Chapter 8

Space and Time

8.1 Partial Differential Equations

In previous chapters we have considered differential equations in a single independent variable. This independent variable was either time, as in the case of the trajectory or predator-prey problems of Chapter 2, or a space variable as in the problems of Chapter 3. We now begin the study of the numerical solution of differential equations in two or more independent variables – partial differential equations. In the present chapter the two independent variables will be time and a single space variable; in the next chapter we shall treat problems in more than one space variable.

Example Equations

As two examples, we will concentrate on

$$u_t = cu_{xx} \tag{8.1.1}$$

and

$$u_{tt} = cu_{xx}, \tag{8.1.2}$$

which are known, respectively, as the *heat* (or *diffusion*) *equation* and the *wave equation*. In both (8.1.1) and (8.1.2) c is a given constant, and subscripts denote partial derivatives. The heat equation (8.1.1) is the prototype example of a *parabolic equation*, and (8.1.2) is an example of a *hyperbolic equation* (see the Supplementary Discussion.) The third standard type of partial differential equation is called *elliptic*, and the simplest example of this type is *Poisson's equation*

$$u_{xx} + u_{yy} = f, \tag{8.1.3}$$

where f is a given function of x and y. If $f \equiv 0$, then (8.1.3) is *Laplace's equation*. Elliptic equations will be treated in Chapter 9.

Initial and Boundary Conditions for the Heat Equation

As we saw with ordinary differential equations, a solution of a differential equation is not determined without appropriate initial and/or boundary conditions, and we expect the same to be true for partial differential equations. For (8.1.1) we need to prescribe an initial condition, and if x ranges over only a finite interval, we also need boundary conditions at the endpoints of this interval. Thus if the domain of x is $0 \leq x \leq 1$, the initial condition would be

$$u(0, x) = g(x), \qquad 0 \leq x \leq 1, \tag{8.1.4}$$

where g is a given function. We will also prescribe the boundary conditions

$$u(t, 0) = \alpha, \qquad u(t, 1) = \beta, \qquad t \geq 0, \tag{8.1.5}$$

for given constants α and β. The equation (8.1.1) together with the conditions (8.1.4) and (8.1.5) is the mathematical model for the temperature distribution in a thin rod whose ends are held at fixed temperatures α and β, and whose initial temperature distribution is $g(x)$. The solution $u(x, t)$ then gives the temperature within the rod as a function of time. For this model the constant c in (8.1.1) is $c = k/(s\rho)$, where $k > 0$ is the thermal conductivity, s is the specific heat of the material, and ρ is the mass density. Thus $c > 0$.

There are several variations on this problem that can be modeled by changing the boundary conditions or the equation itself. For example, suppose that we assume that the right end of the rod is, like the sides, perfectly insulated; by definition, then, we expect no heat loss or change in temperature across this end, so the boundary conditions (8.1.5) are changed to

$$u(t, 0) = \alpha, \qquad u_x(t, 1) = 0. \tag{8.1.6}$$

Another variation is to suppose that the rod is not homogeneous – as has been tacitly assumed – but is made of an alloy whose components vary as a function of x. Then the density as well as the thermal conductivity and specific heat will generally also vary with x, so that $c = c(x)$. Thus the differential equation is now one with a variable rather than constant coefficient. Going one step further, the thermal conductivity will, in general, depend not only on the material but also on the temperature itself. For many problems this dependence is so slight that it can be ignored, but for others it cannot. Thus we may have $c = c(u)$ so that the equation (8.1.1) is now nonlinear.

The heat equation is also a mathematical model of various other physical phenomena such as the diffusion of a gas.

The Wave Equation

Consider next the wave equation (8.1.2). This equation, or more general forms of it, models various types of wave propagation phenomena, as, for

example, in acoustics. One classical problem is a vibrating string. Consider, for example, a taut string along the x-axis that is fastened at $x = 0$ and $x = 1$. If the string is plucked, it will vibrate. We assume that the string is "ideal" – that is, it is perfectly flexible, and the tension T is constant as a function of both x and t and is large compared to the weight of the string. We denote the deflection of the string at a point x and time t by $u(t, x)$, and we assume that the deflections u are small compared to the length of the string. We assume, moreover, that the slope of the deflected string at any point is small compared to unity and that the horizontal displacement of the string is negligible compared to the vertical displacement (this is sometimes called *transverse motion*). We are also tacitly assuming that the motion of the string is only in a plane. The constant c in (8.1.2) is then equal to gT/w, where g is the gravitational constant and w is the weight of the string per unit length. In more general situations, T, and hence c, may not be constant.

In addition to the differential equation, we again need suitable initial and boundary conditions. Since the ends of the string are fixed at $x = 0$ and $x = 1$, we have the boundary conditions

$$u(t, 0) = 0 \qquad u(t, 1) = 0. \tag{8.1.7}$$

For the initial conditions we must specify the initial deflection as well as the initial velocity of the string; thus for a given function f, we will use

$$u(0, x) = f(x), \qquad u_t(0, x) = 0, \tag{8.1.8}$$

where the latter condition implies zero initial velocity of the string. With the differential equation (8.1.2), the boundary conditions (8.1.7), and the initial conditions (8.1.8), the problem is now fully specified. For other problems different boundary or initial conditions may be given, but since the equation is second order in t two initial conditions must generally be given, just as for ordinary differential equations.

The purpose of this book is, of course, to study techniques for the numerical solution of problems. It is worth recalling here, however, that there is a classical analytical technique for representing the solution of both (8.1.1) and (8.1.2) by means of Fourier series. This technique is valid only under very restrictive conditions, but it does apply to the heat and wave equations together with the types of initial and boundary conditions we have considered and for c constant. It is the *method of separation of variables*, and we shall review it rather briefly.

Separation of Variables

Assume that the solution of (8.1.1) can be written as a product of a function that depends only on t and a function that depends only on x:

$$u(t, x) = v(t)w(x). \tag{8.1.9}$$

If we substitute (8.1.9) into (8.1.1), we obtain

$$v'(t)w(x) = cv(t)w''(x), \tag{8.1.10}$$

or, assuming that neither v nor w is zero,

$$\frac{v'(t)}{v(t)} = c\frac{w''(x)}{w(x)}. \tag{8.1.11}$$

Since the left side of (8.1.11) is a function only of t and the right side is a function only of x, it follows that both sides must be equal to some constant, say μ; thus,

$$v'(t) = \mu v(t), \qquad cw''(x) = \mu w(x). \tag{8.1.12}$$

The general solution of the first equation is

$$v(t) = c_1 e^{\mu t}, \qquad c_1 = \text{ constant}. \tag{8.1.13}$$

The second equation of (8.1.12) is the eigenvalue problem discussed in Section 7.1. As we saw there, the eigenfunctions are

$$w(x) = \sin k\pi x, \qquad 0 \le x \le 1, \quad k = 1, 2, \ldots, \tag{8.1.14}$$

and the eigenvalues are $-k^2\pi^2$. Because of the constant c, the corresponding values of μ are

$$\mu = -ck^2\pi^2, \qquad k = 1, 2, \ldots . \tag{8.1.15}$$

Hence any function w of the form (8.1.14) is a solution of $cw'' = \mu w$ provided that μ is given by (8.1.15). Note that such a solution would vanish at certain points in the interval $[0, 1]$, and (8.1.11) would not be valid there. However, the use of (8.1.11) is meant only to be suggestive; the final criterion is whether v and w satisfy (8.1.10), and this will be the case provided that (8.1.12) holds. Thus any function of the form

$$u(t, x) = e^{-ck^2\pi^2 t} \sin k\pi x, \tag{8.1.16}$$

where k is any positive integer, satisfies the heat equation (8.1.1), as may be verified directly (Exercise 8.1.1).

