contains two unknowns, the density and the flow rate, and therefore we need a constitutive equation that relates the two. For reasonably slow flows it is observed experimentally that the volumetric flow rate is given by

$$v = -\frac{\mu}{\nu}(p_x - g\rho), \tag{1.3.18}$$

where $p = p(x,t)$ is the pressure, and the positive constants g, μ, and ν are the acceleration due to gravity, the permeability, and the viscosity of the fluid, respectively. Equation (1.3.18) is *Darcy's law*, and it is a basic assumption in many groundwater problems. It is physically plausible because it confirms our intuition that the flow rate should depend on the pressure gradient as well as gravity. Darcy's law is a statement replacing a momentum balance law. When (1.3.18) is substituted into (1.3.17), we obtain the *groundwater equation*:

$$\kappa\rho_t - \frac{\mu}{\nu}(\rho p_x + g\rho^2)_x = 0. \tag{1.3.19}$$

Equation (1.3.19) contains two unknowns, ρ and p, because the constitutive relation (1.3.18) introduced yet another unknown, the pressure. Consequently, we require another equation to obtain a determined system.

If the fluid is a gas, it is common to neglect gravity and assume an equation of state for the gas of the form $p = p(\rho)$. In particular, we assume a *γ-law gas* having equation of state

$$\frac{p}{p_0} = \left(\frac{\rho}{\rho_0}\right)^\gamma, \tag{1.3.20}$$

where p_0 and ρ_0 are positive constants and $\gamma > 1$. Substituting (1.3.20) into (1.3.19) and setting $g = 0$ yields a single PDE for the density ρ having the form

$$\rho_t - \frac{\alpha}{\kappa}(\rho^\gamma \rho_x)_x = 0, \tag{1.3.21}$$

where $\alpha = \gamma\mu p_0/\nu\rho_0^\gamma$. Equation (1.3.21) is a nonlinear diffusion equation called the *porous medium equation*, and it governs flows subject to three laws: mass conservation, Darcy's law, and the gas equation of state. See Aronson (1986) for a survey on the porous medium equation. □

Finally, we point out that it is often crucial to nondimensionalize a problem. Differential equations, when formulated, involve dimensioned dependent and independent variables such as time, length, and temperature. The Buckingham Pi theorem guarantees that a dimensionally consistent physical law can always be transformed to one where the variables, as well as the parameters, are dimensionless. A valid comparison of the relative magnitudes of the terms in an equation can be made only when a problem is nondimensionalized. Further, the dimensionless problem often offers an economy over the dimensioned version in that there is a reduction in the number of parameters. The process of nondimensionalization is sometimes called *scaling*. Lin & Segel (1974) and Logan (2006a) thoroughly discuss scaling and dimensional analysis.

Example. (*Scaling*) In Fisher's equation,

$$u_t = Du_{xx} + ru\left(1 - \frac{u}{K}\right),$$

the variables t, x, and u have dimensions is time, length, and animals per area, respectively. The parameters are the growth rate r with units of 1/time, the carrying capacity K with units of animals per area, and the diffusion constant D with dimensions length-squared per time. Using the parameters we can build dimensionless variables by defining

$$\tau = \frac{t}{r^{-1}}, \quad \xi = \frac{x}{\sqrt{D/r}}, \quad v = \frac{u}{K}.$$

Note that each has the form of a dimensioned variable divided by a constant with the same dimension. We refer to these constants in the denominator as *scales*. For example, the population density is scaled by K, which means that the population is being measured relative to the carrying capacity; r^{-1} is the time scale, meaning that time is being measured relative to the (inverse) growth rate, and so on. There are usually several ways to determine the scales. By the chain rule we can transform the PDE into the dimensionless variables. Observe that derivatives in the PDE transform via

$$\frac{\partial u}{\partial t} = rK\frac{\partial v}{\partial \tau}, \quad \frac{\partial^2 u}{\partial x^2} = \frac{K}{D/r}\frac{\partial^2 v}{\partial \xi^2},$$

and therefore the PDE becomes

$$rK\frac{\partial v}{\partial \tau} = D\frac{K}{D/r}\frac{\partial^2 v}{\partial \xi^2} + rKv(1-v),$$

or

$$v_\tau = v_{\xi\xi} + v(1-v).$$

Therefore, in dimensionless variables Fisher's equation reduces to a model equation without any constants at all. This simpler equation may be analyzed, and, if required, a return to interpretations in terms of the original dimensioned quantities can be made. □

EXERCISES

1. Write the PDE

$$u_t + uu_x + u_{xxx} = 0$$

in the form of a conservation law, identifying the flux ϕ. Given u as a solution for which u, u_x, and u_{xx} approach zero as $|x| \to \infty$, show that

$$\int_{-\infty}^{\infty} u(x,t)dx = \text{ const}, \qquad \int_{-\infty}^{\infty} u^2(x,t)dx = \text{ const},$$

for all $t > 0$.

2. In three dimensions assume show that advection should be modeled by the equation

$$u_t + \operatorname{div} \mathbf{c}u = 0,$$

where $\mathbf{c} = \mathbf{c}(\mathbf{x})$ is the velocity of the medium. Given $\mathbf{c}$ as a constant vector, show that $u = f(\mathbf{x} - \mathbf{c}t)$ is a solution for any real-valued differentiable function f.

3. Show that Burgers' equation (1.3.12) can be transformed into the linear diffusion equation (1.3.3) by the *Cole-Hopf transformation*

$$u = -\frac{2Dv_x}{v}.$$

4. Show that the PDE

$$u_t + kuu_x + q(t)u = 0$$

can be reduced to the inviscid Burgers' equation

$$v_s + vv_x = 0$$

using the transformation

$$v = u\exp\int q(t)\,dt, \qquad s = \int k\exp\left(-\int q(y)\,dy\right)\,dt.$$

Show that the same transformation transforms

$$u_t + kuu_x + q(t)u - Du_{xx} = 0$$

into

$$v_s + vv_x - g^{-1}(s)v_{xx} = 0,$$

where $g = (k/D)\exp\left(-\int q(t)\,dt\right)$.

5. By rescaling, show that the porous media equation can be written in the form

$$u_\tau = (u^m)_{\xi\xi} \quad (m > 2)$$

for appropriately chosen dimensionless variables τ, ξ, and u.

6. Show that the nonlinear growth–diffusion equation

$$S_t = k(S^3)_{xx} + aS$$

can be reduced to the porous medium equation by the transformations

$$S = \rho(x,t)e^{at}, \quad \tau = \frac{1}{2a}e^{2at}.$$

7. Consider a porous medium where the fluid is water, and assume that the density ρ is constant. What equation must the pressure p satisfy? Describe the pressure distribution.

8. Nondimensionalize the growth–advection–diffusion equation

$$u_t = Du_{xx} - cu_x + ru.$$

9. The population density $u(x,t)$ of zooplankton in a deep lake varies as a function of depth $x > 0$ and time t ($x = 0$ is the surface). Zooplankton diffuse with diffusion constant D, and buoyancy effects cause them to migrate toward the surface with an advection speed of αg, where g is the acceleration due to gravity. Ignore birth and death rates.

   (a) Find a PDE model for the population density of zooplankton in the lake, along with the the appropriate boundary conditions at $x = 0$ and $x = +\infty$.

   (b) Find the steady-state population density $u = U(x)$ for zooplankton as a function of depth, and sketch its graph.

10. Let $u = u(x, y, t)$ satisfy the following PDE and boundary conditions:

$$\begin{aligned} \varepsilon^2 u_{xx} + u_{yy} &= 0, \quad 0 < y < 1,\, x \in \mathbb{R},\, t > 0, \\ u_y(x, 0, t) &= 0, \quad u_y(x, 1, t) + \varepsilon^2 u_{tt}(x, 1, t) = 0, \\ x \in \mathbb{R},\, t > 0, \end{aligned}$$

where ε is small. Assuming a perturbation expansion

$$u = u_0(x, t) + u_1(x, y, t)\varepsilon^2 + \cdots,$$

show that u_0 satisfies the wave equation.

