CHAPTER

THREE

BASICS OF DISCRETIZATION METHODS

3.1 INTRODUCTION

In this chapter, basic concepts and techniques needed in the formulation of finite-difference and finite-volume representations are developed. In the finite-difference approach, the continuous problem domain is "discretized," so that the dependent variables are considered to exist only at discrete points. Derivatives are approximated by differences, resulting in an algebraic representation of the partial differential equation (PDE). Thus a problem involving calculus has been transformed into an algebraic problem.

The nature of the resulting algebraic system depends on the character of the problem posed by the original PDE (or system of PDEs). Equilibrium problems usually result in a system of algebraic equations that must be solved simultaneously throughout the problem domain in conjunction with specified boundary values. Marching problems result in algebraic equations that usually can be solved one at a time (although it is often convenient to solve them several at a time). Several considerations determine whether the solution so obtained will be a good approximation to the exact solution of the original PDE. Among these considerations are truncation error, consistency, and stability, all of which will be discussed in the present chapter.

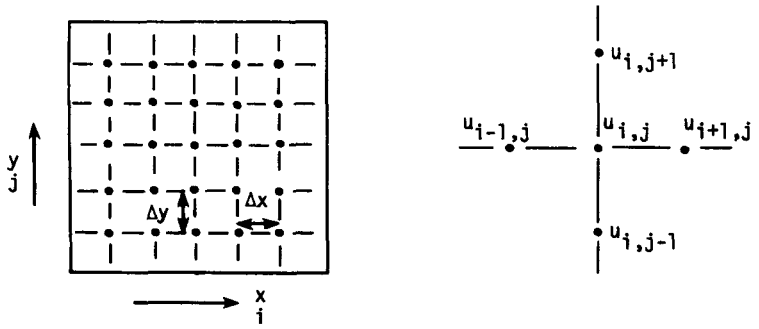

Figure 3.1 A typical finite-difference grid.

3.2 FINITE DIFFERENCES

One of the first steps to be taken in establishing a finite-difference procedure for solving a PDE is to replace the continuous problem domain by a finite difference mesh or grid. As an example, suppose that we wish to solve a PDE for which $u(x, y)$ is the dependent variable in the square domain $0 \leqslant x \leqslant 1$, $0 \leqslant y \leqslant 1$. We establish a grid on the domain by replacing $u(x, y)$ by $u(i\,\Delta x, j\,\Delta y)$. Points can be located according to values of i and j, so difference equations are usually written in terms of the general point (i, j) and its neighbors. This labeling is illustrated in Fig. 3.1. Thus, if we think of $u_{i,j}$ as $u(x_0, y_0)$, then

$$u_{i+1,j} = u(x_0 + \Delta x, y_0) \qquad u_{i-1,j} = u(x_0 - \Delta x, y_0)$$
$$u_{i,j+1} = u(x_0, y_0 + \Delta y) \qquad u_{i,j-1} = u(x_0, y_0 - \Delta y)$$

Often in the treatment of marching problems, the variation of the marching coordinate is indicated by a superscript, such as u_j^{n+1}, rather than a subscript. Many different finite-difference representations are possible for any given PDE and it is usually impossible to establish a "best" form on an absolute basis. First, the accuracy of a difference scheme may depend on the exact form of the equation and problem being solved, and second, our selection of a best scheme will be influenced by the aspect of the procedure that we are trying to optimize, i.e., accuracy, economy, or programming simplicity.

The idea of a finite-difference representation for a derivative can be introduced by recalling the definition of the derivative for the function $u(x, y)$ at $x = x_0$, $y = y_0$:

$$\frac{\partial u}{\partial x} = \lim_{\Delta x \to 0} \frac{u(x_0 + \Delta x, y_0) - u(x_0, y_0)}{\Delta x} \tag{3.1}$$

Here, if u is continuous, it is expected that $[u(x_0 + \Delta x, y_0) - u(x_0, y_0)]/\Delta x$ will be a "reasonable" approximation to $\partial u/\partial x$ for a "sufficiently" small but finite Δx. In fact, the mean-value theorem assures us that the difference

representation is exact for some point within the Δx interval. The difference approximation can be put on a more formal basis through the use of either a Taylor-series expansion or Taylor's formula with a remainder. Developing a Taylor-series expansion for $u(x_0 + \Delta x, y_0)$ about (x_0, y_0) gives

$$u(x_0 + \Delta x, y_0) = u(x_0, y_0) + \frac{\partial u}{\partial x}\bigg)_0 \Delta x + \frac{\partial^2 u}{\partial x^2}\bigg)_0 \frac{(\Delta x)^2}{2!} + \cdots + \frac{\partial^{n-1} u}{\partial x^{n-1}}\bigg)_0 \frac{(\Delta x)^{n-1}}{(n-1)!} + \frac{\partial^n u}{\partial x^n}\bigg)_\xi \frac{(\Delta x)^n}{n!}$$

$$x_0 \leqslant \xi \leqslant (x_0 + \Delta x) \quad (3.2)$$

where the last term can be identified as the remainder. Thus we can form the "forward" difference by rearranging Eq. (3.2):

$$\frac{\partial u}{\partial x}\bigg)_{x_0, y_0} = \frac{u(x_0 + \Delta x, y_0) - u(x_0, y_0)}{\Delta x} - \frac{\partial^2 u}{\partial x^2}\bigg)_0 \frac{\Delta x}{2!} - \cdots \quad (3.3)$$

Switching now to the i, j notation for brevity, we consider

$$\frac{\partial u}{\partial x}\bigg)_{i,j} = \frac{u_{i+1,j} - u_{i,j}}{\Delta x} + \text{T.E.} \quad (3.4)$$

where $(u_{i+1,j} - u_{i,j})/\Delta x$ is obviously the finite-difference representation for $\partial u/\partial x)_{i,j}$. The truncation error (T.E.) is the difference between the partial derivative and its finite-difference representation. We can characterize the limiting behavior of the T.E. by using the order of (O) notation, whereby we write

$$\frac{\partial u}{\partial x}\bigg)_{i,j} = \frac{u_{i+1,j} - u_{i,j}}{\Delta x} + O(\Delta x)$$

where $O(\Delta x)$ has a precise mathematical meaning. Here, when the T.E. is written as $O(\Delta x)$, we mean $|\text{T.E.}| \leqslant K|\Delta x|$ for $\Delta x \to 0$ (sufficiently small Δx), and K is a positive real constant. As a practical matter, the order of the T.E. in this case is found to be Δx raised to the largest power that is common to all terms in the T.E.

To give a more general definition of the O notation, when we say $f(x) = O[\phi(x)]$, we mean that there exists a positive constant K, independent of x, such that $|f(x)| \leqslant K|\phi(x)|$ for all x in S, where f and ϕ are real or complex functions defined in S. We often restrict S by $x \to \infty$ (sufficiently large x) or, as is most common in finite-difference applications, $x \to 0$ (sufficiently small x). More details on the O notation can be found in the work by Whittaker and Watson (1927).

Note that $O(\Delta x)$ tells us nothing about the exact size of the T.E., but rather how it behaves as Δx tends toward zero. If another difference expression had a T.E. $= O[(\Delta x)^2]$, we might expect or hope that the T.E. of the second repre-

sentation would be smaller than the first for a convenient Δx, but we could only be *sure* that this would be true if we refined the mesh "sufficiently," and "sufficiently" is a quantity that is hard to estimate.

An infinite number of difference representations can be found for $\partial u/\partial x)_{i,j}$. For example, we could expand "backward":

$$u(x_0 - \Delta x, y_0) = u(x_0, y_0) - \left.\frac{\partial u}{\partial x}\right)_0 \Delta x + \left.\frac{\partial^2 u}{\partial x^2}\right)_0 \frac{(\Delta x)^2}{2} - \left.\frac{\partial^3 u}{\partial x^3}\right)_0 \frac{(\Delta x)^3}{6} + \cdots \tag{3.5}$$

and obtain the backward-difference representation

$$\left.\frac{\partial u}{\partial x}\right)_{i,j} = \frac{u_{i,j} - u_{i-1,j}}{\Delta x} + O(\Delta x) \tag{3.6}$$

We can subtract Eq. (3.5) from Eq. (3.2), rearrange, and obtain the "central" difference

$$\left.\frac{\partial u}{\partial x}\right)_{i,j} = \frac{u_{i+1,j} - u_{i-1,j}}{2\,\Delta x} + O[(\Delta x)^2] \tag{3.7}$$

We can also add Eq. (3.2) and Eq. (3.5) and rearrange to obtain an approximation to the second derivative:

$$\left.\frac{\partial^2 u}{\partial x^2}\right)_{i,j} = \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{(\Delta x)^2} + O[(\Delta x)^2] \tag{3.8}$$

It should be emphasized that these are only a few examples of the possible ways in which first and second derivatives can be approximated.

It is convenient to utilize difference operators to represent finite differences when particular forms are used repetitively. Here we define the first forward difference of $u_{i,j}$ with respect to x at the point i, j as

$$\Delta_x u_{i,j} = u_{i+1,j} - u_{i,j} \tag{3.9}$$

Thus we can express the forward finite-difference approximation for the first partial derivative as

$$\left.\frac{\partial u}{\partial x}\right)_{i,j} = \frac{u_{i+1,j} - u_{i,j}}{\Delta x} + O(\Delta x) = \frac{\Delta_x u_{i,j}}{\Delta x} + O(\Delta x) \tag{3.10}$$

Similarly, derivatives with respect to other variables such as y can be represented by

$$\frac{\Delta_y u_{i,j}}{\Delta y} = \frac{u_{i,j+1} - u_{i,j}}{\Delta y}$$

The first backward difference of $u_{i,j}$ with respect to x at i, j is denoted by

$$\nabla_x u_{i,j} = u_{i,j} - u_{i-1,j} \tag{3.11}$$

It follows that the first backward-difference approximation to the first derivative can be written as

$$\left(\frac{\partial u}{\partial x}\right)_{i,j} = \frac{u_{i,j} - u_{i-1,j}}{\Delta x} + O(\Delta x) = \frac{\nabla_x u_{i,j}}{\Delta x} + O(\Delta x) \tag{3.12}$$

The central-difference operators $\bar{\delta}$, δ and δ^2 will be defined as

$$\bar{\delta}_x u_{i,j} = u_{i+1,j} - u_{i-1,j} \tag{3.13}$$

$$\delta_x u_{i,j} = u_{i+1/2,j} - u_{i-1/2,j} \tag{3.14}$$

$$\delta_x^2 u_{i,j} = \delta_x(\delta_x u_{i,j}) = u_{i+1,j} - 2u_{i,j} + u_{i-1,j} \tag{3.15}$$

and an averaging operator μ as

$$\mu_x u_{i,j} = \frac{u_{i+1/2,j} + u_{i-1/2,j}}{2} \tag{3.16}$$

Other convenient operators include the identity operator I and the shift operator E. The identity operator provides no operation, i.e., $Iu_{i,j} = u_{i,j}$. The shift operator advances the index associated with the subscripted variable by an amount indicated by the superscript. For example, $E_x^{-1} u_{i,j} = u_{i-1,j}$. When the superscript on E is $+1$, it is usually omitted. Difference representations can by indicated by using combinations of E and I, as for example,

$$\Delta_x u_{i,j} = (E_x - I)u_{i,j} = u_{i+1,j} - u_{i,j}$$

It is convenient to have specific operators for certain common central differences, although two of them can be easily expressed in terms of first-difference operators:

$$\bar{\delta}_x u_{i,j} = \Delta_x u_{i,j} + \nabla_x u_{i,j} \tag{3.17}$$

$$\delta_x^2 u_{i,j} = \Delta_x u_{i,j} - \nabla_x u_{i,j} = \Delta_x \nabla_x u_{i,j} \tag{3.18}$$

Using the newly defined operators, the central-difference representation for the first partial derivative can be written as

$$\left(\frac{\partial u}{\partial x}\right)_{i,j} = \frac{u_{i+1,j} - u_{i-1,j}}{2\,\Delta x} + O[(\Delta x)^2] = \frac{\bar{\delta}_x u_{i,j}}{2\,\Delta x} + O[(\Delta x)^2] \tag{3.19}$$

and the central-difference representation of the second derivative as

$$\left(\frac{\partial^2 u}{\partial x^2}\right)_{i,j} = \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{(\Delta x)^2} + O[(\Delta x)^2] = \frac{\delta_x^2 u_{i,j}}{(\Delta x)^2} + O[(\Delta x)^2] \tag{3.20}$$

Higher-order forward- and backward-difference operators are defined as

$$\Delta_x^n u_{i,j} = \Delta_x\left(\Delta_x^{n-1} u_{i,j}\right) \tag{3.21}$$

and

$$\nabla_x^n u_{i,j} = \nabla_x\left(\nabla_x^{n-1} u_{i,j}\right) \tag{3.22}$$

As an example, a forward second-derivative approximation is given by

$$\frac{\Delta_x^2 u_{i,j}}{(\Delta x)^2} = \frac{\Delta_x(u_{i+1,j} - u_{i,j})}{(\Delta x)^2} = \frac{u_{i+2,j} - u_{i+1,j} - u_{i+1,j} + u_{i,j}}{(\Delta x)^2}$$

$$= \frac{u_{i+2,j} - 2u_{i+1,j} + u_{i,j}}{(\Delta x)^2} = \left.\frac{\partial^2 u}{\partial x^2}\right)_{i,j} + O(\Delta x) \tag{3.23}$$

We can show that forward- and backward-difference approximations to derivatives of any order can be obtained from

$$\left.\frac{\partial^n u}{\partial x^n}\right)_{i,j} = \frac{\Delta_x^n u_{i,j}}{(\Delta x)^n} + O(\Delta x) \tag{3.24}$$

and

$$\left.\frac{\partial^n u}{\partial x^n}\right)_{i,j} = \frac{\nabla_x^n u_{i,j}}{(\Delta x)^n} + O(\Delta x) \tag{3.25}$$

Central-difference representations of derivatives of orders greater than the second can be expressed in terms of Δ and ∇ or δ. A more complete development on the use of difference operators can be found in many textbooks on numerical analysis such as that by Hildebrand (1956).