Even though (8.1.16) satisfies the differential equation, it need not satisfy the initial or boundary conditions. Consider the special case of the boundary conditions (8.1.5) for which

$$u(t, 0) = 0, \qquad u(t, 1) = 0, \tag{8.1.17}$$

and suppose that the initial condition can be written as a finite trigonometric sum

$$u(0, x) = \sum_{k=1}^{n} a_k \sin k\pi x. \tag{8.1.18}$$

It is then easy to verify (Exercise 8.1.1) that

$$u(t,x) = \sum_{k=1}^{n} a_k e^{-ck^2\pi^2 t} \sin k\pi x \tag{8.1.19}$$

is a solution of (8.1.1) and, moreover, satisfies the boundary and initial conditions (8.1.17) and (8.1.18).

The solution (8.1.19) is predicated on the finite expansion (8.1.18), but by the theory of Fourier series a very large class of functions, and hence initial conditions, can be represented by the infinite series

$$g(x) = \sum_{k=1}^{\infty} a_k \sin k\pi x, \qquad a_k = 2\int_0^1 g(z)\sin(k\pi z)dz. \tag{8.1.20}$$

In this case the solution can be given, analogously to (8.1.19), by

$$u(t,x) = \sum_{k=1}^{\infty} a_k e^{-ck^2\pi^2 t} \sin k\pi x, \tag{8.1.21}$$

although it is no longer as simple as in the case of (8.1.19) to verify rigorously that this is a solution.

The method of separation of variables can also be applied to the wave equation, and we will just indicate the result corresponding to (8.1.21) for equation (8.1.2) together with the boundary and initial conditions (8.1.7) and (8.1.8). Again, if we assume that the first initial condition of (8.1.8) can be represented by

$$f(x) = \sum_{k=1}^{\infty} a_k \sin(k\pi x), \tag{8.1.22}$$

then the solution is

$$u(t,x) = \sum_{k=1}^{\infty} a_k \sin(k\pi x)\cos(k\pi\sqrt{c}t). \tag{8.1.23}$$

In the case that (8.1.22) is a finite sum, analogous to (8.1.18), it is easy to verify this result directly (Exercise 8.1.2) without any technical difficulties.

We have not meant to imply that these series expansions are to be the basis for numerical methods, although in certain special cases they can be. Rather, such representations are sometimes useful in ascertaining qualitative information about the solution of the differential equation. For example, (8.1.19) clearly shows, since $c > 0$, that $u(t,x) \to 0$ as $t \to \infty$, and this same conclusion can be reached from the infinite series (8.1.21). We shall use this information in the following section on finite difference methods. We shall also use the

technique of separation of variables, applied to difference equations, to study stability properties of the numerical methods in the following sections.

Supplementary Discussion and References: 8.1

Consider a partial differential equation of the form

$$au_{xx} + bu_{xt} + cu_{tt} + du_x + eu_t + fu = g,$$

where the coefficients $a, b, \ldots$ are functions of x and t. Then the equation is *elliptic* if

$$[b(x,t)]^2 < a(x,t)c(x,t) \tag{8.1.24}$$

for all x, t in the region of interest. Laplace's equation is the special case in which $a = c = 1$ and all other coefficients are zero. The equation is *hyperbolic* if $b^2 > ac$; this is the case for the wave equation. Finally, the equation is *parabolic* if $b = ac$, which holds for the heat equation since $b = c = 0$.

Very interesting, important, and difficult problems occur for equations of *mixed type* in which the condition (8.1.24) holds in part of the domain while the opposite inequality holds in another part; that is, the equation is elliptic in part of the domain and hyperbolic in another part. Such problems arise, for example, in transonic airflow in which the flow is subsonic in part of the region (the elliptic part) and supersonic in another part (the hyperbolic part). In these problems the variable t is a second space variable y.

Many, if not most, partial differential equation models of physical phenomena involve systems of equations rather than a single equation. The classification system of elliptic, hyperbolic, and parabolic can be extended to systems of equations, although relatively few systems that model realistic physical situations fit into this nice classification.

The method of separation of variables together with the use of Fourier series is a classical technique for solving certain simple equations and is discussed in most beginning textbooks on partial differential equations. Most such books will contain additional examples, derivations of equations, and classification theory. See, for example, Haberman [1983] and Keener [1988], and, for more advanced treatments, Courant and Hilbert [1953, 1962], and Garabedian [1986].

EXERCISES 8.1

8.1.1. Show that the function of (8.1.16) satisfies (8.1.1) for any integer k. Then verify that (8.1.19) is a solution of (8.1.1) that satisfies the initial and boundary conditions (8.1.17) and (8.1.18).

8.1.2. **a.** Show that

$$u(x,t) = \sum_{k=1}^{n} a_k \sin(k\pi x)\cos(k\pi\sqrt{c}t) \tag{8.1.25}$$

satisfies the wave equation (8.1.2) as well as the boundary and initial conditions $u(t,0) = u(t,1) = u_t(0,x) = 0$, $u(0,x) = \sum_{k=1}^{n} a_k \sin(k\pi x)$.

b. Write a program to display the special solution $u(x,t) = \sin(\pi x)\cos(\pi t)$ on a graphics terminal in such a way that the motion of the string is clear. Do the same for more complicated solutions consisting of two and three terms of (8.1.25), for $c = 1$.

8.2 Explicit Methods and Stability

We begin in this section the study of finite difference methods for partial differential equations and, in particular, for the equations discussed in the previous section. We will treat first the heat equation

$$u_t = cu_{xx}, \qquad 0 \le x \le 1, \quad t \ge 0, \tag{8.2.1}$$

with the initial and boundary conditions

$$u(0,x) = g(x), \qquad 0 \le x \le 1, \tag{8.2.2}$$

$$u(t,0) = \alpha, \qquad u(t,1) = \beta, \qquad t \ge 0. \tag{8.2.3}$$

Difference Equations for the Heat Equation

We set up a grid in the x,t plane with grid spacings Δx and Δt as illustrated in Figure 8.1. The idea of the simplest finite difference method for (8.2.1) is to replace the second derivative on the right-hand side of (8.2.1) with a central difference quotient in x, and replace u_t with a forward difference in time. Then one advances the approximate solution forward in time one time level after another. More precisely, if we let u_j^m denote the approximate solution at $x_j = j\Delta x$ and $t_m = m\Delta t$, then the finite difference analog of (8.2.1) is

$$\frac{u_j^{m+1} - u_j^m}{\Delta t} = \frac{c}{(\Delta x)^2}(u_{j+1}^m - 2u_j^m + u_{j-1}^m), \tag{8.2.4}$$

or

$$u_j^{m+1} = u_j^m + \mu(u_{j+1}^m - 2u_j^m + u_{j-1}^m), \qquad j = 1, \ldots, n, \tag{8.2.5}$$

where

$$\mu = \frac{c\Delta t}{(\Delta x)^2}. \tag{8.2.6}$$

The boundary conditions (8.2.3) give the values

$$u_0^m = \alpha, \qquad u_{n+1}^m = \beta, \quad m = 0, 1, \ldots,$$

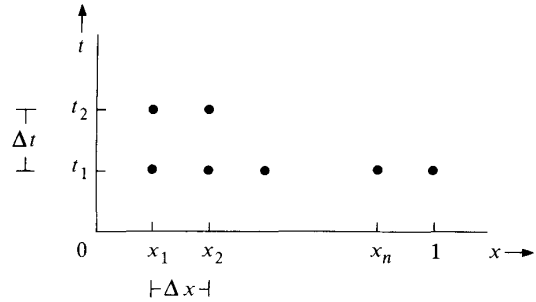


Figure 8.1: *Grid Spacings*

and the initial condition (8.2.2) furnishes

$$u_j^0 = g(x_j), \qquad j = 1, \ldots, n.$$

Therefore (8.2.5) provides a prescription for marching the approximate solution forward one time step after another: the values u_j^1, $j = 1, \ldots, n$, are first all obtained, and knowing these we can obtain u_j^2, $j = 1, \ldots, n$, and so on.