1.4 Initial and Boundary Value Problems

So far we encountered several types of PDEs governing different types of physical processes:

$$\begin{array}{ll} u_t = Du_{xx} & \text{(diffusion)} \\ u_t = Du_{xx} + f(x,t,u) & \text{(reaction–diffusion)} \\ u_t + cu_x = 0 & \text{(advection)} \\ u_t + cu_x = Du_{xx} & \text{(advection–diffusion)} \\ u_t + uu_x = Du_{xx} & \text{(nonlinear advection–diffusion)} \end{array}$$

As we noted earlier, we are seldom interested in the general solution to a PDE, which contains arbitrary functions. Rather, we are interested in solving the PDE subject to auxiliary conditions such as initial conditions, boundary conditions, or both.

One of the fundamental problems in PDEs is the *pure initial value problem* (or *Cauchy problem*) on $\mathbb{R}$ having the form

$$u_t + F(x, t, u, u_x, u_{xx}) = 0, \quad x \in \mathbb{R}, \quad t > 0, \tag{1.4.1}$$

$$u(x, 0) = u_0(x), \quad x \in \mathbb{R}, \tag{1.4.2}$$

where $u_0(x)$ is a given function. Interpreted physically, the function $u_0(x)$ represents a *signal* at time $t = 0$, and the PDE is the equation that propagates the signal in time. Figure 1.9 depicts this interpretation. In wave propagation problems, the signal is usually called a *wave* or *wave profile*. There are several fundamental questions associated with the pure initial value problem (1.4.1)–(1.4.2).

1. *Existence of Solutions.* Given an initial signal $u_0(x)$ satisfying specified regularity conditions (e.g., continuous, bounded, integrable, or whatever), does a solution $u = u(x,t)$ exist for all $x \in \mathbb{R}$ and $t > 0$? If a solution exists for all $t > 0$, it is called a *global solution*. Sometimes solutions are only *local*; that is, they exist for only up to finite times. For nonlinear hyperbolic problems, for example, a signal can propagate up to a finite time and blowup occurs; that is, the signal experiences a gradient catastrophe where u_x becomes infinite and the solution ceases to be smooth. In other problems, for example in some reaction–diffusion equations, the solution u itself may blow up.

2. *Uniqueness.* If a solution of (1.4.1)–(1.4.2) exists, is the solution unique? For a properly posed physical problem we expect an affirmative answer, and therefore we expect the governing initial value problem, which is regarded as a mathematical model for the physical system, to mirror the properties of the system when considering uniqueness and existence questions.

3. *Continuous Dependence on Data.* Another requirement of a physical problem is that of stability; that is, if the initial condition is changed by only a small amount, the system should behave in nearly the same way. Mathematically, this is translated into the statement that the solution should depend continuously on the initial data. In PDEs, if the initial value problem has a unique solution that depends continuously on the initial conditions, we say that the problem is *well-posed*. A similar statement can be made for boundary value problems. A basic question in PDEs is the problem of well-posedness.

4. *Regularity of Solutions.* If a solution exists, how regular is it? In other words, is it continuous, continuously differentiable, or piecewise smooth?

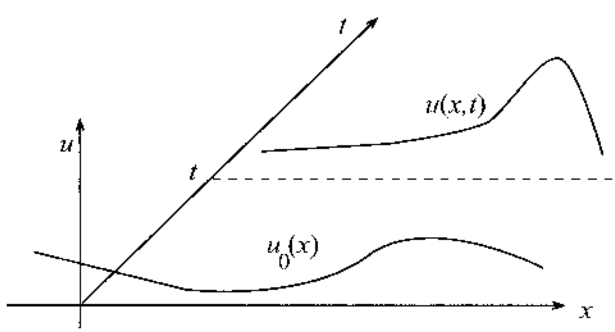

Figure 1.9 Schematic indicating the time evolution, or propagation, of an initial signal or waveform $u_0(x)$.

5. *Asymptotic Behavior.* If an initial signal can be propagated for all times $t > 0$, we may inquire about its asymptotic behavior, or the form of the signal for long times. If the signal decays, for example, what is the decay rate? Does the signal disperse, or does it remain coherent for long times? Does it keep the same shape?

These are a few of the issues in the study of PDEs. The primary issue, however, from the point of view of the applied scientist, may be methods of solution. If a physical problem leads to an initial value problem as a mathematical model, what methods are available or can be developed to obtain a solution, either exact or approximate? Or, if no solution can be obtained (say, other than numerical), what properties can be inferred from the governing PDEs themselves? For example, what is the speed of propagation? Are solutions wave-like, diffusion-like, or dispersive? These questions are addressed in subsequent chapters.

Several examples illustrate the diversity of solutions.

Example. (*Diffusion Equation*) The solution to initial value problem for the diffusion equation

$$u_t = Du_{xx}, \quad x \in \mathbb{R}, \quad t > 0,$$
$$u(x,0) = u_0(x), \quad x \in \mathbb{R},$$

is, as one may verify (e.g., using Fourier transforms), as follows:

$$u(x,t) = \frac{1}{\sqrt{4\pi Dt}} \int_{-\infty}^{\infty} u_0(\xi) e^{-(x-\xi)^2/4Dt}\, d\xi.$$

The solution is valid for all $t > 0$ and $x \in \mathbb{R}$ under rather mild restrictions on the initial signal $u_0(x)$, and the solution has a high degree of smoothness even if the initial data u_0 are discontinuous. Succinctly stated, diffusion smooths out signals. □

Example. (*Advection Equation*) Consider the linear initial value problem for the advection equation

$$u_t + cu_x = 0, \quad x \in \mathbb{R}, \quad t > 0,$$
$$u(x,0) = u_0(x), \quad x \in \mathbb{R},$$

where c is a positive constant. It is easy to check that $u(x,t) = f(x - ct)$ is a solution of the PDE for any differentiable function f. We can apply the initial condition to determine f by writing $u(x,0) = f(x) = u_0(x)$. Therefore the global solution to the initial value problem is

$$u(x,t) = u_0(x - ct) \quad x \in \mathbb{R}, \quad t > 0.$$

Graphically, the solution is the initial signal $u_0(x)$ shifted to the right by the amount ct, as shown in Figures 1.10 and 1.11. Therefore, the initial signal moves forward undistorted in spacetime at speed c. Regarding regularity, even if u_0 is discontinuous, it appears that the solution holds, provided we can make sense of derivatives of discontinuous functions. □

Example. (*Inviscid Burgers' Equation*) A more complicated example is the nonlinear Cauchy problem

$$u_t + uu_x = 0, \quad x \in \mathbb{R}, \quad t > 0,$$
$$u(x,0) = \frac{1}{1+x^2}, \quad x \in \mathbb{R}.$$

In contrast to the two preceding examples, the solution does not exist for all $t > 0$. The initial waveform, in the form of a bell-shaped curve, distorts during propagation and a gradient catastrophe occurs in finite time. The argument we present to show this nonexistence of a global solution is typical of the types of general arguments that are developed later to study nonlinear hyperbolic problems. Assume that this problem has a solution $u = u(x,t)$, and consider the family of curves in xt space defined by the differential equation

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

Denote a curve C in this family by $x = x(t)$. Along this curve we have $du/dt = u_t(x(t),t) + u_x(x(t),t)dx/dt = 0$, and therefore $u =$ constant on C. The curve C must be a straight line because $d^2x/dt^2 = du/dt = 0$. The curve C, which is called a *characteristic curve*, is shown in Figure 1.12 emanating from a point

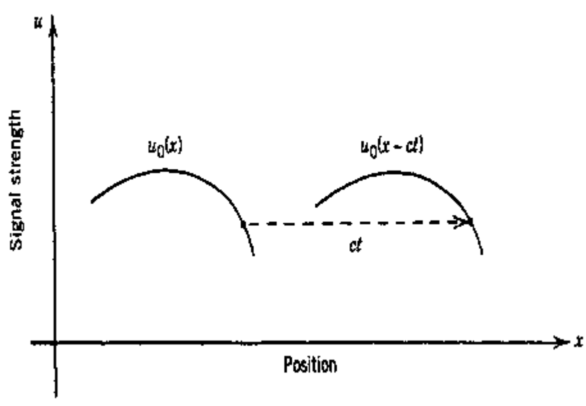

Figure 1.10 Right traveling wave, which represents the solution to the advection equation, shown in xu space.