Most of the PDEs arising in fluid mechanics and heat transfer involve only first and second partial derivatives, and generally, we strive to represent these derivatives using values at only two or three grid points. Within these restrictions, the most frequently used first-derivative approximations on a grid for which $\Delta x = h = \text{const}$ are

$$\left.\frac{\partial u}{\partial x}\right)_{i,j} = \frac{u_{i+1,j} - u_{i,j}}{h} + O(h) \tag{3.26}$$

$$\left.\frac{\partial u}{\partial x}\right)_{i,j} = \frac{u_{i,j} - u_{i-1,j}}{h} + O(h) \tag{3.27}$$

$$\left.\frac{\partial u}{\partial x}\right)_{i,j} = \frac{u_{i+1,j} - u_{i-1,j}}{2h} + O(h^2) \tag{3.28}$$

$$\left.\frac{\partial u}{\partial x}\right)_{i,j} = \frac{-3u_{i,j} + 4u_{i+1,j} - u_{i+2,j}}{2h} + O(h^2) \tag{3.29}$$

$$\left.\frac{\partial u}{\partial x}\right)_{i,j} = \frac{3u_{i,j} - 4u_{i-1,j} + u_{i-2,j}}{2h} + O(h^2) \tag{3.30}$$

$$\left.\frac{\partial u}{\partial x}\right)_{i,j} = \frac{1}{2h}\left(\frac{\bar{\delta}_x u_{i,j}}{1 + \delta_x^2/6}\right) + O(h^4) \tag{3.31}$$

The most common three-point second-derivative approximations for a uniform grid, $\Delta x = h = \text{const}$, are

$$\left.\frac{\partial^2 u}{\partial x^2}\right)_{i,j} = \frac{u_{i,j} - 2u_{i+1,j} + u_{i+2,j}}{h^2} + O(h) \tag{3.32}$$

$$\left.\frac{\partial^2 u}{\partial x^2}\right)_{i,j} = \frac{u_{i,j} - 2u_{i-1,j} + u_{i-2,j}}{h^2} + O(h) \tag{3.33}$$

$$\left.\frac{\partial^2 u}{\partial x^2}\right)_{i,j} = \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2} + O(h^2) \tag{3.34}$$

$$\left.\frac{\partial^2 u}{\partial x^2}\right)_{i,j} = \frac{\delta_x^2 u_{i,j}}{h^2(1 + \delta_x^2/12)} + O(h^4) \tag{3.35}$$

The compact, three-point schemes given by Eqs. (3.31) and (3.35) having fourth-order T.E.s deserve a further word of explanation (see also Orszag and Israeli, 1974). Letting $\partial u/\partial x)_{i,j} = v_{i,j}$, Eq. (3.31) is to be interpreted as

$$\left(1 + \frac{\delta_x^2}{6}\right) v_{i,j} = \frac{\bar{\delta}_x u_{i,j}}{2h}$$

or

$$\frac{1}{6}(v_{i+1,j} + 4v_{i,j} + v_{i-1,j}) = \frac{\bar{\delta}_x u_{i,j}}{2h} \tag{3.36}$$

which provides an *implicit* formula for the derivative of interest, $v_{i,j}$. The $v_{i,j}$ can be determined from the $u_{i,j}$ by solving a tridiagonal system of simultaneous algebraic equations, which can usually be accomplished quite efficiently. Tridiagonal systems commonly occur in connection with the use of implicit difference schemes for second-order PDEs arising from marching problems and are defined and discussed in some detail in Chapter 4. For now it is sufficient to think of a tridiagonal system as the arrangement of unknowns that would occur if each difference equation in a system only involved a single unknown variable evaluated at three adjacent grid locations. The interpretation of Eq. (3.35) proceeds in a similar manner, providing an *implicit* representation of $\partial^2 u/\partial x^2)_{i,j}$. Some difference approximations for derivatives that involve more than three grid points are given in Table 3.1. For completeness, a few common difference representations for mixed partial derivatives are presented in Table 3.2. These will prove useful for schemes discussed in subsequent chapters. The mixed-derivative approximations in Table 3.2 can be verified by using the

Table 3.1 Difference approximations using more than three points

Derivative	Finite-difference representation	Equation
$\left(\frac{\partial^3 u}{\partial x^3}\right)_{i,j} =$	$\frac{u_{i+2,j} - 2u_{i+1,j} + 2u_{i-1,j} - u_{i-2,j}}{2h^3} + O(h^2)$	(3.38)
$\left(\frac{\partial^4 u}{\partial x^4}\right)_{i,j} =$	$\frac{u_{i+2,j} - 4u_{i+1,j} + 6u_{i,j} - 4u_{i-1,j} + u_{i-2,j}}{h^4} + O(h^2)$	(3.39)
$\left(\frac{\partial^2 u}{\partial x^2}\right)_{i,j} =$	$\frac{-u_{i+3,j} + 4u_{i+2,j} - 5u_{i+1,j} + 2u_{i,j}}{h^2} + O(h^2)$	(3.40)
$\left(\frac{\partial^3 u}{\partial x^3}\right)_{i,j} =$	$\frac{-3u_{i+4,j} + 14u_{i+3,j} - 24u_{i+2,j} + 18u_{i+1,j} - 5u_{i,j}}{2h^3} + O(h^2)$	(3.41)
$\left(\frac{\partial^2 u}{\partial x^2}\right)_{i,j} =$	$\frac{2u_{i,j} - 5u_{i-1,j} + 4u_{i-2,j} - u_{i-3,j}}{h^2} + O(h^2)$	(3.42)
$\left(\frac{\partial^3 u}{\partial x^3}\right)_{i,j} =$	$\frac{5u_{i,j} - 18u_{i-1,j} + 24u_{i-2,j} - 14u_{i-3,j} + 3u_{i-4,j}}{2h^3} + O(h^2)$	(3.43)
$\left(\frac{\partial u}{\partial x}\right)_{i,j} =$	$\frac{-u_{i+2,j} + 8u_{i+1,j} - 8u_{i-1,j} + u_{i-2,j}}{12h} + O(h^4)$	(3.44)
$\left(\frac{\partial^2 u}{\partial x^2}\right)_{i,j} =$	$\frac{-u_{i+2,j} + 16u_{i+1,j} - 30u_{i,j} + 16u_{i-1,j} - u_{i-2,j}}{12h^2} + O(h^4)$	(3.45)

Taylor-series expansion for two variables:

$$\begin{aligned} u(x_0 + \Delta x, y_0 + \Delta y) &= u(x_0, y_0) + \left(\Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y}\right) u(x_0, y_0) \\ &\quad + \frac{1}{2!}\left(\Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y}\right)^2 u(x_0, y_0) \\ &\quad + \cdots + \frac{1}{n!}\left(\Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y}\right)^n u(x_0 + \theta \Delta x, y_0 + \theta \Delta y) \end{aligned}$$

$$0 \leqslant \theta \leqslant 1 \quad (3.37)$$

3.3 DIFFERENCE REPRESENTATION OF PARTIAL DIFFERENTIAL EQUATIONS

3.3.1 Truncation Error

As a starting point in our study of T.E., let us consider the heat equation

$$\frac{\partial u}{\partial t} = \alpha \frac{\partial^2 u}{\partial x^2} \quad (3.55)$$

Table 3.2 Difference approximations for mixed partial derivatives

Derivative	Finite-difference representation	Equation
$\left.\frac{\partial^2 u}{\partial x\,\partial y}\right)_{i,j}$	$= \frac{1}{\Delta x}\left(\frac{u_{i+1,j}-u_{i+1,j-1}}{\Delta y} - \frac{u_{i,j}-u_{i,j-1}}{\Delta y}\right) + O(\Delta x, \Delta y)$	(3.46)
$\left.\frac{\partial^2 u}{\partial x\,\partial y}\right)_{i,j}$	$= \frac{1}{\Delta x}\left(\frac{u_{i,j+1}-u_{i,j}}{\Delta y} - \frac{u_{i-1,j+1}-u_{i-1,j}}{\Delta y}\right) + O(\Delta x, \Delta y)$	(3.47)
$\left.\frac{\partial^2 u}{\partial x\,\partial y}\right)_{i,j}$	$= \frac{1}{\Delta x}\left(\frac{u_{i,j}-u_{i,j-1}}{\Delta y} - \frac{u_{i-1,j}-u_{i-1,j-1}}{\Delta y}\right) + O(\Delta x, \Delta y)$	(3.48)
$\left.\frac{\partial^2 u}{\partial x\,\partial y}\right)_{i,j}$	$= \frac{1}{\Delta x}\left(\frac{u_{i+1,j+1}-u_{i+1,j}}{\Delta y} - \frac{u_{i,j+1}-u_{i,j}}{\Delta y}\right) + O(\Delta x, \Delta y)$	(3.49)
$\left.\frac{\partial^2 u}{\partial x\,\partial y}\right)_{i,j}$	$= \frac{1}{\Delta x}\left(\frac{u_{i+1,j+1}-u_{i+1,j-1}}{2\Delta y} - \frac{u_{i,j+1}-u_{i,j-1}}{2\Delta y}\right) + O[\Delta x, (\Delta y)^2]$	(3.50)
$\left.\frac{\partial^2 u}{\partial x\,\partial y}\right)_{i,j}$	$= \frac{1}{\Delta x}\left(\frac{u_{i,j+1}-u_{i,j-1}}{2\Delta y} - \frac{u_{i-1,j+1}-u_{i-1,j-1}}{2\Delta y}\right) + O[\Delta x, (\Delta y)^2]$	(3.51)
$\left.\frac{\partial^2 u}{\partial x\,\partial y}\right)_{i,j}$	$= \frac{1}{2\Delta x}\left(\frac{u_{i+1,j+1}-u_{i+1,j-1}}{2\Delta y} - \frac{u_{i-1,j+1}-u_{i-1,j-1}}{2\Delta y}\right) + O[(\Delta x)^2, (\Delta y)^2]$	(3.52)
$\left.\frac{\partial^2 u}{\partial x\,\partial y}\right)_{i,j}$	$= \frac{1}{2\Delta x}\left(\frac{u_{i+1,j+1}-u_{i+1,j}}{\Delta y} - \frac{u_{i-1,j+1}-u_{i-1,j}}{\Delta y}\right) + O[(\Delta x)^2, \Delta y]$	(3.53)
$\left.\frac{\partial^2 u}{\partial x\,\partial y}\right)_{i,j}$	$= \frac{1}{2\Delta x}\left(\frac{u_{i+1,j}-u_{i+1,j-1}}{\Delta y} - \frac{u_{i-1,j}-u_{i-1,j-1}}{\Delta y}\right) + O[(\Delta x)^2, \Delta y]$	(3.54)

Using a forward-difference representation for the time derivative ($t = n\Delta t$) and a central-difference representation for the second derivative, we can approximate the heat equation by

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \frac{\alpha}{(\Delta x)^2}(u_{j+1}^n - 2u_j^n + u_{j-1}^n) \tag{3.56a}$$

However, we noted in Section 3.2 that T.E.s were associated with the forward- and central-difference representations used in Eq. (3.56*a*). If we rearrange Eq. (3.55) to put zero on the right-hand side and include the T.E.s associated with the difference representation of the derivatives, we obtain

$$\underbrace{\frac{\partial u}{\partial t} - \alpha\frac{\partial^2 u}{\partial x^2}}_{\text{PDE}} = \underbrace{\frac{u_j^{n+1} - u_j^n}{\Delta t} - \frac{\alpha}{(\Delta x)^2}(u_{j+1}^n - 2u_j^n + u_{j-1}^n)}_{\text{FDE}} + \underbrace{\left[-\left.\frac{\partial^2 u}{\partial t^2}\right)_{n,j}\frac{\Delta t}{2} + \alpha\left.\frac{\partial^4 u}{\partial x^4}\right)_{n,j}\frac{(\Delta x)^2}{12} + \cdots\right]}_{\text{T.E.}} \tag{3.56b}$$

where PDE is the partial differential equation and FDE is the finite-difference equation. The T.E.s associated with all derivatives in any one PDE should be obtained by expanding about the same point (n, j in the above discussion).

The difference representation given by Eq. (3.56*a*) will be referred to as the *simple explicit scheme* for the heat equation. An *explicit* scheme is one for which only one unknown appears in the difference equation in a manner that permits evaluation in terms of known quantities. Since the parabolic heat equation governs a marching problem for which an initial distribution of u must be specified, u at the time level n can be considered as known. If the second-derivative term in the heat equation was approximated by u at the $n + 1$ time level, three unknowns would appear in the difference equation, and the procedure would be known as *implicit*, indicating that the algebraic formulation would require the simultaneous solution of several equations involving the unknowns. The differences between implicit and explicit schemes are discussed further in Chapter 4.