How accurate will be the approximate solution obtained by (8.2.5)? A rigorous answer to this question is a difficult problem that is beyond the scope of this book, but we will attempt to obtain some insight by considering two aspects of the error analysis.

Discretization Error

Let $u(t, x)$ be the exact solution of (8.2.1) together with the initial and boundary conditions (8.2.2) and (8.2.3). If we put this exact solution into the difference formula (8.2.4), the amount by which the formula fails to be satisfied is called the *local discretization error* (or *local truncation error*); at the point (t, x) the local discretization error, e, is

$$\frac{u(t + \Delta t, x) - u(t, x)}{\Delta t} - \frac{c}{(\Delta x)^2}[u(t, x + \Delta x) - 2u(t, x) + u(t, x - \Delta x)]. \quad (8.2.7)$$

This is entirely analogous to the previous definitions of local discretization error for ordinary differential equations and enjoys similar properties. For example, suppose that we know the exact solution $u(t, x)$ for some t and all $0 \leq x \leq 1$, and we use (8.2.4) to estimate the solution at $t + \Delta t$. Call this estimate $\hat{u}(t + \Delta t, x)$. Then, by definition

$$\frac{\hat{u}(t + \Delta t, x) - u(t, x)}{\Delta t} = \frac{c}{(\Delta x)^2}[u(t, x + \Delta x) - 2u(t, x) + u(t, x - \Delta x)],$$

so if we subtract this from (8.2.7), we obtain

$$\hat{u}(t + \Delta t, x) - u(t + \Delta t, x) = \Delta t e. \tag{8.2.8}$$

Thus the error caused by one time step using the difference scheme (8.2.4) is Δt times the local discretization error.

It is easy to estimate the quantity e of (8.2.7) in terms of Δt and Δx. If we consider u as only a function of t for fixed x, we can apply the Taylor expansion

$$u(t + \Delta t, x) = u(t, x) + u_t(t, x)\Delta t + O[(\Delta t)^2]$$

to conclude that

$$\frac{u(t + \Delta t, x) - u(t, x)}{\Delta t} = u_t(t, x) + O(\Delta t).$$

Similarly, by Taylor expansions in x, we have

$$\frac{u(t, x + \Delta x) - 2u(t, x) + u(t, x - \Delta x)}{(\Delta x)^2} = u_{xx}(t, x) + O[(\Delta x)^2].$$

If we put these expressions into (8.2.7) and use $u_t = cu_{xx}$ (since u is the exact solution of the differential equation), we obtain

$$e = 0(\Delta t) + 0[(\Delta x)^2]. \tag{8.2.9}$$

The fact that Δt appears to the first power and Δx to the second power in this expression for the local discretization error is usually described by the statement that the finite difference method (8.2.4) is *first-order accurate in time* and *second-order accurate in space.*

Stability

It is tempting to conclude from (8.2.8) and (8.2.9) that the discretization error in u_j^m, as obtained from (8.2.5), converges to zero as Δt and Δx tend to zero. Unfortunately, this conclusion is not warranted since (8.2.8) gives the error in the approximate solution only for a single time step. To show that the discretization error tends to zero on a whole time interval $[0, T]$ is difficult and, in general, requires additional conditions on how Δt and Δx tend to zero. A relationship of the type (8.2.9), or more generally a statement that the local discretization error tends to zero with Δt and Δx, is essentially a necessary condition for the global discretization error itself to tend to zero, and is called *consistency* of the difference scheme. The reason that consistency of the difference method does not necessarily imply convergence of the discretization error is connected with *stability* of the difference scheme, and we now discuss certain aspects of this for the difference scheme (8.2.5).

In a way exactly analogous to the method of separation of variables and the use of Fourier series that were applied to the differential equation $u_t = cu_{xx}$ in the previous section, we can obtain the exact solution of (8.2.5) together with the boundary and initial conditions

$$\begin{aligned} u_0^m &= u_{n+1}^m = 0, \qquad m = 1, 2, \ldots, \\ u_j^0 &= g(x_j), \qquad j = 1, \ldots, n. \end{aligned} \tag{8.2.10}$$

Assume that the solution u_j^m can be written as

$$u_j^m = v_m w_j, \qquad j = 1, \ldots, n, \qquad m = 0, 1, \ldots . \tag{8.2.11}$$

This is the paradigm of separation of variables for difference equations. Putting (8.2.11) into (8.2.5) and collecting terms yields

$$\frac{v_{m+1} - v_m}{\mu v_m} = \frac{w_{j+1} - 2w_j + w_{j-1}}{w_j}, \qquad j = 1, \ldots, n, \quad m = 0, 1, \ldots .$$

Since the left side is independent of j and the right side is independent of m, both sides must be equal to some constant, say $-\lambda$; thus

$$v_{m+1} - v_m = -\lambda \mu v_m, \qquad m = 0, 1, \ldots, \tag{8.2.12}$$

$$w_{j+1} - 2w_j + w_{j-1} = -\lambda w_j, \qquad j = 1, \ldots, n, \tag{8.2.13}$$

where $w_0 = w_{n+1} = 0$ from the boundary conditions (8.2.10). Equation (8.2.13) represents the eigenvalue problem for the $(2, -1)$ tridiagonal matrix of (3.1.10). The eigenvalues of this matrix are (see Exercise 4.4.5)

$$\lambda_k = 2 - 2\cos k\pi \Delta x, \qquad k = 1, \ldots, n, \tag{8.2.14}$$

with corresponding eigenvectors (Exercise 4.4.5)

$$\mathbf{w} = [\sin(k\pi\Delta x), \sin(2k\pi\Delta x), \ldots, \sin(nk\pi\Delta x)]^T, \qquad k = 1, \ldots, n,$$

where $\Delta x = 1/(n+1)$. Thus for each $\lambda = \lambda_k$

$$w_j = \sin(jk\pi\Delta x), \qquad j = 0, 1, \ldots, n+1, \tag{8.2.15}$$

is a solution of (8.2.13). Clearly

$$v_m = (1 - \lambda\mu)^m v_0, \qquad m = 0, 1, \ldots,$$

is a solution of (8.2.12) for any λ, so that

$$u_j^m = v_m w_j = (1 - \lambda_k \mu)^m \sin(jk\pi\Delta x), \qquad m = 0, 1, \ldots, \; j = 0, 1, \ldots, n+1,$$

is a solution of (8.2.5) for each k. As with the differential equation, any linear combination of these solutions is also a solution; thus

$$u_j^m = \sum_{k=1}^{n} a_k(1 - \lambda_k\mu)^m \sin(jk\pi\Delta x) \tag{8.2.16}$$

is a solution of (8.2.5) for any constants a_k. If the a_k are chosen so that

$$a_k = \sum_{l=1}^{n} g(x_l) \sin k\pi l\Delta x, \tag{8.2.17}$$

then (Exercise 8.2.2) u_j^m also satisfies the initial condition

$$u_j^0 = g(x_j), \qquad j = 1, \ldots, n. \tag{8.2.18}$$

We now use the representation (8.2.16) in the following way. From our discussion in the previous section, the equation $u_t = cu_{xx}$ together with the boundary conditions $u(t, 0) = u(t, 1) = 0$ is a model of the temperature distribution in a thin insulated rod whose ends are held at zero temperature. Since there is no source of heat, we expect that the temperature of the rod will decrease to zero so that $u(t, x) \to 0$ as $t \to \infty$. This conclusion can also be obtained mathematically from the series representations (8.1.19) or (8.1.21) of the solution since the exponential terms all tend to zero. Therefore it is reasonable to demand that the finite difference approximations u_j^m also tend to zero as m tends to infinity, for any initial conditions; by (8.2.16), this will be the case if and only if