$(\xi, 0)$ on the initial timeline (x axis) to an arbitrary point (x, t) in spacetime. The equation of C is given by

$$x - \xi = u(\xi, 0)t,$$

where its speed (the reciprocal of its slope), $u(\xi, 0) = (1 + \xi^2)^{-1}$, is determined by the initial condition and the fact that u is constant on C. Now let us determine how the gradient u_x of u evolves along curve C. For simplicity denote $g(t) = u_x(x(t), t)$. Then

$$g'(t) = u_{xx}\frac{dx}{dt} + u_{xt} = (u_t + uu_x)_x - u_x^2 = -u_x^2 = -g(t)^2.$$

The general solution of $g' = -g^2$ is

$$g(t) = \frac{1}{t + c}, \quad c \text{ const.}$$

But $g(0) = 1/c = -2\xi/(1+\xi^2)^2$ is the initial gradient, and therefore g is given by

$$g(t) = \frac{1}{t - (1 + \xi^2)^2/2\xi}.$$

Note that ξ may be chosen positive. Therefore, along the straight line C the gradient u_x becomes infinite at the finite time $t = (1 + \xi^2)^2/2\xi$. Therefore a smooth solution cannot exist for all $t > 0$. □

The nonexistence of a global solution to the initial value problem is a typically nonlinear phenomenon. Because physical processes are often governed by nonlinear equations, we may well ask what happens after the gradient catastrophe. Actually, the distortion of the wave profile and development of an infinite gradient is the witnessing of the formation of a shock wave (i.e., a discontinuous solution that propagates thereafter). However, the idea of a nonsmooth

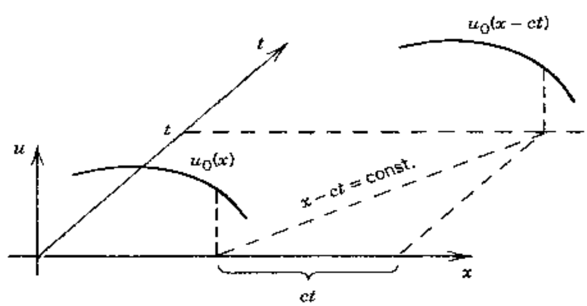

Figure 1.11 Right traveling wave shown in Figure 1.10 represented in xtu-space.

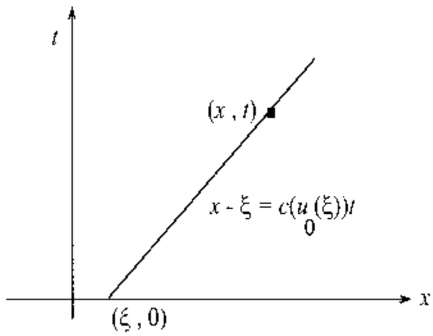

Figure 1.12 Characteristic C emanating from the x axis.

solution to a PDE is a concept that must be formulated carefully, and we carry out this program in Chapter 3.

Another type of problem associated with PDEs is the *signaling problem*. In this case the domain of the problem is the first quadrant $x > 0$, $t > 0$ in spacetime, and initial data are given along the positive x axis; data are also prescribed along the positive t axis as boundary conditions, or signaling data (see Figure 1.13). The form of a *signaling problem* is

$$u_t + F(x, t, u, u_x, u_{xx}) = 0, \quad x > 0, \quad t > 0, \tag{1.4.3}$$

$$u(x, 0) = u_0(x), \quad x > 0, \tag{1.4.4}$$

$$u(0, t) = u_1(t), \quad t > 0, \tag{1.4.5}$$

where $u_0(x)$ is the given initial state and $u_1(t)$ is a specified signal imposed

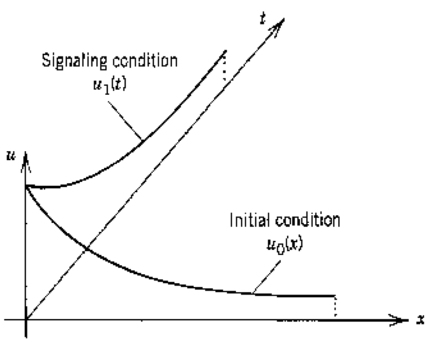

Figure 1.13 Schematic representing a signaling problem where signaling data are prescribed at $x = 0$ along the time axis and initial data are prescribed at $t = 0$ along the spatial axis.

at $x = 0$ for all times $t > 0$. As in the case of the initial value problem, the signaling problem may or may not have a solution that exists for all times t.

In lieu of the condition (1.4.5) given at $x = 0$, one may impose a condition on the derivative u of the form

$$u_x(0, t) = u_2(t), \quad t > 0, \tag{1.4.6}$$

or some combination of u and u_x,

$$u(0, t) + \alpha u_x(0, t) = u_2(t), \quad t > 0.$$

If the PDE is a conservation law and Fick's law $\phi(x, t) = -Du_x(x, t)$ holds, condition (1.4.6) translates into a condition on the flux ϕ. The condition that the flux be zero at $x = 0$ is the physical condition that the amount of u that passes through $x = 0$ is zero; in heat flow problems this condition is called the *insulated boundary condition*. A boundary condition of the type (1.4.5) is called a *Dirichlet condition*, and one of the type (1.4.6) is called a *Neumann condition*. Mixed conditions are called *Robin conditions*.

If the spatial domain is finite, that is, $a \le x \le b$, one may expect to impose boundary data along both $x = a$ and $x = b$, and therefore we consider the *initial boundary value problem*

$$u_t + F(x, t, u, u_x, u_{xx}) = 0, \quad a < x < b, \quad t > 0, \tag{1.4.7}$$

$$u(x, 0) = u_0(x), \quad a < x < b, \tag{1.4.8}$$

$$u(a, t) = u_1(t), \quad u(b, 0) = u_2(t), \quad t > 0, \tag{1.4.9}$$

where u_0, u_1, and u_2 are given functions. If (1.4.7) is the diffusion equation, this problem has a solution under mild restrictions on the data. However, if we consider the advection equation, the problem seldom has a solution for arbitrary boundary data, as the following example shows.

Example. Consider the initial boundary value problem

$$u_t + cu_x = 0, \quad 0 < x < 1, \quad t > 0,$$

with initial and boundary conditions given by (1.4.8) and (1.4.9) with $a = 0$ and $b = 1$. We have noted already that the general solution of the advection equation is $u(x, t) = f(x - ct)$, for an arbitrary function f. Consequently, u must be constant on the straight lines $x - ct =$ constant (see Figure 1.14). Clearly, therefore, data cannot be independently specified along the boundary $x = b$. In this case, the initial data along $0 < x < 1$ are carried along the straight lines to the segment A on $x = 1$; the boundary data along $x = 0$ is carried to the segment B on $x = 1$. Thus $u(1, t)$ cannot be specified arbitrarily. □

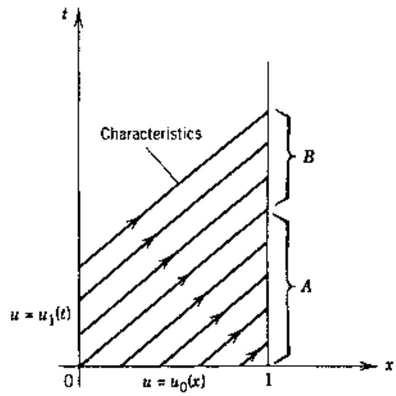

Figure 1.14 Right-moving characteristics carrying left boundary and initial data into the region of interest. Arbitrary data may not be prescribed along the right boundary $x = 1$ labeled A and B.