The quantity in brackets (note that only the leading terms have been written out utilizing Taylor-series expansions) in Eq. (3.56*b*) is identified as the *truncation error* for this finite-difference representation of the heat equation and is defined as the difference between the PDE and the difference approximation to it. That is, T.E. = PDE − FDE. The *order* of the T.E. in this case is $O(\Delta t) + O[(\Delta x)^2]$, which is frequently expressed in the form $O[\Delta t, (\Delta x)^2]$. Naturally, we solve only the finite-difference equations and hope that the T.E. is small. If we do not feel a little uneasy at this point, perhaps we should. How do we know that our difference representation is acceptable and that a marching solution technique will work in the sense of giving us an approximate solution to the PDE? In order to be acceptable, our difference representation for this marching problem needs to meet the conditions of *consistency* and *stability*.

3.3.2 Round-Off and Discretization Errors

Any computed solution, including sometimes an "exact" analytic solution to a PDE, may be affected by rounding to a finite number of digits in the arithmetic operations. These errors are called *round-off errors*, and we are especially aware of their existence in obtaining machine solutions to finite-difference equations because of the large number of dependent, repetitive operations that are usually involved. In some types of calculations, the magnitude of the round-off error is proportional to the number of grid points in the problem domain. In these cases, refining the grid may decrease the T.E. but increase the round-off error.

Discretization error is the error in the solution to the PDE caused by replacing the continuous problem by a discrete one and is defined as the difference between the exact solution of the PDE (round-off free) and the exact solution of the FDEs (round-off free). In terms of the definitions developed thus far, the difference between the exact solution of the PDE and the computer solution to the FDEs would be equal to the sum of the discretization error and

the round-off error associated with the finite-difference calculation. We can also observe that the discretization error is the error in the solution that is caused by the T.E. in the difference representation of the PDE plus any errors introduced by the treatment of boundary conditions.

3.3.3 Consistency

Consistency deals with the extent to which the FDEs approximate the PDEs. The difference between the PDE and the finite-difference approximation has already been defined as the T.E. of the difference representation. A finite-difference representation of a PDE is said to be consistent if we can show that the difference between the PDE and its difference representation vanishes as the mesh is refined, i.e., $\lim_{\text{mesh}\to 0}(\text{PDE} - \text{FDE}) = \lim_{\text{mesh}\to 0}(\text{T.E.}) = 0$. This should always be the case if the order of the T.E. vanishes under grid refinement. An example of a questionable scheme would be one for which the T.E. was $O(\Delta t/\Delta x)$, where the scheme would not formally be consistent unless the mesh were refined in a manner such that $\Delta t/\Delta x \to 0$. The DuFort-Frankel (DuFort and Frankel, 1953) differencing of the heat equation,

$$\frac{u_j^{n+1} - u_j^{n-1}}{2\,\Delta t} = \frac{\alpha}{(\Delta x)^2}\left(u_{j+1}^n - u_j^{n+1} - u_j^{n-1} + u_{j-1}^n\right) \tag{3.57}$$

for which the leading terms in the T.E. are

$$+\frac{\alpha}{12}\frac{\partial^4 u}{\partial x^4}\Bigg)_{n,j}(\Delta x)^2 - \alpha\frac{\partial^2 u}{\partial t^2}\Bigg)_{n,j}\left(\frac{\Delta t}{\Delta x}\right)^2 - \frac{1}{6}\frac{\partial^3 u}{\partial t^3}\Bigg)_{n,j}(\Delta t)^2$$

serves as an example. All is well if

$$\lim_{\Delta t,\,\Delta x\to 0}\left(\frac{\Delta t}{\Delta x}\right) = 0$$

but if Δt and Δx were to approach zero at the same rate, such that $\Delta t/\Delta x = \beta$, then the DuFort-Frankel scheme is consistent with the hyperbolic equation

$$\frac{\partial u}{\partial t} + \alpha\beta^2\frac{\partial^2 u}{\partial t^2} = \alpha\frac{\partial^2 u}{\partial x^2}$$

3.3.4 Stability

Numerical stability is a concept applicable in the strict sense only to marching problems. A stable numerical scheme is one for which errors from any source (round-off, truncation, mistakes) are not permitted to grow in the sequence of numerical procedures as the calculation proceeds from one marching step to the next. Generally, concern over stability occupies much more of our time and energy than does concern over consistency. Consistency is relatively easy to check, and most schemes that are conceived will be consistent just owing to the

methodology employed in their development. Stability is much more subtle, and usually a bit of hard work is required in order to establish analytically that a scheme is stable. More detail is presented in Section 3.6, and some very workable methods will be developed for establishing the stability limits for linear PDEs. It will be possible to extend these guidelines to nonlinear equations in an approximate sense.

Using these guidelines, the DuFort-Frankel scheme, Eq. (3.57), for the heat equation would be found to be unconditionally stable, whereas the simple explicit scheme would be stable only if $r = [\alpha \Delta t/(\Delta x)^2] \leqslant \frac{1}{2}$. This restriction would limit the size of the marching step permitted for any specific spatial mesh.

A scheme using a central time difference and having a more favorable T.E. of $O[(\Delta t)^2, (\Delta x)^2]$,

$$\frac{u_j^{n+1} - u_j^{n-1}}{2\,\Delta t} = \frac{\alpha}{(\Delta x)^2}(u_{j+1}^n - 2u_j^n + u_{j-1}^n) \tag{3.58}$$

is unconditionally unstable and therefore cannot be used for real calculations despite the fact that it looks to be more accurate, in terms of T.E., than the ones given previously that will work.

Sometimes instability can be identified with a physical implausibility. That is, conditions that would result in an unstable numerical procedure would also imply unacceptable modeling of physical processes. To illustrate this, we rearrange the simple explicit representation of the heat equation, Eq. (3.56*a*), so that the unknown appears on the left. Letting $r = \alpha \Delta t/(\Delta x)^2$, our difference equation becomes

$$u_j^{n+1} = r(u_{j+1}^n + u_{j-1}^n) + (1 - 2r)u_j^n \tag{3.59}$$

Suppose that at time t, $u_{j+1}^n = u_{j-1}^n = 100°C$ and $u_j^n = 0°C$. This arrangement is shown in Fig. 3.2. If $r > \frac{1}{2}$, we see that the temperature at point j at time level $n + 1$ will exceed the temperature at the two surrounding points at time level n. This seems unreasonable, since we expect heat to flow from the warmer region to a colder region but not vice versa. The maximum temperature that we would expect to find at point j at time level $n + 1$ is 100°C. If $r = 1$, for example, u_j^{n+1} would equal 200°C by Eq. (3.59).

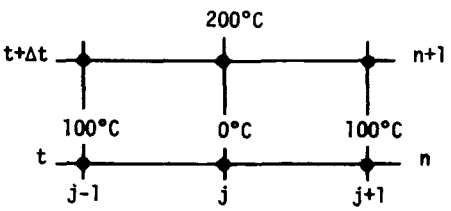

Figure 3.2 Physical implausibility resulting from $r = 1$.

3.3.5 Convergence for Marching Problems

Generally, we find that a consistent, stable scheme is convergent. Convergence here means that the solution to the finite-difference equation approaches the true solution to the PDE having the same initial and boundary conditions as the mesh is refined. A proof of this is available for initial value (marching) problems governed by linear PDEs. The theorem, due to Lax (see Richtmyer and Morton, 1967) is stated here without proof.

> **Lax's equivalence theorem:** Given a properly posed initial value problem and a finite-difference approximation to it that satisfies the consistency condition, stability is the necessary and sufficient condition for convergence.

We might add that most computational work proceeds as though this theorem applies also to nonlinear PDEs, although the theorem has never been proven for this more general category of equations.

3.3.6 A Comment on Equilibrium Problems

Throughout our discussion of stability and convergence, the focus was on marching problems (parabolic and hyperbolic PDEs). Despite this emphasis on initial value problems, most of the material presented in this chapter also applies to equilibrium problems. The exception is the concept of stability. We should observe, however, that the important concept of consistency applies to difference representations of PDEs of all classes.

The "convergence" of the solution of the difference equation to the exact solution of the PDE might be aptly termed truncation or discretization convergence. The solution to equilibrium problems (elliptic equations) leads us to a system of simultaneous algebraic equations that needs to be solved only once, rather than in a marching manner. Thus the concept of stability developed previously is not directly applicable as stated. To achieve "truncation convergence" for equilibrium problems, it would seem that it is only necessary to devise a solution scheme in which the error in solving the simultaneous algebraic equations can be controlled as the mesh size is refined without limit. Many common schemes are iterative (Gauss-Seidel iteration is one example) in nature, and for these we want to ensure that the iterative process converges. Here convergence means that the iterative process is repeated until the magnitude of the difference between the function at the $k + 1$ and the k iteration levels is as small as we wish for each grid point, i.e., $|u_{i,j}^{k+1} - u_{i,j}^{k}| < \epsilon$. This is known as *iteration convergence*. It would appear that (no proof can be cited) truncation convergence will be assumed for a consistent representation to an equilibrium problem if it can be shown that the iterative method of solution converges even for arbitrarily small choices of mesh sizes.

It is possible to use direct (noniterative) methods to solve the algebraic equations associated with equilibrium problems. For these methods we would want to be sure that the errors inherent in the method, especially round-off

errors, do not get out of control as the mesh is refined and the number of points tends toward infinity.

In closing this section, we should mention that there are aspects to the iterative solution of equilibrium problems that resemble the marching process in initial value problems and a sense in which stability concerns in the marching problems correspond to iterative convergence concerns in the solution to equilibrium problems.

3.3.7 Conservation Form and Conservative Property

Two different ideas will be discussed in this section. The first has to do with the PDEs themselves. The terms "conservation form," "conservation-law form," "conservative form," and "divergence form" are all equivalent, and PDEs possessing this form have the property that the coefficients of the derivative terms are either constant or, if variable, their derivatives appear nowhere in the equation. Normally, for the PDEs that represent a physical conservation statement, this means that the divergence of a physical quantity can be identified in the equation. If all spatial derivative terms of an equation can be identified as divergence terms, the equation is said to be in "strong conservation-law form." As an example, the conservative form of the equation for mass conservation (continuity equation) is

$$\frac{\partial \rho}{\partial t} + \frac{\partial \rho u}{\partial x} + \frac{\partial \rho v}{\partial y} + \frac{\partial \rho w}{\partial z} = 0 \tag{3.60}$$

which can be written in vector notation as

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \rho \mathbf{V} = 0$$

A nonconservative or nondivergence form would be

$$\frac{\partial \rho}{\partial t} + u\frac{\partial \rho}{\partial x} + \rho\frac{\partial u}{\partial x} + v\frac{\partial \rho}{\partial y} + \rho\frac{\partial v}{\partial y} + w\frac{\partial \rho}{\partial z} + \rho\frac{\partial w}{\partial z} = 0 \tag{3.61}$$

As a second example, we consider the one-dimensional (1-D) heat conduction equation for a substance whose density ρ, specific heat c, and thermal conductivity k all vary with position. The conservative form of this equation is

$$\rho c\frac{\partial T}{\partial t} = \frac{\partial}{\partial x}\left(k\frac{\partial T}{\partial x}\right) \tag{3.62}$$

whereas a nonconservative form would be

$$\rho c\frac{\partial T}{\partial t} = k\frac{\partial^2 T}{\partial x^2} + \frac{\partial k}{\partial x}\frac{\partial T}{\partial x} \tag{3.63}$$

In Eq. (3.62) the right-hand side can be identified as the negative of the divergence of the heat flux vector specialized for 1-D conduction. A difference formulation based on a PDE in nondivergence form may lead to numerical

difficulties in situations where the coefficients may be discontinuous, as in flows containing shock waves.

The second idea to be developed in this section deals with the *conservative property of a finite-difference representation.* The PDEs of interest in this book all have their basis in physical laws, such as the conservation of mass, momentum, and energy. Such a PDE represents a conservation statement at a point. We strive to construct finite-difference representations that provide a good approximation to the PDE in a small, local neighborhood involving a few grid points. The same conservation principles that gave rise to the PDEs also apply to arbitrarily large regions (control volumes). In fact, in deriving the PDEs, we usually start with the control-volume form of the conservation statement. If our finite-difference representation approximates the PDE closely in the neighborhood of each grid point, then we have reason to expect that the related conservation statement will be approximately enforced over a larger control volume containing a large number of grid points in the interior. Those finite-difference schemes that maintain the discretized version of the conservation statement exactly (except for round-off errors) for any mesh size over an arbitrary finite region containing any number of grid points is said to have the *conservative property.* For some problems this property is crucial.

The key word in the definition above is "exactly." All consistent schemes should approximately enforce the appropriate conservation statement over large regions, but schemes having the conservative property do so exactly (except for round-off errors) because of exact cancellation of terms. To illustrate this concept, we will consider a problem requiring the solution of the continuity equation for steady flow. The PDE can be written as

$$\nabla \cdot \rho \mathbf{V} = 0$$

We will assume that the PDE is approximated by a suitable finite-difference representation and solved throughout the flow. For an arbitrary control volume that could include the entire problem domain or any fraction of it, conservation of mass for steady flow requires that the net mass efflux be zero (mass flow rate in equals mass flow rate out). This is observed formally by applying the divergence theorem to the governing PDE,

$$\iiint_R \nabla \cdot \rho \mathbf{V}\, dR = \iint_S \rho \mathbf{V} \cdot \mathbf{n}\, dS = 0$$

To see if the finite-difference representation for the PDE has the conservative property, we must establish that the discretized version of the divergence theorem is satisfied. We normally check this for a control volume consisting of the entire problem domain. To do this, the integral on the left is evaluated by summing the difference representation of the PDE at all grid points. If the difference scheme has the conservative property, all terms will cancel except those that represent fluxes at the boundaries. This is sometimes referred to as

the "telescoping property." It should be possible to rearrange the remaining terms to obtain identically a finite-difference representation of the integral on the right. For this example the result will be a verification that the mass flux into the control volume equals the mass flux out. If the difference scheme used for the PDE is not conservative, the numerical solution may permit the existence of small mass sources or sinks.