$$|1 - \mu\lambda_k| < 1, \qquad k = 1, \ldots, n. \tag{8.2.19}$$

Since μ and all the λ_k are positive, (8.2.19) will hold if and only if

$$-(1 - \mu\lambda_k) < 1, \qquad k = 1, \ldots, n,$$

or

$$\mu < \min_k \frac{2}{\lambda_k} = \frac{1}{1 - \cos \pi n\Delta x} = \frac{1}{1 + \cos \pi\Delta x}, \tag{8.2.20}$$

since the largest λ_k is λ_n. Thus, with $\mu = c\Delta t/(\Delta x)^2$ from (8.2.6), (8.2.20) becomes

$$\Delta t < \frac{(\Delta x)^2}{c(1 + \cos \pi\Delta x)}. \tag{8.2.21}$$

This gives a restriction on the relative sizes of Δt and Δx, which if not satisfied will, in general, mean that the approximate solution u_j^m of the difference scheme will ultimately diverge as m tends to infinity and, obviously, will become an increasingly poor approximation to the solution of the differential

Table 8.1: *Unstable Behavior*

t/x	0.2	0.4	0.6	0.8
0	0.59	0.95	0.95	0.59
0.16	0.08	0.13	0.13	0.08
0.24	0.03	0.05	0.05	0.03
0.28	0.02	0.03	0.02	0.004
0.32	0.01	0.02	0.09	0.20
0.36	0.005	−0.14	−1.22	−2.43
0.40	0.12	3.20	19.0	32.2
0.44	−3.97	−62.0	−286.9	−428.

equation, which tends to zero. In Table 8.1 we give an example of this instability for (8.2.5) with $c = 1$, $g(x) = \sin \pi x$, $\alpha = \beta = 0$ and with $\Delta x = 0.1$ and $\Delta t = 0.04$ so that (8.2.21) is not satisfied. Note that the instability has begun to develop noticeably by $t = 0.32$ and then rapidly worsens.

We can replace (8.2.21) by the slightly stronger condition

$$\Delta t \leq \frac{(\Delta x)^2}{2c}, \tag{8.2.22}$$

which always implies (8.2.21). This relation is called the *stability condition* for the difference method (8.2.5). Our derivation of it has been in the context of the behavior of u_j^m as $m \to \infty$ for fixed Δt and Δx. But it is also relevant to the problem of the convergence of the discretization error to zero as Δt and Δx tend to zero. In fact, although it is beyond our scope to prove this, the approximate solutions will converge to the exact solution as Δt and Δx tend to zero if (8.2.22) holds as Δt and Δx tend to zero. This is a special case of a more general principle known as the *Lax Equivalence Theorem*, which states that for quite general differential equations and consistent difference schemes, the global discretization error will tend to zero if and only if the method is stable.

The condition (8.2.22) imposes an increasingly stringent limitation on the time step Δt as the space increment Δx becomes small, as Table 8.2 shows for the case $c = 1$. Thus we may require a time step far smaller than otherwise necessary to resolve the time-dependent nature of the solution of the differential equation itself. Although the analysis that we have done has been restricted to the simplest differential equation and simplest difference scheme, the requirement of small time steps for explicit finite difference methods for parabolic and similar equations is a general problem and is a primary motivation for the implicit methods to be discussed in the next section.

Table 8.2: *Maximum Time Steps for Given Δx and $c = 1$*

Δx	Δt
0.1	$0.5 \cdot 10^{-2}$
0.01	$0.5 \cdot 10^{-4}$
0.001	$0.5 \cdot 10^{-6}$

The Wave Equation

We turn now to hyperbolic equations, and in particular to the wave equation

$$u_{tt} = cu_{xx}, \qquad 0 \le x \le 1, \qquad t \ge 0, \tag{8.2.23}$$

together with the initial and boundary conditions

$$u(0,x) = f(x), \quad u_t(0,x) = g(x), \quad u(t,0) = \alpha, \quad u(t,1) = \beta. \tag{8.2.24}$$

As we discussed in the previous section, the problem (8.2.23), (8.2.24) is a mathematical model for a vibrating string.

The simplest finite difference scheme for (8.2.23) is

$$\frac{u_j^{m+1} - 2u_j^m + u_j^{m-1}}{(\Delta t)^2} = \frac{c}{(\Delta x)^2}(u_{j+1}^m - 2u_j^m + u_{j-1}^m). \tag{8.2.25}$$

To obtain (8.2.25) we have used the usual centered difference formula for u_{xx}, just as in (8.2.4), as well as for u_{tt}. Note that (8.2.25) now involves three time levels and requires that both u_j^m and u_j^{m-1} be known in order to advance to the $(m+1)$st time level. This requires additional storage for u_j^{m-1}, as opposed to the method (8.2.5) for $u_t = cu_{xx}$, but is a natural consequence of the fact that the differential equation contains a second derivative in time. We also require both u_j^0 and u_j^1 in order to start, and these can be obtained from the initial conditions (8.2.24):

$$u_j^0 = f(x_j) \qquad u_j^1 = f(x_j) + \Delta t g(x_j), \qquad j = 1, \ldots, n, \tag{8.2.26}$$

where the second condition is obtained by approximating $u_t(0,x) = g(x)$ by $[u(\Delta t, x) - u(0,x)]/\Delta t = g(x)$. From the boundary conditions, we have

$$u_0^m = \alpha, \qquad u_{n+1}^m = \beta, \qquad m = 0, 1, \ldots . \tag{8.2.27}$$

Thus the values at the $(m+1)$st time level are obtained by

$$u_j^{m+1} = 2u_j^m - u_j^{m-1} + \mu(u_{j+1}^m - 2u_j^m + u_{j-1}^m), \tag{8.2.28}$$

where

$$\mu = \frac{c(\Delta t)^2}{(\Delta x)^2}.$$

It is easy to show that the local discretization error for (8.2.25) is $0(\Delta t)^2 + 0(\Delta x)^2$, so that the method is second-order accurate in both space and time (Exercise 8.2.4). For the stability analysis we can again proceed by the method of separation of variables and assume that the boundary conditions α and β are zero. Let $u_j^m = v_m w_j$. Putting this into (8.2.25) leads to the two conditions

$$v_{m+1} - 2v_m + v_{m-1} = -\lambda\mu v_m, \qquad m = 1, \ldots, \tag{8.2.29}$$

$$w_{j+1} - 2w_j + w_{j-1} = -\lambda w_j, \qquad j = 1, \ldots, n. \tag{8.2.30}$$

The second set of equations is the same as (8.2.13), and thus its solutions are given by (8.2.14),(8.2.15). The equations (8.2.29), although ostensibly of the same form as (8.2.30), are for an initial-value problem in which v_0 and v_1 are known. From Section 2.5 [see (2.5.11) - (2.5.14)], the solution of (8.2.29) is given by

$$v_m = \gamma_1 \eta_+^m + \gamma_2 \eta_-^m, \qquad m = 0, 1, \ldots,$$

where $\eta_\pm$ are the roots of the characteristic equation $\eta^2 + (\lambda\mu - 2)\eta + 1 = 0$ and are given by

$$\eta_\pm = \tfrac{1}{2}(2 - \lambda\mu \pm \sqrt{\lambda^2\mu^2 - 4\lambda\mu}).$$