The preceding example shows that we must be careful in defining a well-posed problem. Mathematically correct initial conditions and boundary conditions are associated with types of PDEs (hyperbolic, parabolic, elliptic), as well as their order. Conditions that ensure well-posedness are often suggested by the underlying physical problem.

Example. (*Wave Equation*) The wave equation

$$u_{tt} - c^2 u_{xx} = 0, \quad x \in \mathbb{R}, \quad t > 0, \tag{1.4.10}$$

is second-order in t, and therefore it does not fit into the category of equations defined by (1.4.1). Because it is second order in t, we are guided by our experiences with ordinary differential equations to impose two conditions at $t = 0$, a condition on u and a condition on u_t. Therefore, the pure initial value problem for the wave equation consists of (1.4.10) subject to the initial conditions

$$u(x,0) = u_0(x), \quad u_t(x,0) = u_1(x), \quad x \in \mathbb{R}, \tag{1.4.11}$$

where u_0 and u_1 are given functions. If $u_0 \in C^2(\mathbb{R})$ and $u_1 \in C^1(\mathbb{R})$, the unique, global solution to (1.4.10)–(1.4.11) is given by D'Alembert's formula (1.1.6) with $f = u_0$ and $g = u_1$. $\square$

Example. The initial value problem

$$u_{tt} + u_{xx} = 0, \quad x \in \mathbb{R}, \quad t > 0, \tag{1.4.12}$$

$$u(x,0) = u_0(x), \quad u_t(x,0) = u_1(x), \quad x \in \mathbb{R} \tag{1.4.13}$$

is not well-posed because small changes, or perturbations, in the initial data can lead to arbitrarily large changes in the solution (see Exercise 5). Equation (1.4.12) is Laplace's equation (with variables x and t, rather than the usual x and y), which is elliptic. In general, initial conditions are not correct for elliptic equations, which are naturally associated with equilibrium phenomena and boundary data. □

To summarize, the auxiliary conditions imposed on a PDE must be considered carefully. In the sequel, as the subject is developed, the reader should become aware of which conditions go with which equations in order to ensure, in the end, a well-formulated problem.

EXERCISES

1. Solve the signaling problem

$$u_t + cu_x = 0, \quad x > 0, \quad t > 0,$$
$$u(x, 0) = 1, \quad x > 0,$$
$$u(0, t) = \frac{1 + t^2}{1 + 2t^2}, \quad t > 0,$$

using the fact that u must be constant on the curves $x = ct + \xi$, where ξ is constant. *Hint*: Treat the regions $x > ct$ and $x < ct$ separately.

2. Obtain the solution to the initial value problem

$$u_t = u_{xx}, \qquad x \in \mathbb{R}, \quad t > 0,$$
$$u(x, 0) = u_0 \quad \text{if} \quad |x| < L \quad \text{and} \quad u(x, 0) = 0 \quad \text{if } |x| > L,$$

where u_0 is a constant, in the form

$$u(x, t) = -\frac{u_0}{2}\left[\operatorname{erf}\left(\frac{x - L}{\sqrt{4t}}\right) - \operatorname{erf}\left(\frac{x + L}{\sqrt{4t}}\right)\right],$$

where erf is the *error function*

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

Show that for x fixed and for large t, we obtain

$$u(x, t) \sim \frac{u_0 L}{\sqrt{\pi t}}.$$

3. Find a formula for the solution to the initial boundary value problem

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

4. Let $u \in C^1$ be a solution to the Cauchy problem

$$u_t + Q(u)_x = 0, \quad x \in \mathbb{R}, \quad t > 0,$$
$$u(x,0) = u_0(x), \quad x \in \mathbb{R},$$

where $Q(0) = 0$, $Q''(u) > 0$, and u_0 is integrable on $\mathbb{R}$ with $u_0(x) = 0$ for $x < x_0$ (for some x_0), and $u_0(x) > 0$ otherwise. Define

$$U(x,t) = \int_{-\infty}^{x} u(s,t)\,ds$$

(a) Show that U satisfies the equation $U_t + Q(U_x) = 0$.

(b) Prove that $Q(u) \geq Q(v) + c(v)(u - v)$, where $c(u) = Q'(u)$.

(c) Prove that $U_t + c(v)U_x \leq c(v)v - Q(v)$.

(d) Show that $U(x,t) \leq U(\xi, 0) + t[c(v)v - Q(v)]$, where ξ is the point where the line defined by $dx/dt = c(v)$ through the point (x,t) intersects the x axis.

(e) In the inequalities above, show that equality holds if $v = u$.

(f) Give a geometric–physical interpretation of the results above.

5. By considering solutions $u_n(x,t) = n^{-1} \cos nx \cosh nt$, show that the initial value problem (1.4.12)–(1.4.13) for Laplace's equation is not well-posed.

6. Consider the initial value problem for the *backward* diffusion equation:

$$u_t + u_{xx} = 0, \quad x \in \mathbb{R}, \quad t > 0,$$
$$u(x,0) = 1, \quad x \in \mathbb{R}.$$

Show that the solution does not depend continuously on the initial condition by considering the functions

$$u_n(x,t) = 1 + \frac{1}{n}\exp(n^2 t)\sin nx.$$

1.5 Waves

One of the cornerstones of PDEs is wave propagation. A *wave* is a recognizable signal that is transferred from one part of the medium to another part with a recognizable speed of propagation. Energy is often transferred as a wave propagates, but matter may not be. There is hardly any area of science or engineering where wave phenomena are not a critical part of the subject. We mention a few areas where wave propagation is of fundamental importance.

Fluid mechanics (water waves, aerodynamics, meteorology, traffic flow)
Acoustics (sound waves in air and liquids)
Elasticity (stress waves, earthquakes)
Physics (optics, electromagnetic waves, quantum mechanics)
Biology (spread of diseases, population dispersal, nerve signal transmission)
Porous media (groundwater dynamics, contaminant migration)
Chemistry (combustion and detonation waves)

1.5.1 Traveling Waves

The simplest form of a mathematical wave is a function of the form

$$u(x,t) = f(x - ct). \tag{1.5.1}$$

We interpret the density u as the strength of the signal. At $t = 0$ the wave has the form $f(x)$, which is the initial wave profile. Then $f(x - ct)$ represents the profile at time t, which is just the initial profile translated to the right ct spatial units. The constant c represents the speed of the wave. Evidently, (1.5.1) represents a right traveling wave of speed c. Similarly, $u(x,t) = f(x+ct)$ represents a left traveling wave of speed c. These types of waves propagate undistorted along the straight lines $x - ct = \text{const.}$ (or $x + ct = \text{const.}$) in spacetime.