Schemes having the conservative property occur in a natural way when differencing starts with the divergence form of the PDE. For some equations and problems, the divergence form is not an appropriate starting point. For these situations, use of a control-volume method (Section 3.4.4) for obtaining the difference scheme is helpful. This difference representation will usually have the conservative property if care is taken to ensure that the expressions used to represent fluxes across the interface of two adjacent control volumes are the same in the difference form of the conservation statement for each of the two control volumes.

The conservative property issue has been actively discussed and debated over the short history of computational fluid mechanics and heat transfer. However, the conservative property is not the only important figure of merit for a difference representation. PDEs represent more than a conservation statement at a point. As shown by solution forms in Chapter 2, PDEs also contain information on characteristic directions and domains of dependence. Proper representation of this information is also important. Many useful finite-difference equations do not have the conservative property and, in a few instances, prove to be more accurate in some sense than those that do. The importance of maintaining the conservation statement with high accuracy over a finite region is highly problem dependent. All consistent formulations, whether or not they have the conservative property, can provide an adequate representation for most problems if the grid is refined sufficiently.

3.4 FURTHER EXAMPLES OF METHODS FOR OBTAINING FINITE-DIFFERENCE EQUATIONS

As we start with a given PDE and a finite-difference mesh, several procedures are available to us for developing finite-difference equations. Among these are

1. Taylor-series expansions
2. polynomial fitting
3. integral method (called the micro-integral method by some)
4. finite-volume (control-volume) approach

It is sometimes possible to obtain exactly the same finite-difference representation by using all four methods. In our introduction to the subject, we will lean most heavily on the use of Taylor-series expansions, utilizing polynomial fitting on occasion in treating boundary conditions.

3.4.1 Use of Taylor Series

We now demonstrate how one might proceed on a slightly more formal basis with Taylor-series expansions to develop difference expressions satisfying specified constraints. Suppose we want to develop a difference approximation for $\partial u/\partial x)_{i,j}$ having a T.E. of $O[(\Delta x)^2]$ using at most values $u_{i-2,j}$, $u_{i-1,j}$, and $u_{i,j}$.

With these constraints and objectives, it would appear logical to write Taylor-series expressions for $u_{i-2,j}$ and $u_{i-1,j}$ expanding about the point (i, j) and attempt to solve for $\partial u/\partial x)_{i,j}$ from the resulting equations in such a way as to obtain a T.E. of $O[(\Delta x)^2]$:

$$u_{i-2,j} = u_{i,j} + \frac{\partial u}{\partial x}\bigg)_{i,j}(-2\,\Delta x) + \frac{\partial^2 u}{\partial x^2}\bigg)_{i,j}\frac{(2\,\Delta x)^2}{2!} + \frac{\partial^3 u}{\partial x^3}\bigg)_{i,j}\frac{(-2\,\Delta x)^3}{3!} + \cdots \tag{3.64}$$

$$u_{i-1,j} = u_{i,j} + \frac{\partial u}{\partial x}\bigg)_{i,j}(-\Delta x) + \frac{\partial^2 u}{\partial x^2}\bigg)_{i,j}\frac{(\Delta x)^2}{2!} + \frac{\partial^3 u}{\partial x^3}\bigg)_{i,j}\frac{(-\Delta x)^3}{3!} + \cdots \tag{3.65}$$

It is often possible to determine the required form of the difference representation by inspection or simple substitution. To proceed by substitution, we will rearrange Eq. (3.64) to put $\partial u/\partial x)_{i,j}$ on the left-hand side, such that

$$\frac{\partial u}{\partial x}\bigg)_{i,j} = \frac{u_{i,j}}{2\,\Delta x} - \frac{u_{i-2,j}}{2\,\Delta x} + \frac{\partial^2 u}{\partial x^2}\Delta x + O[(\Delta x)^2]$$

As is, the representation is $O(\Delta x)$ because of the term $(\partial^2 u/\partial x^2)\,\Delta x$. We can substitute for $\partial^2 u/\partial x^2$ in the above equation using Eq. (3.65) to obtain the desired result. A more formal procedure to obtain the desired expression is sometimes useful. To proceed more formally, we first multiply Eq. (3.64) by a and Eq. (3.65) by b and add the two equations. If $-2a - b = 1$, then the coefficient of $\partial u/\partial x)_{i,j}\,\Delta x$ will be 1 after the addition, and if $2a + b/2 = 0$, then the terms involving $\partial^2 u/\partial x^2)_{i,j}$, which would contribute a T.E. of $O(\Delta x)$ to the final result, will be eliminated. A solution to the equations

$$-2a - b = 1 \qquad 2a + \frac{b}{2} = 0$$

is given by $a = \frac{1}{2}$, $b = -2$. Thus, if we multiply Eq. (3.64) by $\frac{1}{2}$, Eq. (3.65) by -2, add the results, and solve for $\partial u/\partial x)_{i,j}$, we obtain

$$\frac{\partial u}{\partial x}\bigg)_{i,J} = \frac{u_{i-2,j} - 4u_{i-1,j} + 3u_{i,j}}{2\,\Delta x} + O[(\Delta x)^2]$$

which can be recognized as Eq. (3.30). A careful check on the details of this example will reveal that it was really necessary to include terms involving $\partial^3 u/\partial x^3)_{i,j}$ in the Taylor-series expansions in order to determine whether or

not these terms would cancel in the algebraic operations and reduce the T.E. even further to $O[(\Delta x)^3]$. Fortuitous cancellation of terms occurs frequently enough to warrant close attention to this point.

We should observe that it is sometimes necessary to carry out the inverse of the above process. That is, suppose we had obtained the approximation represented by Eq. (3.30) by some other means and we wanted to investigate the consistency and T.E. of such an expression. For this, the use of Taylor-series expansions would be invaluable, and the recommended procedure would be to substitute the Taylor-series expressions from Eq. (3.64) and Eq. (3.65) above for $u_{i-2,j}$ and $u_{i-1,j}$ into the difference representation to obtain an expression of the form $\partial u/\partial x)_{i,j}$ + T.E. on the right-hand side. At this point, the T.E. has been identified, and if $\lim_{\Delta x \to 0}$(T.E.) = 0, the difference representation is consistent.

As a slightly more complex example, we will develop a finite-difference approximation with T.E. of $O[(\Delta y)^2]$ for $\partial u/\partial y$ at point (i, j) using at most $u_{i,j}, u_{i,j+1}, u_{i,j-1}$ when the grid spacing is not uniform. We will adopt the notation that $\Delta y_+ = y_{i,j+1} - y_{i,j}$ and $\Delta y_- = y_{i,j} - y_{i,j-1}$, as indicated in Fig. 3.3.

We recall that for equal spacing, the central-difference representation for a first derivative was equivalent to the arithmetic average of a forward and backward representation. That is, for $\Delta y_+ = \Delta y_- = \Delta y$,

$$\left.\frac{\partial u}{\partial y}\right)_{i,j} = \frac{\bar{\delta}_y u_{i,j}}{2\,\Delta y} = \frac{\Delta_y u_{i,j} + \nabla_y u_{i,j}}{2\,\Delta y} + O\left[(\Delta y)^2\right]$$

We might wonder if, for unequal spacing, use of a geometrically weighted average will preserve the second-order accuracy:

$$\left.\frac{\partial u}{\partial y}\right)_{i,j} \stackrel{?}{=} \frac{\Delta_y u_{i,j}}{\Delta y_+}\left(\frac{\Delta y_-}{\Delta y_+ + \Delta y_-}\right) + \frac{\nabla_y u_{i,j}}{\Delta y_-}\left(\frac{\Delta y_+}{\Delta y_+ + \Delta y_-}\right) + O\left[(\Delta y)^2\right] \quad (3.66)$$

The truth of the above statement may be evident to some, but it can be verified from basics by use of Taylor-series expansions about point (i, j). Letting $\Delta y_+/\Delta y_- = \alpha$, and adopting the more compact subscript notation to denote differentiation, $u_y = \partial u/\partial y)_{i,j}$, $u_{yy} = \partial^2 u/\partial y^2)_{i,j}$, etc., we obtain

$$u_{i,j+1} = u_{i,j} + u_y \alpha\,\Delta y_- + u_{yy}\frac{(\alpha\,\Delta y_-)^2}{2!} + u_{yyy}\frac{(\alpha\,\Delta y_-)^3}{3!} + u_{yyyy}\frac{(\alpha\,\Delta y_-)^4}{4!} + \cdots \quad (3.67)$$

$$u_{i,j-1} = u_{i,j} + u_y(-\Delta y_-) + u_{yy}\frac{(-\Delta y_-)^2}{2!} + u_{yyy}\frac{(-\Delta y_-)^3}{3!} + u_{yyyy}\frac{(-\Delta y_-)^4}{4!} + \cdots \quad (3.68)$$

As before, we will multiply Eq. (3.67) by a and Eq. (3.68) by b, add the results, and solve for $\partial u/\partial y)_{i,j}$. Requiring that the coefficient of $\partial u/\partial y)_{i,j}\Delta y_-$

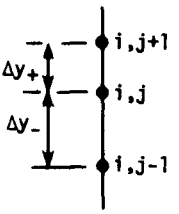

Figure 3.3 Notation for unequal y spacing.

be equal to 1 after the addition, gives $a\alpha - b = 1$. For the final result to have a T.E. of $O[(\Delta y)^2]$ or better, the coefficient of u_{yy} must be zero after the addition, which requires that $\alpha^2 a + b = 0$. A solution to these two algebraic equations can be obtained readily as $a = 1/\alpha(\alpha + 1)$, $b = -\alpha/(\alpha + 1)$. Thus

$$\left.\frac{\partial u}{\partial y}\right)_{i,j} = \frac{a \times \text{Eq. (3.67)} + b \times \text{Eq. (3.68)}}{\Delta y_-} + O\left[(\Delta y)^2\right]$$

The final result can be written as

$$\left.\frac{\partial u}{\partial y}\right)_{i,j} = \frac{u_{i,j+1} + (\alpha^2 - 1)u_{i,j} - \alpha^2 u_{i,j-1}}{\alpha(\alpha + 1)\Delta y_-} \tag{3.69}$$

which can be rearranged further into the form given by Eq. (3.66).

Our Taylor-series examples thus far have illustrated procedures for obtaining a finite-difference approximation to a single derivative. However, our main interest is in correctly approximating an entire PDE at an arbitrary point in the problem domain. For this reason, we must be careful to use the same expansion point in approximating all derivatives in the PDE by the Taylor-series method. If this is done, then the T.E. for the entire equation can be obtained by adding the T.E. for each derivative.

There is no requirement that the expansion point be (i, j), as indicated by the following examples, where the order of the T.E. and the most convenient expansion points are indicated. The geometric arrangement of points used in the difference equation is indicated by the sketch of the difference "molecule."

Fully implicit form for the heat equation, Eq. (3.55):

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \frac{\alpha}{(\Delta x)^2}\left(u_{j+1}^{n+1} - 2u_j^{n+1} + u_{j-1}^{n+1}\right) \qquad \text{T.E.} = O[\Delta t, (\Delta x)^2] \tag{3.70}$$

The difference molecule for this scheme is shown in Fig. 3.4, and point $(n + 1, j)$ is indicated as the most convenient expansion point.

Crank-Nicholson form for the heat equation:

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \frac{\alpha}{2(\Delta x)^2}\left[u_{j+1}^{n+1} + u_{j+1}^n - 2\left(u_j^{n+1} + u_j^n\right) + u_{j-1}^{n+1} + u_{j-1}^n\right] \tag{3.71a}$$

$$\text{T.E.} = O[(\Delta t)^2, (\Delta x)^2]$$

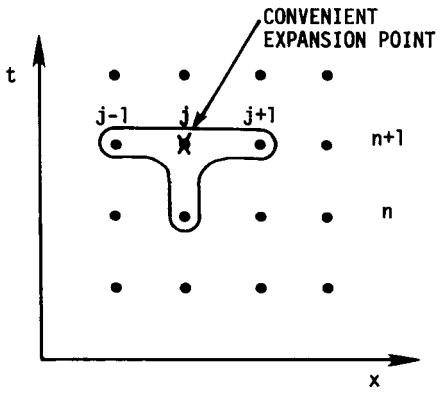

Figure 3.4 Difference molecule, fully implicit form for heat equation.

The difference molecule for the Crank-Nicolson scheme is shown in Fig. 3.5, and point $(n + \frac{1}{2}, j)$ is designated as the most convenient expansion point.