The γ_i can be obtained from the initial conditions v_0 and v_1. Thus the solution of (8.2.25) can be written as

$$u_j^m = \sum_{k=1}^{n} a_k(\gamma_{k,1}\eta_{k,+}^m + \gamma_{k,2}\eta_{k,-}^m)\sin(jk\pi\Delta x), \tag{8.2.31}$$

where the subscript k indicates that the corresponding γ and η have been computed for $\lambda = \lambda_k$. For u_j^m to remain bounded for arbitrary a_k and initial conditions, it is necessary and sufficient that $|\eta_{k,\pm}| \leq 1$. If

$$\lambda_k\eta - 4 \leq 0, \tag{8.2.32}$$

then it is easy to verify that $|\eta_{k,\pm}| = 1$, and if $\lambda_k\eta - 4 > 0$, then $\eta_{k,-} < -1$. Hence (8.2.32) is the necessary and sufficient condition for stability. Since the eigenvalues λ_k satisfy $0 < \lambda_k < 4$, a sufficient condition for (8.2.32) is that $\mu \leq 1$, or

$$\Delta t \leq \frac{\Delta x}{\sqrt{c}}. \tag{8.2.33}$$

This is also essentially a necessary condition in the sense that $\lambda_n \to 4$ as $n \to \infty$, so any condition weaker that (8.2.33) would allow (8.2.32) to be violated for sufficiently large n. The stability condition (8.2.33) is much less

stringent on Δt than was (8.2.22) for the heat equation. Indeed, (8.2.33) shows that Δt need decrease only proportionately to Δx, rather than as the square of Δx, as was the case with (8.2.22).

Although this section has dealt only with the heat and wave equations, the same principles of obtaining finite difference methods apply to more general initial-boundary value problems for either single equations or systems of equations. In all such cases the user must be alert to the possibility of instability, although for most equations a simple analysis of the form given in this section will not be possible.

Supplementary Discussion and References: 8.2

We have only touched the surface of finite difference methods for parabolic and hyperbolic equations. In particular, we have considered only first- and second-order methods, although a variety of higher-order methods have been developed. More importantly, most of the useful methods for parabolic equations are implicit, and these will be dealt with in the next section. The books by Ames [1977] and Hall and Porsching [1990] give many additional methods. See also Isaacson and Keller [1966] and Richtmyer and Morton [1967] for a discussion of methods as well as a rigorous analysis of discretization error and stability criteria.

The separation of variables analysis leading to (8.2.16) can also be viewed in matrix terms. The matrix-vector formulation of (8.2.5) is

$$\mathbf{u}^{m+1} = \mathbf{u}^m - \mu A \mathbf{u}^m, \qquad m = 0, 1, \ldots,$$

where A is the $(2, -1)$ matrix of (8.2.13), and $\mathbf{u}^m$ is the vector with components $u_1^m, \ldots, u_n^m$. The matrix $H = I - \mu A$ has eigenvalues $1 - \mu\lambda_k$ and eigenvectors $\mathbf{w}_k$, where the λ_k and $\mathbf{w}_k$ are the eigenvalues and eigenvectors of A. Therefore if

$$\mathbf{u}^0 = \sum_{k=1}^{n} a_k \mathbf{w}_k$$

is the expansion of $\mathbf{u}^0$ in terms of the eigenvectors, then

$$\mathbf{u}^m = H\mathbf{u}^{m-1} = \cdots = H^m \mathbf{u}^0 = \sum_{k=1}^{n} a_k (1 - \mu\lambda_k)^m \mathbf{w}_k,$$

which is (8.2.16).

EXERCISES 8.2

8.2.1. Use (8.2.5) to approximate a solution to $u_t = u_{xx}$ for the boundary and initial conditions $u(t, 0) = 0$, $u(t, 1) = 1$, and $u(0, x) = \sin \pi x + x$. Use different values of Δt and Δx and discuss your approximate solutions. For what ratios of Δt and Δx do you conclude that your approximate solution is stable?

8.2.2. Substitute (8.2.16) into (8.2.5) and verify that it is indeed a solution. Verify also that (8.2.18) holds if the a_k are given by (8.2.17).

8.2.3. Repeat and verify the calculation of Table 8.1.

8.2.4. Use the approach discussed in the text for the heat equation to conclude that the method (8.2.25) for the wave equation is second-order accurate in both time and space.

8.2.5. Write a computer program to solve the wave equation (8.2.23) with the initial and boundary conditions (8.2.24) by the difference method (8.2.25). Apply your program for various values of Δt and Δx, and conclude that the calculation is stable if (8.2.33) is satisfied.

8.3 Implicit Methods

The finite difference method (8.2.5) discussed in the previous section is called explicit because the values of u_j^{m+1} at the next time level are obtained by an explicit formula in terms of the values at the previous time level. In contrast, consider again the heat equation

$$u_t = cu_{xx}, \qquad 0 \le x \le 1, \qquad t \ge 0, \tag{8.3.1}$$

and the difference approximation

$$\frac{u_j^{m+1} - u_j^m}{\Delta t} = \frac{c}{(\Delta x)^2}(u_{j+1}^{m+1} - 2u_j^{m+1} + u_{j-1}^{m+1}), \qquad j = 1, \ldots, n. \tag{8.3.2}$$

This is similar in form to (8.2.4) but has the important difference that the values of u_j on the right side are now evaluated at the $(m+1)$st time level rather than the mth. Consequently, if we know u_j^m, $j = 1, \ldots, n$, and are ready to compute u_j^{m+1}, $j = 1, \ldots, n$, we see that the variables u_j on the right-hand side of (8.3.2) are all unknown. Thus we must view (8.3.2) as a system of equations that implicitly defines the values u_j^{m+1}, $j = 1, \ldots, n$. This is one of the basic differences between implicit and explicit methods: in an explicit method we have a formula for u_j^{m+1}, such as (8.2.5), in terms of known values of u_j at previous time levels, whereas with an implicit method we must solve a system of equations to advance to the next time level.

If, as in the previous section, we set $\mu = c\Delta t/(\Delta x)^2$, then we can rewrite (8.3.2) as

$$(1 + 2\mu)u_j^{m+1} - \mu(u_{j+1}^{m+1} + u_{j-1}^{m+1}) = u_j^m, \qquad j = 1, \ldots, n, \tag{8.3.3}$$

or in matrix-vector form,

$$(I + \mu A)\mathbf{u}^{m+1} = \mathbf{u}^m + \mathbf{b}, \qquad m = 0, 1, \ldots . \tag{8.3.4}$$

Here A is the $(2,-1)$ tridiagonal matrix of (3.1.10), and $\mathbf{u}^{m+1}$ and $\mathbf{u}^m$ are vectors with components u_i^{m+1} and u_i^m, $i = 1,\ldots,n$, respectively. If we use the same boundary conditions

$$u(t,0) = \alpha, \qquad u(t,1) = \beta \tag{8.3.5}$$

as in the previous section, then $u_0^k = \alpha$ and $u_{n+1}^k = \beta$ for $k = 0,1,\ldots$; thus the vector $\mathbf{b}$ in (8.3.4) is zero except for $\mu\alpha$ and $\mu\beta$ in the first and last components. We also assume the initial condition

$$u(0,x) = f(x), \qquad 0 \le x \le 1, \tag{8.3.6}$$

so, as before, $u_j^0 = f(x_j)$, $j = 1,\ldots,n$.