A key question is whether a given PDE can propagate such a traveling wave or, in different words, whether a traveling wave solution (TWS) exists for a given PDE. This question is generally posed without regard to initial conditions (time is usually regarded as varying from $-\infty$ to $+\infty$), so that the wave is assumed to have existed for all times. However, boundary conditions of the form

$$u(-\infty, t) = \text{const.}, \quad u(+\infty, t) = \text{const.} \tag{1.5.2}$$

may be imposed. A *wavefront solution* is a TWS of the form $u(x,t) = f(x - ct)$ (or $f(x + ct)$) subject to the conditions (1.5.2) of constancy at plus and minus infinity (not necessarily the same constant); at present, the function f

is assumed to have the requisite degree of smoothness defined by the PDE ($C^1(\mathbb{R}), C^2(\mathbb{R}), \ldots$). Figure 1.15 shows a typical wavefront. If u approaches the same constant at both plus and minus infinity, the wavefront solution is called a *pulse*.

The key computational device is to substitute the form $u(x,t) = f(x - ct)$ into the PDE and observe whether it transforms the PDE into an ODE for the unknown wave profile f, which is a function of a single variable $z = x - ct$, interpreted as a moving coordinate. The speed c of the wave is not known. To carry out the substitution it is necessary to calculate how the derivatives of u transform. By the chain rule, we have

$$\begin{aligned} u_t &= f'(z)z_t = -cf'(z), \\ u_x &= f'(z)z_x = f'(z). \end{aligned}$$

We easily find $u_{tt} = c^2 f''(z)$, $u_{xx} = f''(z)$, and so on for higher derivatives. Therefore, the general equation

$$u_t = G(u, u_x, u_{xx})$$

transforms into

$$-cf' = G(f, f', f''), \quad z \in \mathbb{R},$$

which is an ODE for f and the unknown wave speed c. If boundary conditions

$$f(-\infty) = f_0, \quad f(+\infty) = f_1$$

are imposed, we interpret this problem as a nonlinear eigenvalue problem for f, with c as an eigenvalue.

Example. (*Advection Equation*) We have already observed that

$$u(x,t) = f(x - ct),$$

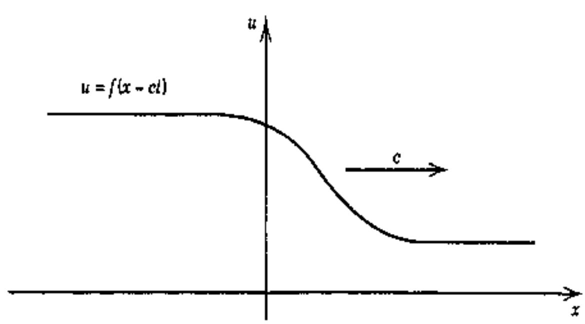

Figure 1.15 A right traveling wavefront with speed c. A wavefront solution is a TWS with constant states at $\pm\infty$.

where f is a differentiable function, is a solution to the PDE

$$u_t + cu_x = 0. \tag{1.5.3}$$

Therefore, the advection equation (1.5.3) admits wavefront solutions, and it is the simplest example of a wave equation. □

Example. (*Wave Equation*) The wave equation

$$u_{tt} = c^2 u_{xx}$$

has solutions that are the superposition of right and left traveling waves (see Section 1.1). □

Example. (*Diffusion Equation*) The diffusion equation cannot propagate nonconstant wavefronts. To verify this fact, we substitute $u = f(z)$, where $z = x - ct$, into the diffusion equation $u_t = Du_{xx}$ to obtain the following ordinary differential equation for the wave profile f:

$$-cf'(z) = Df''(z), \quad z \in \mathbb{R}$$

This linear differential equation has the general solution

$$f(z) = a + be^{-cz/D},$$

where a and b are arbitrary constants. The only possibility for f to be constant at both plus and minus infinity is to require $b = 0$. Thus, there are traveling wave solutions, but no nonconstant wavefront solutions to the diffusion equation. □

Example. (*Korteweg–deVries Equation*) The KdV equation is

$$u_t + uu_x + u_{xxx} = 0. \tag{1.5.4}$$

This nonlinear equation admits traveling wave solutions of different types. One particular type is the soliton, or solitary wave, which is now derived. Assume that $u = f(z)$, where $z = x - ct$. Substituting into (1.5.4) gives

$$-cf' + ff' + f''' = 0, \tag{1.5.5}$$

where the prime denotes d/dz. Integrating (1.5.5) 2 times yields, after rearrangement

$$\frac{df}{dz} = \frac{1}{\sqrt{3}}(-f^3 + 3cf^2 + 6af + 6b)^{1/2}, \tag{1.5.6}$$

where a and b are constants of integration. Here we took the plus sign on the square root; later we observe that this can be done without loss of generality.

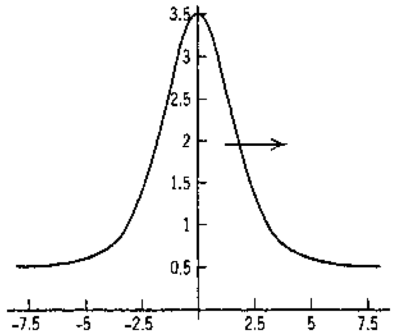

Figure 1.16 Plot of the soliton (1.5.8) at $t = 0$ with $\gamma = 0.5$ and $\alpha - \gamma = 3.0$. A soliton is an example of a pulse.

The first-order differential equation (1.5.6) for the form of the wave f is separable. However, the expression under the radical on the right side is a cubic in f, and therefore different cases must be considered depending on the number of real roots, double roots, and so on. We examine only one case, namely, when the cubic has three real roots, where one is a double root. So let us assume that

$$-f^3 + 3cf^2 + 6af + 6b = (f-\gamma)^2(\alpha - f), \quad 0 < \gamma < \alpha.$$

Then the differential equation (1.5.6) becomes

$$(f-\gamma)^{-1}(\alpha - f)^{-1/2}\,df = \frac{1}{\sqrt{3}}\,dz. \tag{1.5.7}$$

The substitution $f = \gamma + (\alpha - \gamma)\text{sech}^2 w$ easily reduces (1.5.7) to

$$-2(\alpha-\gamma)^{-1/2}\,dw = \frac{1}{\sqrt{3}}dz,$$

which integrates to $w = \sqrt{(\alpha-\gamma)/12}\,z$. Consequently, traveling waves in the case we are considering have the form

$$u(x,t) = \gamma + (\alpha - \gamma)\text{sech}^2[\kappa(x - ct)], \quad \kappa = \left(\frac{\alpha-\gamma}{12}\right)^{1/2}. \tag{1.5.8}$$

A graph of the waveform, which is a pulse, is shown in Figure 1.16. It is instructive to write the roots α and γ in terms of the original parameters. To this end we have

$$\begin{aligned} -f^3 + 3cf^2 + 6af + 6b &= (f-\gamma)^2(\alpha - f) \\ &= -f^3 + (\alpha + 2\gamma)f^2 - (2\gamma\alpha + \gamma^2)f + \gamma^2\alpha. \end{aligned}$$

Therefore, the wave speed c is given by

$$c = \frac{\alpha + 2\gamma}{3} = \frac{\alpha - \gamma}{3} + \gamma. \tag{1.5.9}$$

The speed of the wave relative to the state γ ahead of the wave is proportional to the amplitude $A = \alpha - \gamma$. Both the amplitude and the width of the wave depend on the wave speed c. Thus, the taller and wider the wave, the faster it moves. For linear waves, say, governed by the wave equation, the speed of propagation is independent of the amplitude of the wave. Such a waveform (1.5.8) is known as a *soliton*, or *solitary wave*, and many of the important equations of mathematical physics exhibit soliton-type solutions (e.g., the Boussinesq equation, the Sine–Gordon equation, the Born–Infeld equation, and nonlinear Schrödinger equations). In applications, the value of such solutions is that if a pulse or signal travels as a soliton, the information contained in the pulse can be carried over long distances with no distortion or loss of intensity. Solitons occur in fluid mechanics, nonlinear optics, and other nonlinear phenomena. □

There is an interesting history of solitary waves. In 1836 John Scott Russell, a Scottish engineer, observed such a wave moving along a canal. He followed it on horseback, noting that it propagated a long distance without changing form. Many at the time doubted his observation and thought that such waves could not exist. However, Boussinesq, in 1872, showed that in a special limit of the equations of fluid flow such water waves can exist; Korteweg and deVries obtained (1.5.4) in the 1890. It wasn't until the 1960s that physicists Kruskal and Zabusky obtained the equation as a continuum limit of a model of nonlinear spring mass chains, and they coined the term *soliton*.