It is interesting to note that the *order* of the T.E. for difference representations of a complete PDE (not a single derivative term, however) is not dependent upon the choice of expansion point in the evaluation of this error by the Taylor-series method. We will demonstrate this point by considering the Crank-Nicolson scheme. The T.E. for the Crank-Nicolson scheme was most conveniently determined by expanding about the point $(n + \frac{1}{2}, j)$ to obtain the results stated above. Using this point resulted in the elimination of the maximum number of terms from the Taylor series by cancellation. Had we used point (n, j) or even $(n - 1, j)$ as the expansion point, the conclusion on the order of the T.E. would have been the same. To reach this conclusion, however, we often must examine the T.E. very carefully. To illustrate, evaluating the T.E. of the Crank-Nicolson scheme by using expansions for $u^n_{j-1}, u^n_{j+1}, u^{n+1}_{j-1}, u^{n+1}_{j+1}, u^{n+1}_j$ about point (n, j) in Eq. (3.71*a*) gives, after rearrangement,

$$u_t - \alpha u_{xx} = -u_{tt}\frac{\Delta t}{2} + \alpha u_{txx}\frac{\Delta t}{2} + O[(\Delta x)^2] + O[(\Delta t)^2] \qquad (3.71b)$$

At first glance, we are tempted to conclude that the T.E. for the Crank-Nicolson scheme becomes $O(\Delta t) + O[(\Delta x)^2]$, when evaluated by expanding about point (n, j), because of the appearance of the terms $-u_{tt}\,\Delta t/2$ and $\alpha u_{txx}\,\Delta t/2$. However, we can recognize these two terms as $-(\Delta t/2)(\partial/\partial t)(u_t - \alpha u_{xx})$, where the quantity in the second set of parentheses is the left-hand side of Eq. (3.71*b*). Thus we can differentiate Eq. (3.71*b*) with respect to t and multiply both sides by $-\Delta t/2$ to learn that $-(\Delta t/2)(\partial/\partial t)(u_t - \alpha u_{xx}) = O[(\Delta t)^2] + O[(\Delta x)^2]$. From this, we conclude that the T.E. for the Crank-Nicolson scheme is $O[(\Delta t)^2] + O[(\Delta x)^2]$ when evaluated about either point (n, j) or point $(n + \frac{1}{2}, j)$. Use of other points will give the same results for the order of the T.E. This example illustrates that the leading terms in the T.E. should be examined very carefully to see if they can be identified as a multiple of a derivative of the original PDE. If they can, they should be replaced by expressions of higher order.

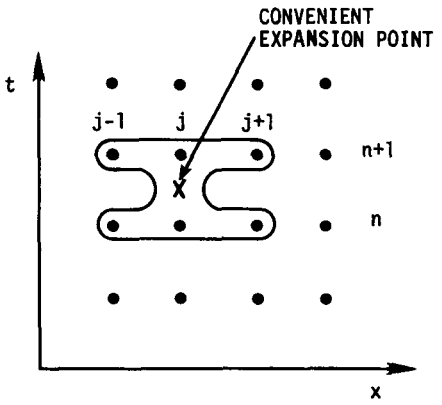

Figure 3.5 Difference molecule, Crank-Nicholson form for heat equation.

3.4.2 Use of Polynomial Fitting

Many applications of polynomial fitting are observed in computational fluid mechanics and heat transfer. The technique can be used to develop the entire finite-difference representation for a PDE. However, the technique is perhaps most commonly employed in the treatment of boundary conditions or in gleaning information from the solution in the neighborhood of the boundary. Consider some specific examples.

Example 3.1 In this example, the derivative approximations needed to represent a PDE will be obtained by assuming that the solution to the PDE can be approximated locally by a polynomial. The polynomial is then "fitted" to the points surrounding the general point (i, j), utilizing values of the function at the grid points. A sufficient number of points can be used to determine the coefficients in the polynomial exactly. The polynomial can then be differentiated to obtain the desired approximation to the derivatives. Consider Laplace's equation, which governs the 2-D temperature distribution in a solid under steady-state conditions:

$$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} = 0 \tag{3.72}$$

Solution We suppose that both the x and y dependency of temperature can be expressed by a second-degree polynomial. For example, holding y fixed, we assume that temperatures at various x locations in the neighborhood of point (i,j) can be determined from

$$T(x, y_0) = a + bx + cx^2$$

For convenience, we let $x = 0$ at point (i, j), and $\Delta x = \text{const}$. Clearly,

$$\left.\frac{\partial T}{\partial x}\right)_{i,j} = b$$

$$\left.\frac{\partial^2 T}{\partial x^2}\right)_{i,j} = 2c$$

The coefficients a, b, and c can be evaluated in terms of temperatures at specific grid points and Δx. To do so, we must make some choices as to which neighboring grid points to use, and this choice determines the geometric arrangement of the difference molecule, that is, whether the resulting derivative approximations are central, forward, or backward differences. Here we will choose points $(i - 1, j)$, (i, j), and $(i + 1, j)$ and obtain

$$T(i, j) = a$$

$$T(i + 1, j) = a + b\,\Delta x + c(\Delta x)^2$$

$$T(i - 1, j) = a - b\,\Delta x + c(\Delta x)^2$$

from which we determine that

$$b = \left.\frac{\partial T}{\partial x}\right)_{i,j} = \frac{T_{i+1,j} - T_{i-1,j}}{2\,\Delta x}$$

$$c = \frac{1}{2}\left.\frac{\partial^2 T}{\partial x^2}\right)_{i,j} = \frac{T_{i+1,j} - 2T_{i,j} + T_{i-1,j}}{2(\Delta x)^2}$$

Thus

$$\left.\frac{\partial^2 T}{\partial x^2}\right)_{i,j} = \frac{T_{i+1,j} - 2T_{i,j} + T_{i-1,j}}{(\Delta x)^2} \tag{3.73}$$

This represents an exact result if indeed a second-degree polynomial expresses the correct variation of temperature with x. In the general case, we only suppose that the second-degree polynomial is a good approximation to the solution. The T.E. of the expression, Eq. (3.73), can be determined by substituting Taylor-series expansions about point (i, j) for $T_{i+1,j}$ and $T_{i-1,j}$ into Eq. (3.73). The T.E. is found to be $O[(\Delta x)^2]$ and will involve only fourth-order and higher derivatives, which are equal to zero when the temperature variation is given by a second-degree polynomial.

A finite-difference approximation for $\partial^2 T/\partial y^2$ can be found in a like manner. We notice that arbitrary decisions need to be made in the process of polynomial fitting, which will influence the form and T.E. of the result: in particular, these decisions influence which of the neighboring points will appear in the difference expression. We also observe that there is nothing unique about the procedure of polynomial fitting that guarantees that the difference approximation for the PDE is the best in any sense or that the numerical scheme is stable (when used for a marching problem).

Example 3.2 Suppose we have solved the finite-difference form of the energy equation for the temperature distribution near a solid boundary and we need to estimate the heat flux at the location. Our finite-difference solution gives us only the temperature at discrete grid points. From Fourier's law, the boundary heat flux is given by $q_w = -k\,\partial T/\partial y)_{y=0}$. Thus, we need to approximate $\partial T/\partial y)_{y=0}$ by a difference representation that uses the temperature obtained from the finite-difference solution to the energy equation.

Solution One way to proceed is to assume that the temperature distribution near the boundary is a polynomial and to "fit" such a polynomial, i.e., straight line, parabola, or third-degree polynomial, to the finite-difference solution that has been determined at discrete points. By requiring that the polynomial match the finite-difference solution for T at certain discrete points, the unknown coefficients in the polynomial can be determined.

For example, if we assume that the temperature distribution near the boundary is again a second-degree polynomial of the form $T = a + by + cy^2$, then referring to Fig. 3.6, we note that $\partial T/\partial y)_{y=0} = b$. Further, for equally spaced mesh points we can write

$$T_1 = a$$

$$T_2 = a + b\,\Delta y + c(\Delta y)^2$$

$$T_3 = a + b(2\,\Delta y) + c(2\,\Delta y)^2$$

from which we can determine that

$$a = T_1$$

$$b = \frac{-3T_1 + 4T_2 - T_3}{2\,\Delta y}$$

$$c = \frac{T_1 - 2T_2 + T_3}{2(\Delta y)^2}$$

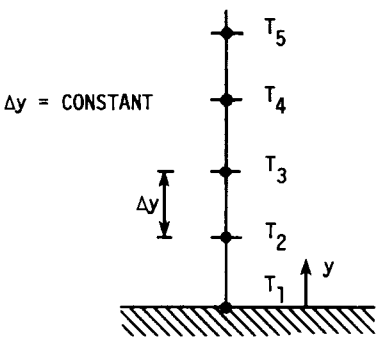

Figure 3.6 Finite-difference grid near wall.

Thus we can evaluate the wall heat flux by the approximation

$$q_w = -k\frac{\partial T}{\partial y}\bigg)_{y=0} \simeq -kb = \frac{k}{2\,\Delta y}(3T_1 - 4T_2 + T_3)$$

It is natural to inquire about the T.E. of this approximation for $\partial T/\partial y)_{y=0}$. This may be established by expressing T_2 and T_3 in terms of Taylor-series expansions about the boundary point and substituting these evaluations into the difference expression for $\partial T/\partial y)_{y=0}$. Alternatively, we can identify the second-degree polynomial as a truncated Taylor-series expansion about $y = 0$.

Second-degree polynomial:

$$T = a + by + cy^2$$

Taylor series:

$$T = T(0) + \frac{\partial T}{\partial y}\bigg)_0 y + \frac{\partial^2 T}{\partial y^2}\bigg)_0 \frac{y^2}{2!} + \underbrace{\frac{\partial^3 T}{\partial y^3}\bigg)_0 \frac{y^3}{3!} + \cdots}_{\text{T.E.}}$$

Thus the approximation $T \simeq a + by + cy^2$ is equivalent to utilizing the first three terms of a Taylor-series expansion with the resulting T.E. in the expression for T being $O[(\Delta y)^3]$. Solving the Taylor series for an expression for $\partial T/\partial y)_{y=0}$ involves division by Δy, which reduces the T.E. in the expression for $\partial T/\partial y)_{y=0}$ to $O[(\Delta y)^2]$.

Example 3.3 Suppose that the energy equation is being solved for the temperature distribution near the wall as in Example 3.2, but now the wall heat flux is specified as a boundary condition. We may then want to use polynomial fitting to obtain an expression for the boundary temperature that is called for in the difference equations for internal points. In other words, if $q_w = -k\,\partial T/\partial y)_{y=0}$ is given, how can we evaluate T at $y = 0$, i.e., (T_1) in terms of q_w/k and T_2, T_3, etc.?

Solution Here we might assume that $T = a + by + cy^2 + dy^3$ near the wall and that $\partial T/\partial y)_{y=0} = b = -q_w/k$ (given). Our objective is to evaluate T_1, which in this case equals a. Referring to Fig. 3.6, we can write

$$T_2 = a - \frac{q_w}{k}\Delta y + c(\Delta y)^2 + d(\Delta y)^3$$

$$T_3 = a - \frac{q_w}{k}(2\,\Delta y) + c(2\,\Delta y)^2 + d(2\,\Delta y)^3$$

$$T_4 = a - \frac{q_w}{k}(3\,\Delta y) + c(3\,\Delta y)^2 + d(3\,\Delta y)^3$$

These three equations can be solved for a, c, and d in terms of T_2, T_3, T_4, q_w/k,

and Δy. The desired result, T_1 as a function of T_2, T_3, q_w/k, and Δy, follows directly from $T_1 = a$ and is given by

$$T_1 = \frac{1}{11}\left(18T_2 - 9T_3 + 2T_4 + \frac{6\,\Delta y\, q_w}{k}\right) + O\left[(\Delta y)^4\right] \tag{3.74}$$

The T.E. in Eq. (3.74) can be established by substituting Taylor-series expansions about (i, j) for the temperatures on the right-hand side or by identifying the polynomial as a truncated series by inspection. We will close this discussion on polynomial fitting by listing some expressions for wall values of a function and its first derivative in terms of values of the function. These expressions are useful, for example, in extracting a value of the function at the wall, if the wall value of the first derivative is specified. The results in Table 3.3 were obtained from polynomial fitting, assuming that $T(y)$ can be expressed as a polynomial of degree up to the fourth, and that $\Delta y = h = \text{const}$.