The implicit method (8.3.4) is now carried out by solving the linear system of equations (8.3.4) at each time step to obtain the u_j^{m+1} from u_j^m. The matrix in (8.3.4) is tridiagonal and also diagonally dominant (Section 3.1) since $c > 0$, and thus $\mu > 0$. Therefore, as we saw in Section 4.3, the system of equations can be efficiently solved by Gaussian elimination without pivoting. In the particular case of (8.3.4), we could compute the L and U factors once and for all, although we may not be able to do this for more general problems.

Even though each time step of (8.3.4) can be carried out relatively efficiently, this method is more costly per time step than the explicit method (8.2.5). However, in return for this additional cost we obtain a substantial benefit in the stability properties of the method, which in many cases will allow us to use a much larger time step than does the explicit method and thus will greatly cut the overall computing costs. We will now indicate the stability analysis following the lines of the previous section.

Stability

We assume, as before, that $\alpha = \beta = 0$. Then, corresponding to (8.2.16),

$$u_j^m = \sum_{k=1}^{n} a_k \gamma_k^m \sin(k\pi j \Delta x) \tag{8.3.7}$$

identically satisfies the difference scheme (8.3.3) for any constants a_k, provided that

$$\gamma_k = \frac{1}{1 + 2\mu(1 + \cos k\pi\Delta x)}, \qquad k = 1,\ldots,n. \tag{8.3.8}$$

Moreover, u_j^0 satisfies the initial condition if

$$a_k = \frac{2}{n+1} \sum_{l=1}^{n} f(x_l) \sin(k\pi l \Delta x). \tag{8.3.9}$$

The verification of these results is left to Exercise 8.3.2.

Now recall from our discussion in the previous section that we require that the approximate solution $u_j^m \to 0$ as $m \to \infty$ if it is to mirror the solution of the differential equation itself. From (8.3.7) we see that this will be the case, in general, if and only if

$$|\gamma_k| < 1, \qquad k = 1, \ldots, n. \tag{8.3.10}$$

But from (8.3.8), since $\mu > 0$,

$$0 < \gamma_k < 1, \qquad k = 1, \ldots, n, \tag{8.3.11}$$

so that (8.3.10) indeed holds. Most importantly, we see that (8.3.11) is true for *any* $\mu > 0$; thus, since $\mu = c\Delta t/(\Delta x)^2$, (8.3.11) is true for *any* ratio of Δt and Δx. We say in this case that the method is *unconditionally stable*, meaning that it is stable without restrictions on the relative size of Δt and Δx.

Discretization Error

The fact that the method (8.3.3) is unconditionally stable does *not* mean that we can expect to obtain a good approximate solution for any Δt and Δx. As usual, these must be chosen sufficiently small to control discretization error. Now it is the case (see Exercise 8.3.3) that (8.3.2), like the corresponding explicit method (8.2.5), is first-order accurate in time and second-order accurate in space; that is, the local discretization error will be

$$e = 0(\Delta t) + 0(\Delta x)^2.$$

Suppose that

$$e = c_1 \Delta t + c_2 (\Delta x)^2.$$

Then for the contributions to the total error from the discretization in time and the discretization in space to be commensurate, we require that

$$\Delta t \doteq c_3 (\Delta x)^2,$$

which is reminiscent of the stability condition (8.2.22) for the explicit method. Thus we see that although the stability requirements for our implicit method do not impose any restrictions on the relative sizes of Δt and Δx, the accuracy requirements may.

The Crank-Nicolson Method

A potentially better implicit method in this regard is the famous *Crank-Nicolson method*, which is an average of the explicit method (8.2.4) and the implicit method (8.3.2):

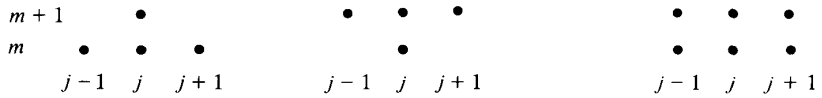
$$u_j^{m+1} - u_j^m = \frac{c\Delta t}{2(\Delta x)^2}(u_{j+1}^{m+1} - 2u_j^{m+1} + u_{j-1}^{m+1} + u_{j+1}^m - 2u_j^m + u_{j-1}^m). \tag{8.3.12}$$

This can be written in matrix-vector form as

$$(I + \frac{\mu}{2}A)\mathbf{u}^{m+1} = (I - \frac{\mu}{2}A)\mathbf{u}^m + \mathbf{b}, \qquad m = 0, 1, \ldots, \tag{8.3.13}$$

where A is again the $(2, -1)$ matrix. Hence (8.3.12) is carried out by solving a tridiagonal system of equations at each time step. The advantage of (8.3.12) is that it is not only unconditionally stable, as is (8.3.4), but it is second-order accurate in time as well as in space. (The verification of these assertions is left to Exercise 8.3.5). These properties have made it one of the most often used methods for parabolic equations.

One easy way to recall the three different methods (8.2.4), (8.3.2), and (8.3.12) is by their "stencils" of grid points as illustrated in Figure 8.2. These show which grid points enter into the difference method.



(a) Explicit (8.2.4) (b) Fully implicit (8.3.2) (c) Crank-Nicolson (8.3.12)

Figure 8.2: *Stencils for the Methods*

It has become common practice in the numerical solution of parabolic-type partial differential equations to use implicit methods since their good stability properties outweigh the additional work required per time step. Most of the methods in actual use are more complicated than the Crank-Nicolson method, but the principles are the same. However, for problems involving more than one space dimension, straightforward extensions of the implicit methods of this section are not satisfactory, and additional techniques are required. One such technique will be discussed in Section 9.1.

It is possible to formulate implicit methods for hyperbolic equations, such as the wave equation, in much the same way. However, as was seen with the wave equation, the stability requirements of explicit methods typically do not impose a stringent restriction on the time step. Consequently, implicit methods for hyperbolic equations are rather infrequently used in practice, and we shall not discuss them further.

Supplementary Discussion and References: 8.3

The references given in Section 8.2 are also relevant for implicit methods. In particular, Richtmyer and Morton [1967, pp. 189-91] summarize graphically

a number of implicit finite difference methods for parabolic equations in terms of their stencils.

EXERCISES 8.3

8.3.1. Write a program to carry out (8.3.4) and apply it to the problem of Exercise 8.2.1. Use various values of Δt and Δx and verify numerically the stability of the method. Discuss your results compared to those for Exercise 8.2.1, including the relative ease and efficiency of carrying out the two methods.

8.3.2. Proceed along the lines of the analysis of the previous section to verify that (8.3.7) satisfies (8.3.3).

8.3.3. Proceed along the lines of the analysis of the previous section to show that the local discretization error for the method (8.3.3) satisfies (8.2.9).

8.3.4. Modify your program of Exercise 8.3.1 to carry out the Crank-Nicolson method (8.3.12). Discuss your results and compare this method to (8.3.4).

8.3.5. For the Crank-Nicolson method (8.3.12) with the boundary conditions $\alpha = \beta = 0$, verify that the solution of (8.3.12) is of the form (8.3.7), where now $\gamma_k = (1 - \frac{1}{2}\mu\lambda_k)/(1 + \frac{1}{2}\mu\lambda_k)$, with $\lambda_k = 2 - 2\cos(k\pi\Delta x)$. Conclude that the method is unconditionally stable. Show also that the method is second-order accurate in both space and time.

8.3.6. The Dufort-Frankel method for the heat equation is

$$u_j^{m+1} - u_j^{m-1} = [\Delta t c/(\Delta x)^2](u_{j+1}^m - u_j^{m+1} - u_j^{m-1} + u_{j-1}^m).$$

Show that this method is unconditionally stable. Give an explicit formulation of it.