The interaction between two solitons is both interesting and unexpected. If two solitons are moving to the right and the one behind represents a stronger signal (therefore moving faster) than the one ahead, the large one overtakes the slower, smaller wave and a complicated nonlinear interaction occurs, after which both waves return to their original shape. The only change is a phase shift in the two waves. The initial value problem for the KdV equation, with $u(x,0) = u_0(x)$, $x \in \mathbb{R}$, also has interesting behavior. If the initial profile u_0 approaches zero fast enough at $x = \pm\infty$, then, over time, the solution divides into a finite number of solitons moving to the right at their respective speeds (smaller to larger), plus a small dispersive disturbance moving to $-\infty$. □

The problem of determining whether a given PDE admits wavefront solutions is fundamental, and it occupies much attention in subsequent chapters. The search for these types of solutions is one of the basic methods in the analysis of nonlinear problems.

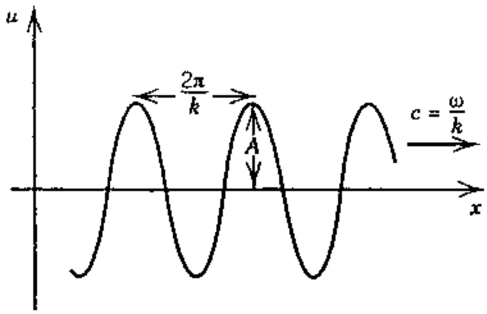

Figure 1.17 A plane wave of amplitude A, wave number k, and frequency ω.

1.5.2 Plane Waves

Another type of wave of interest is a *plane wave*, or *wave train*. These waves are traveling, periodic waves of the form

$$u(x,t) = A\cos(kx - \omega t), \tag{1.5.10}$$

where A is the amplitude of the wave, k the wave number, and ω the frequency. The *wave number* k is a measure of the number of spatial oscillations (per 2π units) observed at a fixed time, and the *frequency* ω is a measure of the number of oscillations in time (per 2π units) observed at a fixed spatial location. The number $\lambda = 2\pi/k$ is the *wavelength*, and $P = 2\pi/\omega$ is the *period*. The wavelength measures the distances between successive crests, and the period measures the time for an observer located at a fixed position x to see a repeat pattern (see Figure 1.17). Note that (1.5.10) may be written

$$u(x,t) = A\cos k\left(x - \frac{\omega}{k}t\right),$$

so (1.5.10) represents a traveling wave moving to the right with speed $c = \omega k$. This number is called the *phase velocity*, and it is the speed that one would have to move to remain on the crest of the wave. For calculations the complex form

$$u(x,t) = Ae^{i(kx-\omega t)} \tag{1.5.11}$$

is preferred because differentiations with the exponential function are simpler. After the calculations are completed, one can use Euler's formula $\exp(i\theta) = \cos\theta + i\sin\theta$ and take real or imaginary parts to recover a real solution. The technique of searching for plane wave solutions is applicable to linear equations, but a modification applies to some nonlinear problems. Finding plane wave solutions is a powerful method for determining properties of linear problems.

Example. (*Diffusion Equation*) We seek solutions of the form of equation (1.5.11) to the diffusion equation

$$u_t = Du_{xx}. \tag{1.5.12}$$

Substituting (1.5.11) into (1.5.12) gives

$$\omega = -iDk^2. \tag{1.5.13}$$

Equation (1.5.13) is a condition, called a *dispersion relation*, between the frequency ω and the wave number k that must be satisfied for (1.5.12) to admit plane wave solutions. Thus we have determined a class of solutions

$$u(x,t;k) = Ae^{-Dk^2t}e^{ikx}, \quad k \in \mathbb{R} \tag{1.5.14}$$

for the diffusion equation that depends on an arbitrary parameter k, the wave number. The factor $\exp(ikx)$ represents a spatial oscillation, and the factor $A\exp(-Dk^2t)$ represents a decaying amplitude. Note that the rate of decay depends on the wave number k; waves of shorter wavelength decay more rapidly than do waves of longer wavelength. □

Example. (*Wave Equation*) Substituting (1.5.11) into the wave equation

$$u_{tt} = c^2u_{xx} \tag{1.5.15}$$

forces

$$\omega = \pm ck. \tag{1.5.16}$$

Therefore, the wave equation admits solutions of the form

$$u(x,t) = Ae^{ik(x \mp ct)},$$

which are right and left sinusoidal traveling waves of speed c. □

The technique of looking for solutions of the form (1.5.11) for linear, homogeneous PDEs of the form

$$Lu = 0, \tag{1.5.17}$$

where L is a linear constant-coefficient operator, always leads to a relation connecting the frequency ω and the wave number k of the form

$$G(k, \omega) = 0. \tag{1.5.18}$$

This condition is called the *dispersion relation* corresponding to the PDE (1.5.17), and it characterizes plane wave solutions entirely. Generally, linear PDEs (1.5.17) can be classified according to their dispersion relation in the following way. Assume that (1.5.18) may be solved for ω in the form

$$\omega = \omega(k). \tag{1.5.19}$$

We say that the PDE is *dispersive* if $\omega(k)$ is real and if $\omega''(k)\neq 0$. When $\omega(k)$ is complex, the PDE is *diffusive*. The diffusion equation, from the example above, has dispersion relation $\omega = -ik^2D$, which is complex; thus the diffusion equation is classified as diffusive. The classical wave equation (1.5.15) is neither diffusive nor dispersive under the foregoing classification. Even though its solutions are wave-like, it is not classified as dispersive since its dispersion relation (see 1.5.16) satisfies $\omega'' = 0$. For the wave equation, the speed of propagation is c and is independent of the wave number k. For dispersive equations, the speed of propagation, or the phase velocity ω/k, depends on the wave number (or equivalently, the wavelength). The wave equation is generally regarded as the prototype of a hyperbolic equation, and the term *dispersive* is reserved for equations where the phase velocity depends on k. We caution the reader that the term *dispersion* is often used in a diffusion context (e.g., a group of animals dispersing), and it is important to be aware of this.

Example. (*Schrödinger Equation*) In quantum mechanics, the Schrödinger equation for a free particle, under appropriate scalings, is

$$u_t = iu_{xx}.$$

It is easy to see that the dispersion relation is $\omega = k^2$, so that the Schrödinger equation is dispersive. The Schrödinger equation is neither parabolic nor hyperbolic. □

1.5.3 Plane Waves and Transforms

Using a plane wave assumption, we already constructed the class of solutions $u(x,t;k) = Ae^{-Dk^2t}e^{ikx}$, $k \in \mathbb{R}$, to the diffusion equation

$$u_t = Du_{xx}, \quad x \in \mathbb{R},\ t > 0.$$

Formally, superimposing these solutions, we obtain

$$u(x,t) = \int_{\mathbb{R}} A(k)e^{-k^2Dt}e^{ikx}\,dk, \tag{1.5.20}$$

which can be verified to be a solution to the diffusion equation, provided that $A(k)$ is a well-behaved function (e.g., continuous, bounded, and integrable on $\mathbb{R}$). Having (1.5.20) as a solution opens up the possibility of selecting the function $A(k)$ so that other conditions (e.g., an initial condition) can be met. Therefore, let us impose the initial condition

$$u(x,0) = u_0(x), \quad x \in \mathbb{R},$$

to the diffusion equation. From (1.5.20) it follows that

$$u_0(x) = \int_{\mathbb{R}} A(k)e^{ikx}\,dk. \tag{1.5.21}$$

We recognize $u_0(x)$ as the Fourier transform of $A(k)$, and therefore $A(k)$ must be the inverse Fourier transform of the function $u_0(x)$. We remind the reader of these facts by recalling the *Fourier integral theorem* [see, e.g., Stakgold 1998].