3.4.3 Integral Method

The integral method provides yet another means for developing difference approximations to PDEs. We consider again the heat equation as the specimen

Table 3.3 Some useful results from polynomial fitting

Polynomial degree	Wall value of function or derivative	Equation
1	$\left(\frac{\partial T}{\partial y}\right)_{i,j} = \frac{T_{i,j+1} - T_{i,j}}{h} + O(h)$	(3.75)
1	$T_{i,j} = T_{i,j+1} - h\left(\frac{\partial T}{\partial y}\right)_{i,j} + O(h^2)$	(3.76)
2	$\left(\frac{\partial T}{\partial y}\right)_{i,j} = \frac{1}{2h}(-3T_{i,j} + 4T_{i,j+1} - T_{i,j+2}) + O(h^2)$	(3.77)
2	$T_{i,j} = \frac{1}{3}\left[4T_{i,j+1} - T_{i,j+2} - 2h\left(\frac{\partial T}{\partial y}\right)_{i,j}\right] + O(h^3)$	(3.78)
3	$\left(\frac{\partial T}{\partial y}\right)_{i,j} = \frac{1}{6h}(-11T_{i,j} + 18T_{i,j+1} - 9T_{i,j+2} + 2T_{i,j+3}) + O(h^3)$	(3.79)
3	$T_{i,j} = \frac{1}{11}\left[18T_{i,j+1} - 9T_{i,j+2} + 2T_{i,j+3} - 6h\left(\frac{\partial T}{\partial y}\right)_{i,j}\right] + O(h^4)$	(3.80)
4	$\left(\frac{\partial T}{\partial y}\right)_{i,j} = \frac{1}{12h}(-25T_{i,j} + 48T_{i,j+1} - 36T_{i,j+2} + 16T_{i,j+3} - 3T_{i,j+4}) + O(h^4)$	(3.81)
4	$T_{i,j} = \frac{1}{25}\left[48T_{i,j+1} - 36T_{i,j+2} + 16T_{i,j+3} - 3T_{i,j+4} - 12h\left(\frac{\partial T}{\partial y}\right)_{i,j}\right] + O(h^5)$	(3.82)

equation:

$$\frac{\partial u}{\partial t} = \alpha \frac{\partial^2 u}{\partial x^2} \tag{3.83}$$

The strategy is to develop an algebraic relationship among the values of u at neighboring grid points by integrating the heat equation with respect to the independent variables t and x over the local neighborhood of point (n, j). The point (n, j) will also be identified as point (t_0, x_0). Grid points are spaced at intervals of Δx and Δt. We arbitrarily decide to integrate both sides of the equation over the interval t_0 to $t_0 + \Delta t$ and $x_0 - \Delta x/2$ to $x_0 + \Delta x/2$. Choosing $t_0 - \Delta t/2$ to $t_0 + \Delta t/2$ would lead to an inherently unstable difference equation. Unfortunately, at this point we have no way of knowing which choice for the integration interval would be the right or wrong one relative to stability of the solution method. This can only be determined by a trial calculation or application of the methods for stability analysis, presented in Section 3.6. The order of integration is chosen for each side in a manner to take advantage of exact differentials:

$$\int_{x_0-\Delta x/2}^{x_0+\Delta x/2} \left(\int_{t_0}^{t_0+\Delta t} \frac{\partial u}{\partial t}\, dt \right) dx = \alpha \int_{t_0}^{t_0+\Delta t} \left(\int_{x_0-\Delta x/2}^{x_0+\Delta x/2} \frac{\partial^2 u}{\partial x^2}\, dx \right) dt \tag{3.84}$$

The inner level of integration can be done exactly, giving

$$\int_{x_0-\Delta x/2}^{x_0+\Delta x/2} [u(t_0 + \Delta t, x) - u(t_0, x)]\, dx = \alpha \int_{t_0}^{t_0+\Delta t} \left[\frac{\partial u}{\partial x}\left(t, x_0 + \frac{\Delta x}{2}\right) - \frac{\partial u}{\partial x}\left(t, x_0 - \frac{\Delta x}{2}\right) \right] dt \tag{3.85}$$

For the next level of integration, we take advantage of the mean-value theorem for integrals, which assures us that for a continuous function $f(y)$,

$$\int_{y_1}^{y_1+\Delta y} f(y)\, dy = f(\bar{y})\, \Delta y \tag{3.86}$$

where $\bar{y}$ is some value of y in the interval $y_1 \leqslant \bar{y} \leqslant y_1 + \Delta y$. Thus, any value of y on the interval will provide an approximation to the integral, and we can write

$$\int_{y_1}^{y_1+\Delta y} f(y)\, dy \simeq f(\tilde{y})\, \Delta y \qquad y_1 \leqslant \tilde{y} \leqslant y_1 + \Delta y$$

As we invoke the mean-value theorem to further simplify Eq. (3.85), we arbitrarily select x_0 on the left-hand side and $t_0 + \Delta t$ on the right-hand side as the locations within the intervals of integration at which to evaluate the integrands:

$$[u(t_0 + \Delta t, x_0) - u(t_0, x_0)]\, \Delta x = \alpha \left[\frac{\partial u}{\partial x}\left(t_0 + \Delta t, x_0 + \frac{\Delta x}{2}\right) - \frac{\partial u}{\partial x}\left(t_0 + \Delta t, x_0 - \frac{\Delta x}{2}\right) \right] \Delta t \tag{3.87}$$

To express the result in purely algebraic terms requires that the first derivatives, $\partial u/\partial x$, on the right-hand side be approximated by finite differences. We could achieve this by falling back on our experience to date and simply utilizing central differences. Alternatively, we can continue to pursue a purely integral approach and invoke the mean-value theorem for integrals, again observing that

$$u(t_0 + \Delta t, x_0 + \Delta x) = u(t_0 + \Delta t, x_0) + \int_{x_0}^{x_0 + \Delta x} \frac{\partial u}{\partial x}(t_0 + \Delta t, x)\, dx$$

$$\simeq u(t_0 + \Delta t, x_0) + \frac{\partial u}{\partial x}\left(t_0 + \Delta t, x_0 + \frac{\Delta x}{2}\right)\Delta x \quad (3.88)$$

from which we can write

$$\frac{\partial u}{\partial x}\left(t_0 + \Delta t, x_0 + \frac{\Delta x}{2}\right) \simeq \frac{u(t_0 + \Delta t, x_0 + \Delta x) - u(t_0 + \Delta t, x_0)}{\Delta x} \quad (3.89)$$

In evaluating the integral in Eq. (3.88) through the mean-value theorem, we have arbitrarily evaluated the integrand at the midpoint of the interval. Hence the final result is only an approximation. Treating the other first derivative in a similar manner permits the approximation to the heat equation to be written as

$$[u(t_0 + \Delta t, x_0) - u(t_0, x_0)]\Delta x = \frac{\alpha}{\Delta x}[u(t_0 + \Delta t, x_0 + \Delta x) - 2u(t_0 + \Delta t, x_0)$$

$$+ u(t_0 + \Delta t, x_0 - \Delta x)]\Delta t \quad (3.90)$$

Reverting back to the n, j notation, whereby n denotes time (t) and j denotes space (x), we can rearrange the above in the form

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \frac{\alpha}{(\Delta x)^2}\left(u_{j+1}^{n+1} - 2u_j^{n+1} + u_{j-1}^{n+1}\right) \quad (3.91)$$

which can be recognized as the fully implicit representation of the heat equation, Eq. (3.70), given in Section 3.4.1. The choice of $t_0 + \Delta t$ as the location to use in utilizing the mean-value theorem for the second integration on the right-hand side is responsible for the implicit form. If t_0 had been chosen instead, an explicit formulation would have resulted. We note that a statement of the T.E. does not evolve naturally as part of this method for developing difference equations but must be determined as a separate step.

3.4.4 Finite-Volume (Control-Volume) Approach

In developing what has become known as the *finite-volume* method, the conservation principles are applied to a fixed region in space known as a *control volume*. Some authorities also refer to such a procedure as a control-volume method, so that the two terms, finite volume and control volume, are used somewhat interchangeably in the literature. In the finite-volume approach a point of view is taken that is distinctly different from that taken with any of the

other methods considered thus far. In the Taylor-series and integral methods, we accepted the PDE as the correct and appropriate form of the conservation principle (physical law) governing our problem and merely turned to mathematical tools to develop algebraic approximations to derivatives. We never again considered the physical law represented by the PDE. The Taylor-series and integral methods then proceed in a rather formal, mechanical way, operating on the PDE, which represents the conservation statement (physical law) at a point.

In the finite-volume method the conservation statement is applied in a form applicable to a region in space (control volume). This integral form of the conservation statement is usually well known from first principles, or it can in most cases, be developed from the PDE form of the conservation law. In this approach, we are recognizing the discrete nature of the computational model at the outset. This feature is shared in common with finite-element methods. The finite-volume procedure can, in fact, be considered as a variant of the finite-element method (Hirsch, 1988), although it is, from another point of view, just a particular type of finite-difference scheme.

As an example, consider unsteady 2-D heat conduction in a rectangular-shaped solid. The problem domain is to be divided up into control volumes with associated grid points. We can establish the control volumes first and place grid points in the centers of the volumes (cell-centered method) or establish the grid first and then fix the boundaries of the control volumes (cell-vertex method) by, for example, placing the boundaries halfway between grid points. When the mesh spacing varies, the points will not be in the geometric center of the control volumes in the cell-vertex method. In the present example, equal spacing will be used, so that the two approaches will result in identical grid and control-volume arrangements.

We first consider the control volume labeled A in Fig. 3.7, which is representative of all internal (nonboundary) points. The appropriate form of the conservation statement for the control volume (namely, that the time rate of increase of energy stored in the volume is equal to the net rate at which energy is conducted into the volume) can be represented mathematically as

$$\iiint_R \rho c \frac{\partial T}{\partial t}\, dR + \oiint_S \mathbf{q} \cdot \mathbf{n}\, dS = 0$$

The first term in this equation, an integral over the control volume, represents the time rate of increase in the energy stored in the volume. The second term, an integral over the surface of the volume, represents the net rate at which energy is conducted out through the surface of the volume. This is the *integral* or control-volume form of the conservation law that we are applying in this case and is the usual starting point for the derivation of the conservation law in partial differential form. On the other hand, if the PDE form of the conservation law is available to us, we can usually work backward with the aid of the

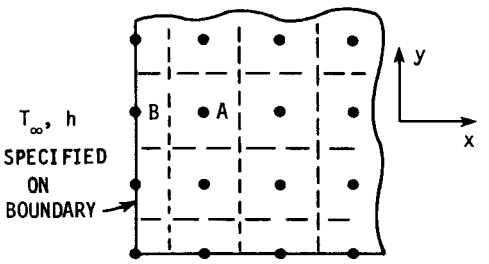

Figure 3.7 Finite-difference grid for control-volume method.

divergence theorem to obtain the appropriate integral form. For example, with constant properties, this problem is governed by the 2-D heat equation, an extension of Eq. (3.62), which can be written in the form

$$\rho c \frac{\partial T}{\partial t} = \frac{\partial}{\partial x}\left(k \frac{\partial T}{\partial x}\right) + \frac{\partial}{\partial y}\left(k \frac{\partial T}{\partial y}\right) = \nabla \cdot (k \nabla T) \qquad (3.92a)$$

where k is the thermal conductivity, ρ is the density, c is the specific heat, and the heat flux vector $\mathbf{q}$ is given by $\mathbf{q} = -k \nabla T$. We can integrate Eq. (3.92*a*) over the control volume to obtain

$$\iiint_R \left(\rho c \frac{\partial T}{\partial t} + \nabla \cdot \mathbf{q}\right) dR = 0 \qquad (3.92b)$$

Applying the divergence theorem gives

$$\iiint_R \rho c \frac{\partial T}{\partial t}\, dR + \oiint_S \mathbf{q} \cdot \mathbf{n}\, dS = 0$$

the integral form of the conservation law. The PDE form of the law is derived from the integral form by observing that Eq. (3.92*a*) must hold for all volumes regardless of size or shape. Therefore the integrand itself must be identically zero at every point. Of course, representing conservation of energy by Eqs. (3.92*a*–3.92*b*) assumes the existence of continuous derivatives that appear in the divergence term.

For a 2-D problem, the "volume" employs a unit depth. In two dimensions, we can represent $\mathbf{n}\, dS$ as $\mathbf{i}\, dy - \mathbf{j}\, dx$ for an integration path around the boundary in a *counterclockwise* direction. Thus the surface integral on the right, representing the net flow of heat out through the surface of the volume, can be evaluated as

$$\oint_s (q_x\, dy - q_y\, dx)$$

where q_x and q_y are components of the heat flux in the x and y directions, respectively. The conservation statement then becomes

$$\iiint_R \rho c \frac{\partial T}{\partial t}\, dR + \oiint_S (q_x\, dy - q_y\, dx) = 0 \qquad (3.93)$$

It should be noted that Eq. (3.93) is valid for volumes of any shape. No assumption was necessary about the shape of the volume in order to obtain Eq. (3.93).

The term on the left containing the time derivative can be evaluated by assuming that the temperature at point (i, j) is the mean value for the volume and then using a forward time difference to obtain

$$\rho c \frac{\left(T_{i,j}^{n+1} - T_{i,j}^{n}\right)}{\Delta t} \Delta x \, \Delta y$$

The time level at which the term on the right, representing the net heat flow out of the volume, is evaluated determines whether the scheme will be explicit or implicit. Reasonable choices include time levels n, $n + 1$, or an average of the two. Fourier's law can be used to represent the heat flux components in terms of the temperature:

$$q_x = -k \frac{\partial T}{\partial x} \qquad q_y = -k \frac{\partial T}{\partial y}$$

The second integral in Eq. (3.93), representing the flow of heat out of the four boundaries of the control volume about point (i, j), can be represented by

$$-k \, \Delta y \frac{\partial T}{\partial x}\Bigg)_{i+\frac{1}{2},j} - k \, \Delta x \frac{\partial T}{\partial y}\Bigg)_{i,j+\frac{1}{2}} + k \, \Delta y \frac{\partial T}{\partial x}\Bigg)_{i-\frac{1}{2},j} + k \, \Delta x \frac{\partial T}{\partial y}\Bigg)_{i,j-\frac{1}{2}}$$

The $\frac{1}{2}$ in the subscripts refers to evaluation at the boundaries of the control volume that are halfway between mesh points. The expression for the net flow of heat out of the volume is exact if the derivatives represent suitable average values for the boundaries concerned. Approximating the spatial derivatives by central differences at time level n and combining with the time-derivative representation yields

$$\rho c \frac{\left(T_{i,j}^{n+1} - T_{i,j}^{n}\right)}{\Delta t} \Delta x \, \Delta y + k \, \Delta y \frac{\left(T_{i,j}^{n} - T_{i+1,j}^{n}\right)}{\Delta x} + k \, \Delta x \frac{\left(T_{i,j}^{n} - T_{i,j+1}^{n}\right)}{\Delta y}$$

$$+ k \, \Delta y \frac{\left(T_{i,j}^{n} - T_{i-1,j}^{n}\right)}{\Delta x} + k \, \Delta x \frac{\left(T_{i,j}^{n} - T_{i,j-1}^{n}\right)}{\Delta y} = 0$$

Dividing by $\rho c \, \Delta x \, \Delta y$ and rearranging gives

$$\frac{T_{i,j}^{n+1} - T_{i,j}^{n}}{\Delta t} = \alpha \left(\frac{T_{i+1,j}^{n} - 2T_{i,j}^{n} + T_{i-1,j}^{n}}{(\Delta x)^2} + \frac{T_{i,j+1}^{n} - 2T_{i,j}^{n} + T_{i,j-1}^{n}}{(\Delta y)^2} \right) \tag{3.94}$$

where $\alpha = k/\rho c$. Equation (3.94) corresponds to the explicit finite-difference representation of the 2-D heat equation.