8.3.7. Consider the nonlinear parabolic equation

$$u_t = u_{xx} - u - x^2 - u^3,$$

with boundary and initial conditions $u(t,0) = u(t,1) = 0$, $u(0,x) = sin\pi x$.

a. Formulate an explicit method and do a stability analysis for the linear equation $u_t = u_{xx} - u$. Test numerically your stability criterion against the nonlinear equation.

b. Formulate a completely implicit method and write a program to solve the resulting nonlinear system at each time step by Newton's method. Verify numerically that your method is unconditionally stable.

c. The corresponding steady-state equation is the two-point boundary-value problem $v'' = v + x^2 + v^3$, $v(0) = v(1) = 0$. If you are only interested in the steady-state solution, would it be better to attack this equation directly by the methods of Chapter 5 or to integrate the partial differential equation to steady-state by the methods of parts **a** and **b**?

8.4 Semidiscrete Methods

We now consider another approach to initial boundary value problems which utilizes the projection principles of Chapter 6 and reduces the partial differential equation to an approximating system of ordinary differential equations. This approach can be applied in principle to both parabolic- and hyperbolic-type equations. We will first illustrate it for the heat equation

$$u_t = cu_{xx}, \qquad 0 \le x \le 1, \qquad t \le 0, \tag{8.4.1}$$

with the initial and boundary conditions

$$u(0,x) = f(x), \qquad u(t,0) = 0, \qquad u(t,1) = 0. \tag{8.4.2}$$

As in Section 6.1, let $\phi_1(x), \ldots, \phi_n(x)$ be a set of basis functions that satisfy the boundary conditions:

$$\phi_k(0) = 0 \qquad \phi_k(1) = 0, \qquad k = 1, \ldots, n. \tag{8.4.3}$$

We attempt to find an approximate solution $\hat{u}$ of (8.4.1) of the form

$$\hat{u}(t,x) = \sum_{i=1}^{n} \alpha_i(t)\phi_i(x), \tag{8.4.4}$$

where the α_i are to be determined. Note that this is the same approach taken in Chapter 6, with the exception that now we allow the coefficients α_i of the linear combination of the basis functions to be functions of t to reflect the time-dependent nature of the problem.

To determine the unknown coefficients α_i, we can apply any of the criteria of Section 6.1, and we will consider first collocation in the following way. We again let $0 \le x_1 < \cdots < x_n \le 1$ be (not necessarily equally spaced) grid points in the x variable, and we require that the approximate solution (8.4.4) satisfy the differential equation at these points; thus, since

$$\hat{u}_t(t,x) = \sum_{i=1}^{n} \alpha_i'(t)\phi_i(x), \qquad \hat{u}_{xx}(t,x) = \sum_{i=1}^{n} \alpha_i(t)\phi_i''(x),$$

we require that

$$\sum_{i=1}^{n} \alpha_i'(t)\phi_i(x_j) = c\sum_{i=1}^{n} \alpha_i(t)\phi_i''(x_j), \qquad j = 1, \ldots, n. \tag{8.4.5}$$

If we introduce the $n \times n$ matrices

$$A = (\phi_j(x_i)), \qquad B = c(\phi_j''(x_i)), \tag{8.4.6}$$

and the n -vectors

$$\boldsymbol{\alpha}(t) = (\alpha_1(t), \ldots, \alpha_n(t))^T, \qquad \boldsymbol{\alpha}'(t) = (\alpha_1'(t), \ldots, \alpha_n'(t))^T,$$

we can write (8.4.5) as

$$A\boldsymbol{\alpha}'(t) = B\boldsymbol{\alpha}(t). \tag{8.4.7}$$

Thus if we assume that A is nonsingular and set $C = A^{-1}B$,

$$\boldsymbol{\alpha}'(t) = C\boldsymbol{\alpha}(t). \tag{8.4.8}$$

Equation (8.4.8) is a system of n ordinary differential equations for the α_i. To solve this system we will need an initial condition. If we require that the approximate solution $\hat{u}$ satisfy the initial condition (8.4.2) at the points $x_1, \ldots, x_n$, we have

$$\sum_{i=1}^{n} \alpha_i(0)\phi_i(x_j) = f(x_j), \qquad j = 1, \ldots, n,$$

or

$$A\boldsymbol{\alpha}(0) = \mathbf{f}, \tag{8.4.9}$$

where $\mathbf{f} = (f(x_1), \ldots, f(x_n))^T$. By our assumption that A is nonsingular, we then have

$$\boldsymbol{\alpha}(0) = A^{-1}\mathbf{f}. \tag{8.4.10}$$

The conceptual problem is then to solve the system of ordinary differential equations (8.4.8) with the initial condition (8.4.10). If we could do this exactly, then $\hat{u}$ as given by (8.4.4) would be the approximate solution. Such a method is called *semidiscrete* because we have discretized only in space by means of the basis functions ϕ_i and grid points x_i while leaving time as a continuous variable. In practice, however, the system of differential equations (8.4.8) must be solved numerically, so a discretization of time is introduced by that process and the term semidiscrete is somewhat of a misnomer. Nevertheless, the conceptual viewpoint of considering discretization of only the space variables and thus reducing the problem to a system of ordinary differential equations is useful.

Let us consider the Euler method discussed in Chapter 2 for the numerical integration of the system (8.4.8). If $t_0, t_1, \ldots$ are equally spaced points in time with the spacing Δt, then Euler's method is

$$\boldsymbol{\alpha}^{k+1} = \boldsymbol{\alpha}^k + \Delta t C\boldsymbol{\alpha}^k, \qquad k = 0, 1, \ldots, \tag{8.4.11}$$

where $\boldsymbol{\alpha}^k$ denotes the approximate solution at the kth time step. In practice we will not carry out (8.4.11) in the manner indicated since we will not actually form $C = A^{-1}B$. Rather, we will work directly with the differential equation (8.4.7) and the corresponding Euler method

$$A\boldsymbol{\alpha}^{k+1} = A\boldsymbol{\alpha}^k + \Delta t B\boldsymbol{\alpha}^k, \qquad k = 0, 1, \ldots . \tag{8.4.12}$$

Thus at each time step we will solve the linear system of equations (8.4.12) with coefficient matrix A. The LU decomposition of A can, in this case, be done once and for all and the factors saved to be used at each time step.

In principle one could apply to (8.4.8) any of the higher-order methods discussed in Chapter 2; the formulation of some of these methods is left to Exercise 8.4.2. However, the equations (8.4.8) are typically rather "stiff," in the sense discussed in Chapter 2, and the use of explicit methods will require a rather small time step. Hence it may be advantageous to use a method such as the trapezoid rule. Applied to (8.4.8), the trapezoid rule (2.5.38) becomes

$$\boldsymbol{\alpha}^{k+1} = \boldsymbol{\alpha}^k + \frac{\Delta t}{2}(A^{-1}B\boldsymbol{\alpha}^{k+1} + A^{-1}B\boldsymbol{\alpha}^k),$$

or, multiplying through by A, collecting coefficients of $\boldsymbol{\alpha}^{k+1}$ and $\boldsymbol{\alpha}^k$, and setting $\mu = \Delta t/2$,

$$(A - \mu B)\boldsymbol{\alpha}^{k+1} = (A + \mu B)\boldsymbol{\alpha}^k. \tag{8.4.13}$$

To carry out this method requires the solution at each time step of a linear system similar to (8.4.12) but with the coefficient matrix $A - \mu B$. Assuming that this matrix is nonsingular and that Δt is held constant, we may proceed as before to compute the LU factors once and for all and use these in subsequent solutions of (8.4.13) at the different time steps. If the subtraction of the matrix μB from A does not materially affect the difficulty of computing the LU decomposition, then the work in carrying out (8.4.13) will not be much more than that for Euler's method. However, the trapezoid method is second-order accurate in time, as discussed in Chapter 2, and allows a larger time step for stiff equations; it should, therefore, be more suitable for this problem.