Theorem. Let f be a continuous, bounded, integrable function on $\mathbb{R}$, and let

$$\hat{f}(k) = \int_{\mathbb{R}} f(x)e^{ikx}\,dx$$

be the Fourier transform of f. Then, for all $x \in \mathbb{R}$, we have

$$f(x) = \frac{1}{2\pi}\int_{\mathbb{R}} \hat{f}(k)e^{-ikx}\,dk.$$

Applying this theorem (the variables k and x have been interchanged) to 1.5.21) leads us to conclude that

$$A(k) = \frac{1}{2\pi}\int_{\mathbb{R}} u_0(x)e^{-ikx}\,dx.$$

Consequently, we have obtained the solution to the Cauchy problem for the diffusion equation

$$u_t - Du_{xx} = 0, \quad x \in \mathbb{R}, \quad t > 0, \tag{1.5.22}$$

$$u(x,0) = u_0(x), \quad x \in \mathbb{R}, \tag{1.5.23}$$

in the form

$$u(x,t) = \frac{1}{2\pi}\int_{\mathbb{R}} \left(\int_{\mathbb{R}} u_0(\xi)e^{-ik\xi}\,d\xi\right) e^{-k^2Dt}e^{ikx}\,dk. \tag{1.5.24}$$

Interchanging the order of integration allows us to formally write the solution as

$$u(x,t) = \frac{1}{2\pi}\int_{\mathbb{R}} u_0(\xi)\left(\int_{\mathbb{R}} e^{ik(x-\xi)-k^2Dt}\,dk\right) d\xi. \tag{1.5.25}$$

The inner integral may be calculated outright by noting

$$\int_{\mathbb{R}} e^{ik(x-\xi)-k^2Dt}\,dk = 2\int_0^\infty e^{-k^2Dt}\cos(k(x-\xi))\,dk$$
$$= \sqrt{\frac{\pi}{Dt}}e^{-(x-\xi)^2/4Dt}.$$

Therefore, the solution to the initial value problem (1.5.22) (1.5.23) for the diffusion equation is given by

$$u(x,t) = \frac{1}{\sqrt{4\pi Dt}} \int_{\mathbb{R}} u_0(\xi) e^{-(x-\xi)^2/4Dt}\, d\xi. \tag{1.5.26}$$

We derived this solution formally, but we have not proved that it is indeed a solution. A proof would consist of a rigorous argument that (1.5.25) satisfies the diffusion equation (1.5.22) and the initial condition (1.5.23). To show well posedness, a uniqueness argument would have to be supplied as well as a proof that the solution is stable to small perturbations of the initial data. We shall not carry out these arguments here, but rather, refer the reader to the references for the details.

This method (finding plane wave solutions followed by superposition and use of the Fourier integral theorem) is equivalent to the classical Fourier transform method learned in elementary courses where one takes a Fourier transform of the PDE and initial condition to reduce the problem to an ordinary differential equation in the transform domain, which is then solved. Then the inverse Fourier transform is applied to return the solution in the original domain. This method is generally applicable to the pure initial value problems on the real line for linear equations with constant coefficients.

Fourier integral expressions of the type obtained above can be approximated for large times by the method of stationary phase [e.g., see Bhatnagar 1979].

1.5.4 Nonlinear Dispersion

For nonlinear equations we do not expect plane wave solutions of the form (1.5.16), and therefore a dispersion relation will not exist as it does for linear equations. Moreover, superposition for nonlinear equations is invalid. However, in some nonlinear problems, there may exist traveling periodic wave trains of the form

$$u(x,t) = U(\theta), \quad \theta = kx - \omega t, \tag{1.5.27}$$

where U is a periodic function. For example, consider the nonlinear PDE

$$u_{tt} - u_{xx} + f'(u) = 0, \tag{1.5.28}$$

where $f(u)$ is some function of u, yet to be specified. If (1.5.27) is substituted into (1.5.28), we obtain the ordinary differential equation

$$(\omega^2 - k^2)U_{\theta\theta} + f'(U) = 0.$$

Here we are using subscripts θ to denote ordinary derivatives of U with respect

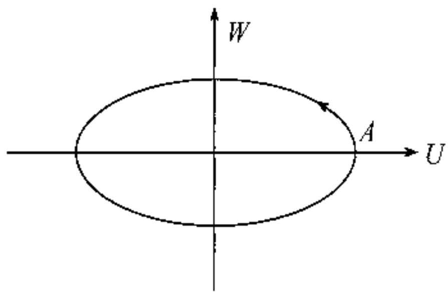

Figure 1.18 Phase space orbit representing the periodic solution (1.5.31) with amplitude A.

θ. Multiplication by U_θ and subsequent integration gives

$$\tfrac{1}{2}(\omega^2 - k^2)U_\theta^2 + f(U) = A, \tag{1.5.29}$$

where A is a constant of integration. The goal is to determine U as a periodic function of θ. Equation (1.5.29) has the same form as an energy conservation law, where f is a potential function, which suggests introducing the variable W defined by $W = U_\theta$. Then

$$W^2 = \frac{2}{\omega^2 - k^2}[A - f(U)], \tag{1.5.30}$$

which are the integral curves. Assuming that $\omega^2 > k^2$, we have

$$W = \pm\sqrt{\frac{2}{\omega^2 - k^2}}\sqrt{A - f(U)}, \tag{1.5.31}$$

which defines a locus of points in the UW plane. For example, let us choose $f(U) = U^4$. Then the locus (1.5.31) is a closed path, representing a periodic solution of (1.5.29), as shown in Figure 1.18. Notice that A, which has been taken positive, is the amplitude of the oscillation represented by the closed path. To find U as a function of θ in the case $f(U) = U^4$ we write (1.5.29) as

$$\theta = \pm\sqrt{\frac{\omega^2 - k^2}{2}}\int_0^U \frac{1}{\sqrt{A - s^4}}\,ds. \tag{1.5.32}$$

This formula defines the periodic function $U = U(\theta)$ implicitly; in this case $U(\theta)$ can be determined explicitly as an elliptic function, and we leave this as an exercise. The period of the oscillation can be determined by integrating over one-quarter period in the integral on the right side of (1.5.32), taking care to choose the appropriate sign. If P denotes the period, we obtain on integration an equation of the form

$$P = P(\omega, k, A). \tag{1.5.33}$$

In other words, the frequency ω will depend on the amplitude A as well as the wave number k. Consequently, the wave speed $c = \omega/k$ will be amplitude-dependent. This amplitude dependence in the nonlinear dispersion relation (1.5.33) is one important distinguishing aspect of nonlinear phenomena.

These calculations can be carried out for nonlinear equations of the form (1.5.28) for various potential functions $f(u)$. Periodic solutions are obtained when U oscillates between two simple zeros of $A - f(U)$. A thorough discussion of nonlinear dispersion can be found in Whitham (1974).

EXERCISES

1. Find the dispersion relation for the advection diffusion equation

$$u_t + au_x = Du_{xx}, \quad (a, D > 0),$$

and show that it is diffusive. Use superposition and the Fourier integral theorem to find an integral representation of the solution of the initial value problem for this equation. *Hint*: You will need the Fourier transform of $e^{iax}f(x)$.