This equation was derived by approximating spatial derivatives at control volume boundaries by central differences; however, it is possible to develop appropriate representations for such derivatives by integral methods in a manner that is not restricted to Cartesian or even orthogonal grids (see Appendix D).

Now consider the control volume on the boundary, labeled B in Fig. 3.7. In this example we will assume that the boundary conditions are convective. For the continuous (nondiscrete) problem, this is formulated mathematically by $h(T_\infty - T_{i,j}) = -k \, \partial T/\partial x)_{i,j}$, where the point (i, j) is the point on the physical

boundary associated with control volume B. If we were to proceed with the Taylor-series approach to this boundary condition, we would likely next seek a difference representation for $\partial T/\partial x)_{i,j}$. If a simple forward difference is used, the difference equation governing the boundary temperature would be

$$h(T_\infty - T_{i,j}^n) = \frac{k}{\Delta x}(T_{i,j}^n - T_{i+1,j}^n) \tag{3.95}$$

In the control-volume approach, however, we are forced to observe that there is some material associated with the boundary point so that conduction may occur along the boundary, and energy can be stored within the volume. The energy balance on the control volume will account for possible transfer across all four boundaries as well as storage. Applying Eq. (3.93) to volume B gives

$$\rho c \frac{(T_{i,j}^{n+1} - T_{i,j}^n)}{\Delta t}\frac{\Delta x\,\Delta y}{2} - k\,\Delta y \frac{\partial T}{\partial x}\bigg)_{i+\frac{1}{2},j} - k\frac{\Delta x}{2}\frac{\partial T}{\partial y}\bigg)_{i,j+\frac{1}{2}} + k\frac{\Delta x}{2}\frac{\partial T}{\partial y}\bigg)_{i,j-\frac{1}{2}}$$
$$+ h\,\Delta y(T_{i,j}^n - T_\infty) = 0$$

Using the same discretization strategy here as was used for volume A, we can write

$$\rho c \frac{(T_{i,j}^{n+1} - T_{i,j}^n)}{2\,\Delta t}\Delta x\,\Delta y + k\,\Delta y\frac{(T_{i,j}^n - T_{i+1,j}^n)}{\Delta x} + \frac{k\,\Delta x}{2}\frac{(T_{i,j}^n - T_{i,j+1}^n)}{\Delta y}$$
$$+ \frac{k\,\Delta x}{2}\frac{(T_{i,j}^n - T_{i,j-1}^n)}{\Delta y} + h\,\Delta y(T_{i,j}^n - T_\infty) = 0$$

Dividing through by $\rho c\,\Delta x\,\Delta y$, we can write the result as

$$\frac{T_{i,j}^{n+1} - T_{i,j}^n}{2\,\Delta t} = \alpha\left[\frac{T_{i+1,j}^n - T_{i,j}^n}{(\Delta x)^2} + \frac{T_{i,j+1}^n - 2T_{i,j}^n + T_{i,j-1}^n}{2(\Delta y)^2}\right] + \frac{h(T_\infty - T_{i,j}^n)}{\rho c\,\Delta x} \tag{3.96}$$

which is somewhat different from Eq. (3.95), which followed from the most obvious application of the Taylor-series method to approximate the mathematical statement of the boundary condition.

Looking back over the methodology of the finite-volume and Taylor-series methods, we can note that the Taylor-series method readily provided difference approximations to derivatives and the representation for the complete PDE was made up from the addition of several such representations. In contrast, the finite-volume method employs the conservation statement or physical law (usually invoked in integral form) corresponding to the entire PDE. The distinctive characteristic of the finite-volume approach is that a "balance" of some physical quantity is made on the region (control volume) in the neighborhood of a grid point. The discrete nature of the problem domain is always taken into account in the finite-volume approach, which ensures that the

physical law is satisfied over a finite region rather than only at a point as the mesh is shrunk to zero. It would appear that the discretization developed by the finite-volume approach would almost certainly have the conservative property.

It is difficult to appreciate the subtle differences that may occur in the difference representations obtained for the same PDE by using the four different methods discussed in this section without working a large number of examples. In many cases, and especially for simple, linear equations, the resulting difference equations can be identical. That is, four different approaches can give the same result. There is no guarantee that difference equations developed by any of the methods will be numerically stable, so that the same difference scheme developed by all four methods could turn out to be worthless. The differences in the results obtained from using the different methods are more likely to become evident in coordinate systems other than rectangular.

3.5 INTRODUCTION TO THE USE OF IRREGULAR MESHES

Clearly, it is convenient to let the mesh increments such as Δx and Δy be constant throughout the computational domain. However, in many instances this is not possible because of domain boundaries that do not coincide with the regular mesh lines or because of the need to reduce the mesh spacing in certain regions in order to maintain the desired level of accuracy. These irregularities occur frequently enough in physical problems to command a significant amount of attention from workers in computational fluid mechanics and heat transfer. In fact, efficiently dealing with irregular geometries that cannot be defined in terms of coordinate lines from a known orthogonal coordinate system is one of the important practical problems challenging computational fluid dynamics at the present time. This problem is complex and has no optimum solution for all cases. Some of the ideas are introduced in this chapter, but the general issue of irregular meshes is addressed at various points throughout the remainder of the book, particularly in Chapters 5 and 10.

3.5.1 Irregular Mesh Due to Shape of a Boundary

Here we address those cases in which some portion of the boundary consists of a curve (in two dimensions) that does not coincide with a coordinate line (for an orthogonal coordinate system) that is satisfactory for the remainder of the boundaries. An example of this would arise in solving Laplace's equations in a rectangular region containing a circular interior "hole." This could also occur in solving for the inviscid flow in a channel containing a circular cylinder, or in a rectangular conduction medium containing a circular pipe. A square mesh, $\Delta x = \Delta y = \text{const}$, would be adequate except near the cylinder, where the spacing between some of the boundary points and the internal points is unequal, as illustrated in Fig. 3.8. If the boundary conditions are Dirichlet (u specified), the following three simple procedures may provide an adequate approximation.

1. Use an especially fine but regular mesh near the boundary and define the

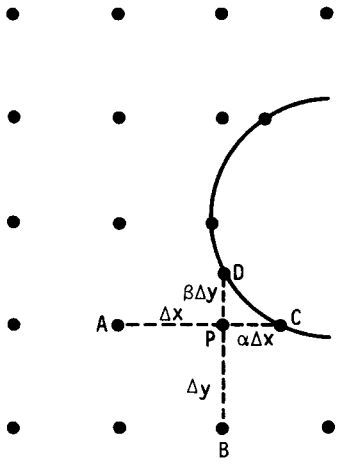

Figure 3.8 Irregular mesh caused by the shape of a boundary.

point closest to the actual boundary as the boundary point for computational purposes. This results in the boundary taking on a "zig-zag" appearance. Unless a coarser mesh is used away from the boundary, resulting in irregular mesh problems where the transition in spacing is made, this method could require a very large number of grid points to achieve reasonable accuracy.

2. Use linear (or bilinear) interpolation to assign values of u to any internal point that is less than a regular mesh increment from the boundary. The interpolation is between the specified boundary values of u and values of u determined at neighboring points by the finite-difference equations applicable to internal points in the regular mesh. This procedure may work but is not strongly recommended. Usually, we can do much better than this with very little additional effort, as indicated below.
3. Develop a finite-difference approximation to the governing PDE that is valid at internal points even when the mesh is irregular. Such a difference representation for Laplace's equation valid on a Cartesian grid with irregular spacing (Δx and Δy not constant) can be developed quite readily through the integral method by integrating about point x_0, y_0 and letting each integration interval extend halfway to a neighboring point. The mesh notation used is defined in Fig. 3.9. The starting point for the integral development of the difference expression is

$$\int_{y_0-\Delta y_-/2}^{y_0+\Delta y_+/2}\left(\int_{x_0-\Delta x_-/2}^{x_0+\Delta x_+/2}\frac{\partial^2 u}{\partial x^2}\,dx\right)dy+\int_{x_0-\Delta x_-/2}^{x_0+\Delta x_+/2}\left(\int_{y_0-\Delta y_-/2}^{y_0+\Delta y_+/2}\frac{\partial^2 u}{\partial y^2}\,dy\right)dx=0$$

Using the definition of an exact differential, this can be written as

$$\int_{y_0-\Delta y_-/2}^{y_0+\Delta y_+/2}\left[\frac{\partial u}{\partial x}\left(x_0+\frac{\Delta x_+}{2},y\right)-\frac{\partial u}{\partial x}\left(x_0-\frac{\Delta x_-}{2},y\right)\right]dy$$

$$+\int_{x_0-\Delta x_-/2}^{x_0+\Delta x_+/2}\left[\frac{\partial u}{\partial y}\left(x,y_0+\frac{\Delta y_+}{2}\right)-\frac{\partial u}{\partial y}\left(x,y_0-\frac{\Delta y_-}{2}\right)\right]dx=0$$

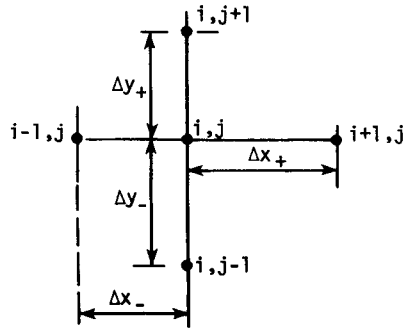

Figure 3.9 Notation for arbitrary irregular mesh.

Employing the mean-value theorem for integrals and using the central point of the interval to evaluate the integrands gives

$$\left[\frac{\partial u}{\partial x}\left(x_0 + \frac{\Delta x_+}{2}, y_0\right) - \frac{\partial u}{\partial x}\left(x_0 - \frac{\Delta x_-}{2}, y_0\right)\right]\frac{\Delta y_+ + \Delta y_-}{2}$$
$$+\left[\frac{\partial u}{\partial y}\left(x_0, y_0 + \frac{\Delta y_+}{2}\right) - \frac{\partial u}{\partial y}\left(x_0, y_0 - \frac{\Delta y_-}{2}\right)\right]\frac{\Delta x_+ + \Delta x_-}{2} = 0$$

Approximating these derivatives centrally, as was done in Section 3.4.3 gives, after rearrangement, the following approximation for Laplace's equation in subscript notation:

$$\frac{2}{\Delta x_+ + \Delta x_-}\left(\frac{u_{i+1,j} - u_{i,j}}{\Delta x_+} - \frac{u_{i,j} - u_{i-1,j}}{\Delta x_-}\right)$$
$$+\frac{2}{\Delta y_+ + \Delta y_-}\left(\frac{u_{i,j+1} - u_{i,j}}{\Delta y_+} - \frac{u_{i,j} - u_{i,j-1}}{\Delta y_-}\right) = 0 \qquad (3.97)$$

When the above is specialized to the points near the irregular boundary depicted in Fig. 3.8, the derivative approximations appear as

$$\left(\frac{\partial^2 u}{\partial x^2}\right)_p \cong \frac{2}{\Delta x(1+\alpha)}\left(\frac{u_C - u_P}{\alpha \Delta x} - \frac{u_P - u_A}{\Delta x}\right)$$
$$\left(\frac{\partial^2 u}{\partial y^2}\right)_p \cong \frac{2}{\Delta y(1+\beta)}\left(\frac{u_D - u_P}{\beta \Delta y} - \frac{u_P - u_B}{\Delta y}\right)$$

Equation (3.97) can also be developed by the control-volume method or by utilizing Taylor-series expansions. However, the unequal spacing makes the Taylor-series method noticeably more laborious, whereas the integral approach proceeds for unequal spacing with no increase in effort. Likewise, using the control-volume method would require little additional effort. However, Taylor-series expansions about (i, j) should be substituted into Eq. (3.97) to establish the consistency and T.E. of these approximations. This will be left as an exercise

for the reader. As a note of warning, we recall that our second-derivative approximations on a regular mesh acquired second-order accuracy only through fortuitous cancellation of terms from the forward and backward Taylor-series expansions. This cancellation will not occur if the mesh increments are unequal.

When approximately the same number of grid points are being used, we might expect this third method of treating irregular points near boundaries to be the most accurate because the governing PDE is being approximated at each internal point (not the case for procedure 2), and the location of the boundary is not being altered as was done in procedure 1.

The above approximate procedures can be useful when solving a single equation for a problem in which Dirichlet boundary conditions are specified on an irregular boundary. However, when a system of equations is being solved or the boundary conditions involve derivatives (Neumann), the simple procedures given above are usually not adequate. Better ways of dealing with this problem usually add significantly to the complexity of the problem formulation. One common way of handling this type of problem is through the use of generalized body-fitted coordinates. This procedure is discussed in Chapters 5 and 10. The finite-volume method can also be extended to provide a satisfactory representation. An example of how the finite-volume approach can be applied to obtain satisfactory difference representations for control volumes associated with irregular boundaries is given below.