Hyperbolic Equations

The same approach can be applied to hyperbolic equations, and we illustrate this for the wave equation

$$u_{tt} = cu_{xx}, \qquad 0 \le x \le 1, \qquad t \ge 0, \tag{8.4.14}$$

with the initial and boundary conditions

$$u(0,x) = f(x), \qquad u_t(0,x) = g(x), \qquad u(t,0) = 0, \qquad u(t,1) = 0. \tag{8.4.15}$$

It is common practice to reduce a hyperbolic equation like (8.4.14) to a system of equations in which only the first derivative with respect to time appears. This is analogous to the situation for ordinary differential equations: recall that in Chapter 2 we reduced higher-order equations to first-order systems of ordinary differential equations (see Appendix 1). In the case of (8.4.14) we will use a reduction made by introducing a function $v(t,x)$ such that

$$u_t = av_x, \qquad v_t = au_x, \tag{8.4.16}$$

where $a = \sqrt{c}$. If u and v are solutions of the system (8.4.16) and are sufficiently differentiable, then by differentiating the first equation of (8.4.16) with respect to t and the second with respect to x, we obtain

$$u_{tt} = av_{xt} = av_{tx} = a^2 u_{xx} = cu_{xx},$$

so that u is a solution of (8.4.14). The initial and boundary conditions (8.4.15) for (8.4.14) become

$$\begin{aligned} &u(0,x) = f(x), \qquad v(0,x) = \frac{1}{a}\int_0^x g(s)ds, \\ &u(t,0) = 0, \qquad u(t,1) = 0. \end{aligned} \tag{8.4.17}$$

Now let $\phi_1(x), \ldots, \phi_n(x)$ and $\psi_1(x), \ldots, \psi_n(x)$ be two sets of basis functions that satisfy

$$\phi_i(0) = \phi_i(1) = \psi_i(0) = \psi_i(1) = 0, \qquad i = 1, \ldots, n. \tag{8.4.18}$$

We shall seek approximate solutions $\hat{u}$ and $\hat{v}$ of (8.4.16) of the form

$$\hat{u}(x,t) = \sum_{i=1}^{n} \alpha_i(t)\phi_i(x), \qquad \hat{v}(x,t) = \sum_{i=1}^{n} \beta_i(t)\psi_i(x). \tag{8.4.19}$$

If we require that these approximate solutions satisfy the equations (8.4.16) at the grid points $x_1, \ldots, x_n$, we obtain

$$\sum_{i=1}^{n} \alpha_i'(t)\phi_i(x_j) = a\sum_{j=1}^{n} \beta_i(t)\psi_i'(x_j), \qquad j = 1, \ldots, n, \tag{8.4.20a}$$

$$\sum_{i=1}^{n} \beta_i'(t)\psi_i(x_j) = a\sum_{i=1}^{n} \alpha_i(t)\phi_i'(x_j), \qquad j = 1, \ldots, n, \tag{8.4.20b}$$

which is a coupled system of $2n$ ordinary differential equations for the unknown functions $\alpha_1, \ldots, \alpha_n$ and $\beta_1, \ldots, \beta_n$. As before, the initial conditions are obtained from (8.4.17) by

$$\sum_{i=1}^{n} \alpha_i(0)\phi_i(x_j) = f(x_j), \quad \sum_{i=1}^{n} \beta_i(0)\psi_i(x_j) = \frac{1}{a}\int_0^x g(s)ds, \quad j = 1, \ldots, n,$$

where we will assume that the $n \times n$ matrices $(\phi_i(x_j))$ and $(\psi_i(x_j))$ are nonsingular. Thus the semidiscrete method for the wave equation (8.4.14) is entirely analogous to that for the heat equation, with the exception that there are now twice as many unknown functions in the system of ordinary differential equations.

The Method of Lines

We have used the collocation principle for the discretization of the space variable in both of the preceding examples, but the Galerkin principle discussed in Chapter 6 could have been used. Finite difference discretizations can also be used in the same fashion, as we now discuss briefly for the heat equation (8.4.1). If u is the exact solution of (8.4.1), then the approximate relation

$$u_t(t, x_i) \doteq \frac{c}{(\Delta x)^2}[u(t, x_{i+1}) - 2u(t, x_i) + u(t, x_{i-1})] \tag{8.4.21}$$

holds at the grid points $x_1, \ldots, x_n$. This leads to the following procedure. We seek n functions $v_1(t), \ldots, v_n(t)$ such that

$$v_i(t) \doteq u(t, x_i), \qquad i = 1, \ldots, n.$$

The approximate relationship (8.4.21) suggests attempting to find these functions as the solution of the system of ordinary differential equations

$$v_i'(t) = \frac{c}{(\Delta x)^2}[v_{i+1}(t) - 2v_i(t) + v_{i-1}(t)], \qquad i = 1, \ldots, n, \tag{8.4.22}$$

in which the functions v_0 and v_{n+1} are taken to be identically zero from the boundary conditions (8.4.2). Moreover, from the initial condition (8.4.2) we will take

$$v_i(0) = f(x_i), \qquad i = 1, \ldots, n. \tag{8.4.23}$$

The system (8.4.22) can be written in matrix form as

$$\mathbf{v}'(t) = \frac{-c}{(\Delta x)^2} A\mathbf{v}(t), \tag{8.4.24}$$

where A is the $(2, -1)$ tridiagonal matrix (3.1.10). If we apply Euler's method to this system, we have

$$\mathbf{v}^{m+1} = \mathbf{v}^m - \frac{c\Delta t}{(\Delta x)^2} A\mathbf{v}^m, \qquad m = 0, 1, \ldots .$$

Written out in component form, this is

$$v_i^{m+1} = v_i^m + \frac{c\Delta t}{(\Delta x)^2}(v_{i+1}^m - 2v_i^m + v_{i-1}^m), \qquad i = 1, \ldots, n, \qquad m = 0, 1, \ldots,$$

which is the explicit method (8.2.5). Similarly, the implicit method (8.3.2) is obtained by applying the backward Euler method (2.5.33) to (8.4.24), and the Crank-Nicolson method (8.3.12) arises by applying the trapezoid rule (2.5.38) to (8.4.24).

The above procedure leading to the system of ordinary differential equations (8.4.24) is called the *method of lines*, and some authors use this term for any semi-discrete method, whether or not it arises from finite differences.

Supplementary Discussion and References: 8.4

In the text we have described only the use of very simple methods for solving the ordinary differential equations arising from discretization only in the space variable. However, one of the advantages of this approach is the possibility of using high quality packages for solving the ordinary differential equations. For further reading, see Schiesser [1991].

EXERCISES 8.4

8.4.1. **a.** Write out the system of equations (8.4.7) explicitly for the basis functions $\phi_k(x) = \sin k\pi x$, $k = 1, \ldots, n$, assuming that the grid points $x_1, \ldots, x_n$ are equally spaced. Write a program to carry out Euler's method (8.4.12) with $n = 10$ and $\Delta t = 0.1$. Run the program for 20 time steps and for different initial conditions.

b. Do the same if the ϕ_k are quadratic splines as discussed in Section 6.2.

8.4.2. Write out explicitly the second- and fourth-order Runge-Kutta methods and the second- and fourth-order Adams-Bashforth methods for (8.4.8).

8.4.3. Write a program to carry out the trapezoid method (8.4.13) for the problems of Exercise 8.4.1.

8.4.4. Repeat Exercise 8.4.1 for the equations (8.4.20) for the wave equation, assuming that the ϕ_i and ψ_i are both trigonometric functions in part **a**, and quadratic splines in part **b**.