2. Consider the KdV equation in the form $u_t - 6uu_x + u_{xxx} = 0$, $x \in \mathbb{R}$, $t > 0$. Let $u = u(x,t)$ be a solution that decays, along with its derivatives, very rapidly to zero as $|x| \to \infty$. Show that $\int_{\mathbb{R}} u\,dx$ and $\int_{\mathbb{R}} u^2\,dx$ are both constant in time.

3. Examine the form of traveling wave solutions of the KdV equation (1.5.4) in the case that the cubic expression on the right side of (1.5.6) has a triple real root.

4. Determine and sketch traveling wave solutions of the equation

$$u_{tt} - u_{xx} = -\sin u$$

in the form

$$u(x,t) = 4\arctan\left\{\exp \pm \left[\frac{x - ct}{\sqrt{1-c^2}}\right]\right\}, \quad 0 < c < 1.$$

5. For the following PDEs find the dispersion relation and classify the equations as diffusive, dispersive, or neither:

   (a) $u_{tt} + a^2 u_{xxxx} = 0$ (beam equation)

   (b) $u_t + u_{xxx} = 0$ (dispersive wave equation)

   (c) $u_t + au_x + bu_{xxx} = 0$ (linearized KdV equation)

   (d) $u_{tt} - c^2 u_{xx} + bu = 0$ (Klein—Gordon equation)

6. Consider the heat equation in an infinite domain $x > 0$ where the boundary condition at $x = 0$ is a periodic function over all time:

$$\begin{aligned} u_t &= ku_{xx}, \quad x > 0, \ t \in \mathbb{R}, \\ u(0,t) &= T_0 + Ae^{i\omega t}, \quad t \in \mathbb{R}, \end{aligned}$$

where T_0, A, and ω of constants, and where $u = u(x,t)$ is the temperature. This problem models temperatures in the ground subject to surface periodic temperatures. Find $u(x,t)$ and determine the amplitude and phase shift, relative to the values at the surface, at a depth x. Answer the same questions if instead the flux is imposed at the boundary:

$$-Ku_x(0,t) = Ae^{i\omega t}, \quad t \in \mathbb{R}.$$

7. By superimposing plane wave solutions to the dispersive wave equation $u_t + u_{xxx} = 0$, find an integral representation of the solution to the Cauchy problem and write your answer in terms of the Airy function

$$\mathrm{Ai}(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \cos\left(\frac{z^3}{3} + xz\right) dz.$$

8. Find solutions of the outgoing signaling problem for the wave equation:

$$\begin{gathered} u_{tt} - c^2 u_{xx} = 0, \quad x > 0, \ t \in \mathbb{R}, \\ u_x(0,t) = s(t), \quad t \in \mathbb{R}. \end{gathered}$$

9. In equation (1.5.28) take $f(u) = u^2/2$ and determine periodic solutions of the form $u = U(\theta)$, where $\theta = kx - \omega t$. What is the period of oscillation? Does ω depend on the amplitude of oscillation in this case?

10. In equation (1.5.28) assume that the potential function $f(u)$ has the expansion

$$f(u) = \frac{u^2}{2} + \sigma u^4 + \cdots,$$

where σ is a small known parameter. Assuming that the amplitude is small, show that periodic wave trains are given by

$$U(\theta) = a\cos\theta + \tfrac{1}{8}\sigma a^3 \cos 3\theta + \cdots,$$

where the frequency and amplitude are

$$\begin{aligned} \omega^2 &= 1 + k^2 + 3\sigma a^2 + \cdots, \\ A &= \frac{a^2}{2} + \frac{9}{8}\sigma a^4 + \cdots. \end{aligned}$$

11. The *nonlinear Schrödinger equation* occurs in the description of water waves, nonlinear optics, and plasma physics. It is given by

$$iu_t + u_{xx} + \gamma|u^2|u = 0, \tag{1.5.34}$$

where $\gamma > 0$ and $u = u(x,t)$ is complex-valued.

(a) If $u = U(z)e^{i(kz-\omega t)}$, where $z = x - ct$, show that

$$\frac{d^2U}{dz^2} + i(2k-c)\frac{dU}{dz} + (kx - \omega + k^2)U + \gamma U^3 = 0.$$

(b) If $c = 2k$, show that

$$\left(\frac{dU}{dz}\right)^2 = aU^2 - \frac{\gamma}{2}U^4 + C$$

for some appropriately chosen constant a, where C is an arbitrary constant.

(c) Taking $C = 0$ and $a > 0$, show that

$$U(z) = \sqrt{\frac{2a}{\gamma}}\operatorname{sech}(\sqrt{a}(z - ct)),$$

and comment on the properties of this solution.

12. Let $F = F(u)$ be a smooth function and suppose

$$u_{tt} - u_{xx} = F'(u).$$

Assuming that u and its partial derivative u_x both go to zero as $|x| \to \infty$, show that

$$\int_{\mathbb{R}} \left(\tfrac{1}{2}u_t^2 + \tfrac{1}{2}u_x^2 + F(u)\right) = \text{ const.}$$

13. Show that solutions of the nonlinear Schrödinger equation (1.5.34) with $\gamma = 1$ have the properties

$$\int_{\mathbb{R}} |u|^2 dx = \text{ const}, \qquad \int_{\mathbb{R}} \left(u_x^2 - \tfrac{1}{2}|u|^4\right) dx = \text{ const},$$

provided u and its derivatives approach zero sufficiently fast as $|x| \to 0$.

Reference Notes. Partial differential equations have a long history and a correspondingly vast literature. The early developments in PDEs were in the post calculus years of the early 1700s and involved the geometry of surfaces. However, it soon became clear that PDEs were models of physical phenomena like fluid flow, vibrating strings, heat conduction, and so on. Euler, Bernoulli

D'Alembert, Laplace, Lagrange, Fourier, Cauchy, and others developed many of the basic ideas in the linear theory and its applications. It wasn't until the last half of the twentieth century, as part of the general interest in nonlinear science and advancement of computation, that nonlinear PDEs came to the forefront, particularly in their role in wave propagation and diffusion.

It is impractical to cite more than just a few key references. There are books at all levels and several research journals, in both mathematics and the pure and applied sciences, that can be consulted for an entry point to the literature. Here we reference only a few of the texts and cite only articles that are relevant to the particular topic under discussion.

Elementary, entry-level texts focus almost exclusively on linear problems, emphasizing Fourier series, integral transforms, and boundary value problems. The long-time standard has been Churchill (1969), which remains an excellent introduction to Fourier series and boundary value problems. There are many more recent introductory texts, too many to cite. We mention Strauss (1992), an outstanding treatment, and the author's text (Logan 2004), which is a brief introduction. More advanced texts include the classic by John (1982), as well as Evans (1998), McOwen (2003), Guenther & Lee (1996), Renardy & Rogers (2004), and Stakgold (1998). Strichartz (1994) gives an outstanding perspective on Fourier transforms and distributions.

More specialized, theoretical books include Friedman (1964) for parabolic equations, Gilbarg & Trudinger (1983) for elliptic equations, Protter & Weinberger (1967) for maximum principles, and Pao (1992) for both nonlinear parabolic and elliptic equations. The treatises by Courant & Hilbert (1953, 1962) provide a wealth of information on elliptic and hyperbolic equations, and Courant & Friedrichs (1948) is still a standard in shock waves and gas dynamics. Kevorkian & Cole (1981) and Zauderer (2006) discuss perturbation methods. For nonlinear equations one can consult Lax (1973), Smoller (1994) and Whitham (1974), all of which are key books. Another text on nonlinear PDEs that is very similar to the first edition of the present text (Logan 1994) is Debnath (1997). Bhatnagar (1979) is an excellent introduction to nonlinear dispersive waves, and solitons are discussed in Drazin & Johnson (1989).

Two excellent volumes that introduce several common and ad hoc methods for nonlinear PDEs are Ames (1965, 1972)