Finite-volume treatment of irregular boundary. As before, the governing PDE is Laplace's equation, and we are considering the effect of an irregular boundary on a computational domain that is otherwise discretized with an orthogonal coordinate system. In particular, we will consider the configuration depicted in Fig. 3.10, where control volumes will be rectangles except near the irregular boundary. In this case, application of the finite-volume methodology will result

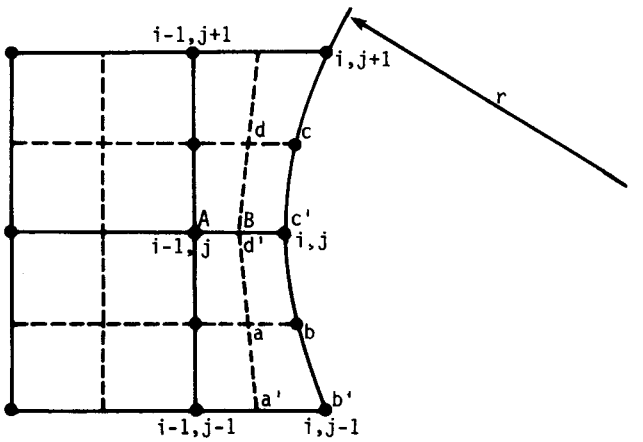

Figure 3.10 Finite-volume treatment of irregular boundary.

in Eq. (3.97) if the volume faces form a rectangle (i.e., if adjacent faces are orthogonal). The only exceptions to this will be for those volumes on the irregular boundary and their immediate internal neighbors. Some immediate internal neighbors to boundary cells, like volume A in Fig. 3.10, will have sufficient geometric symmetry so that Eq. (3.97) will be obtained from the finite-volume analysis.

If the boundary conditions at the irregular boundary are Dirichlet, no unknowns exist in the volumes that reside on the boundary, so a heat balance on those cells is not necessary. However, if the boundary condition is Neumann, corresponding to a specified value of boundary heat flux (q_w), the boundary temperature is unknown, and it is appropriate to apply the integral form of the conservation statement developed in Section 3.4.4:

$$\iiint_R \rho c \frac{\partial T}{\partial t}\, dR + \oiint_S (q_x\, dy - q_y\, dx) = 0$$

to the volume (labeled B in Fig. 3.10) on the boundary. The dashed lines in Fig. 3.10 denote the boundaries of the control volume. The corner points, located halfway between the specified nodal points, are labeled a, b, c, d. It is assumed that the coordinates of the nodal points are known. The coordinates of points a and d are given by

$$x_a = (x_{i,j-1} + x_{i,j} + x_{i-1,j} + x_{i-1,j-1})/4$$
$$y_a = (y_{i,j-1} + y_{i,j} + y_{i-1,j} + y_{i-1,j-1})/4$$
$$x_d = (x_{i,j+1} + x_{i,j} + x_{i-1,j} + x_{i-1,j+1})/4$$
$$y_d = (y_{i,j+1} + y_{i,j} + y_{i-1,j} + y_{i-1,j+1})/4$$

Since we want points b and c to lie exactly on the boundary, we will fix $y_b = (y_{i,j} + y_{i,j-1})/2$ and $y_c = (y_{i,j+1} + y_{i,j})/2$ and establish the x coordinates so that the points are on the boundary. This is easily done, since the equation for the boundary curve is known in the form $(x - x_0)^2 + (y - y_0)^2 = r^2$, where x_0, y_0 are the coordinates of the center of the circle and r is the radius. We can separate the integral around the boundaries of B into line integrals over the four component line segments, a-b, b-c, c-d, d-a:

$$\oiint_S \mathbf{q} \cdot \mathbf{n}\, ds = \int_a^b (q_x\, dy - q_y\, dx) + q_w\, \Delta s_{bc} + \int_c^d (q_x\, dy - q_y\, dx) + \int_d^a (q_x\, dy - q_y\, dx) \qquad (3.98)$$

where Δs_{bc} can be computed exactly in this case, making use of the known equation for a circle or approximated by a straight-line segment between the fixed points as is done for the other boundaries of the control volume. We now start at point a to evaluate the line integral [Eq. (3.98)] around the boundaries of control volume B in Fig. 3.10. Along the boundary from a to b the heat flux

components can be evaluated by Fourier's law as

$$k\int_a^b\left(-\frac{\partial T}{\partial x}\,dy+\frac{\partial T}{\partial y}\,dx\right)\cong k\left(\frac{\partial T}{\partial y}\right)_{i-\frac{1}{4},j-\frac{1}{2}}\Delta x_{ab}-k\left(\frac{\partial T}{\partial x}\right)_{i-\frac{1}{4},j-\frac{1}{2}}\Delta y_{ab}$$

where $\Delta x_{ab}=x_b-x_a$ and $\Delta y_{ab}=y_b-y_a$. The values of $\partial T/\partial y)_{i-\frac{1}{4},j-\frac{1}{2}}$ and $\partial T/\partial x)_{i-\frac{1}{4},j-\frac{1}{2}}$ are approximately in the center of the region $a'b'c'd'$ denoted in Fig. 3.10. It is assumed that these derivatives can be approximated by averages over $a'b'c'd'$.

$$\left.\frac{\partial T}{\partial y}\right)_{i-\frac{1}{4},j-\frac{1}{2}}\cong\frac{1}{A'}\left(\iint_{A'}\frac{\partial T}{\partial y}\,dy\,dx\right)$$

$$\left.\frac{\partial T}{\partial x}\right)_{i-\frac{1}{4},j-\frac{1}{2}}\cong\frac{1}{A'}\left(\iint_{A'}\frac{\partial T}{\partial x}\,dy\,dx\right)$$

where A' denotes the area of the region $a'b'c'd'$. Using the Gauss divergence theorem again, the integrals over the area $a'b'c'd'$ can be evaluated by line integrals around the boundary of $a'b'c'd'$. This allows the heat flux across the a-b portion of the boundary of control volume B to be represented as

$$\begin{aligned}&\int_a^b(q_x\,dy-q_y\,dx)\\&\quad\cong-\frac{k}{A'}\left(\oint_{A'}T\,dx\,\Delta x_{ab}+\oint_{A'}T\,dy\,\Delta y_{ab}\right)\\&\quad\cong\frac{-k}{A'}\Big[(T_{i-\frac{1}{4},j-1}\,\Delta x_{a'b'}+T_b\,\Delta x_{b'c'}+T_{i-\frac{1}{4},j}\,\Delta x_{c'd'}+T_a\,\Delta x_{d'a'})\,\Delta x_{ab}\\&\qquad+(T_{i-\frac{1}{4},j-1}\Delta y_{a'b'}+T_b\,\Delta y_{b'c'}+T_{i-\frac{1}{4},j}\,\Delta y_{c'd'}+T_a\,\Delta y_{d'a'})\,\Delta y_{ab}\Big]\end{aligned}\tag{3.99}$$

Because $\Delta y=0$ along path a-b, half of the terms on the right-hand side of Eq. (3.99) vanish, so that the expression simplifies to

$$\begin{aligned}&\frac{-k}{A'}\Big[T_{i-\frac{1}{4},j-1}(\Delta x_{a'b'}\,\Delta x_{ab})+T_b(\Delta x_{b'c'}\,\Delta x_{ab})\\&\qquad+T_{i-\frac{1}{4},j}(\Delta x_{c'd'}\,\Delta x_{ab})+T_a(\Delta x_{d'a'}\,\Delta x_{ab})\Big]\end{aligned}\tag{3.100}$$

We note that further simplifications would occur if Δx were zero along paths b'-c' and d'-a' (i.e., the paths were parallel to the y axis). The temperatures required in Eq. (3.100) must be obtained by interpolation from values at nodal points. For this configuration, bilinear interpolation yields

$$T_a=0.25(T_{i,j}+T_{i,j-1}+T_{i-1,j-1}+T_{i-1,j})$$

$$T_b=0.5(T_{i,j}+T_{i,j-1})$$

$$T_{i-\frac{1}{4},j}=0.25T_{i-1,j}+0.75T_{i,j}$$

$$T_{i-\frac{1}{4},j-1}=0.25T_{i-1,j-1}+0.75T_{i,j-1}$$

The area A' can be approximated in several ways, one of which is by assuming that $a'b'c'd'$ forms a quadrilateral and computing its area as one-half the cross product of the diagonals of the quadrilateral region:

$$A' = 0.5(\Delta x_{d'b'}\,\Delta y_{a'c'} - \Delta y_{d'b'}\,\Delta x_{a'c'})$$

where $x_{a'} = 0.5(x_{i-1,j-1} + x_{i,j-1})$
$x_{b'} = x_{i,j-1}$
$x_{c'} = x_{i,j}$
$x_{d'} = 0.5(x_{i-1,j} + x_{i,j})$

In this formulation, care must be exercised in order to obtain a positive value for the area. This can be assured by employing the right-hand rule or by taking the absolute value of the cross product. The y coordinates of points a', b', c', d' are found by replacing x with y in the expressions above. The fluxes across control-volume boundaries c-d and d-a can be evaluated by extending the methodology illustrated above for boundary a-b appropriately.

Although the irregular shape of the boundary volumes clearly adds significant complexity to the solution procedure, the techniques needed to deal with this can be generalized and implemented reasonably systematically and efficiently. On the other hand, it is correct to conclude that when the boundaries of the domain of interest do not coincide with grid lines of an orthogonal coordinate system and the boundary conditions are not Dirichlet, a major escalation in the effort required to formulate the solution procedure seems to follow.

3.5.2 Irregular Mesh Not Caused by Shape of a Boundary

Here we assume that the boundaries of the problem domain conform to grid lines in an orthogonal coordinate system. The use of variable grid spacing may still be desirable in this situation because it is often necessary to employ very small grid spacings in regions where gradients of the dependent variables are especially large in order to obtain the desired accuracy or "resolution." However, in the interest of computational economy, we strive to use a coarser grid away from these critical regions. This requires that the mesh spacings vary. We can cite at least two ways to proceed:

1. We can employ a coordinate transformation so that unequal spacing in the original coordinate system becomes equal spacing in the new system but the PDE becomes altered somewhat in form. This procedure is described in detail in Chapter 5.
2. The difference equation can be formulated in such a way that it remains valid when the spacing is irregular (grid lines remain orthogonal, but the increments in each coordinate direction vary instead of remaining constant). Actually, this is the same as procedure 3 used above in connection with the irregular mesh caused by curved boundaries. Such a formulation for Laplace's equation is given as Eq. (3.97).

3.5.3 Concluding Remarks

The purpose of this section has been to introduce some of the problems and applicable solution procedures associated with irregular boundaries and unequal mesh spacing in general. Coverage of the topic has been by no means complete. More advanced considerations on this topic tend to quickly become quite specialized and detailed. Good pedagogy suggests that we move on and see more of the forest before we spend any more time studying this tree. Some ideas on this topic will be developed further in Chapters 5 and 10 and in connection with specific example problems in fluid mechanics and heat transfer.

3.6 STABILITY CONSIDERATIONS

A finite-difference approximation to a PDE may be consistent, but the solution will not necessarily converge to the solution of the PDE. The Lax Equivalence theorem (see Section 3.3.5) states that a stable numerical method must also be used. We will address the question of stability in this section.

The problem of stability in numerical analysis is similar to the problem of stability encountered in a modern control system. The transfer function in a control system plays the role of the difference operator. Consider a marching problem in which initial values at time level n are known and values of the unknown at time level $n + 1$ are required. The difference operator may be viewed as a "black box" that has a certain transfer function. A schematic representation would appear as shown in Fig. 3.11. The stability of such a system depends upon the operations performed by the black box on the input data. A control systems engineer would require that the transfer function have no poles in the right-half plane. Without this requirement, input signals would be falsely amplified, and the output would be useless; in fact, it would grow without bound. Similarly, the way in which the difference operator alters the input information to produce the solution at the next time level is the central concern of stability analysis.

As a starting point for stability analysis, consider the simple explicit approximation to the heat equation:

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \frac{\alpha}{(\Delta x)^2}(u_{j+1}^n - 2u_j^n + u_{j-1}^n)$$

This may be solved for u_j^{n+1} to yield

$$u_j^{n+1} = u_j^n + \alpha\frac{\Delta t}{(\Delta x)^2}(u_{j+1}^n - 2u_j^n + u_{j-1}^n) \tag{3.101}$$

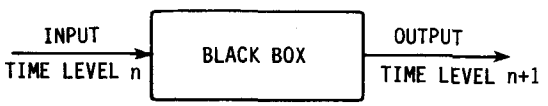

Figure 3.11 Schematic diagram of stability.

Let the exact solution of this equation be denoted by D. This is the solution that would be obtained using a computer with infinite accuracy. Similarly, denote by N the numerical solution of Eq. (3.101) computed using a real machine with finite accuracy. If the analytical solution of the PDE is A, then we may write

$$\text{Discretization error} = A - D$$
$$\text{Round-off error} = N - D$$

The question of stability of a numerical method examines the error growth while computations are being performed. O'Brien et al. (1950) pose the question of stability in the following manner:

1. Does the overall error due to round-off

$$\begin{bmatrix}\text{Grow}\\ \text{Not grow}\end{bmatrix} \Rightarrow \text{strong} \begin{bmatrix}\text{instability}\\ \text{stability}\end{bmatrix}$$

2. Does a single general round-off error

$$\begin{bmatrix}\text{Grow}\\ \text{Not grow}\end{bmatrix} \Rightarrow \text{weak} \begin{bmatrix}\text{instability}\\ \text{stability}\end{bmatrix}$$

The second question is the one most frequently answered because it can be treated much more easily from a practical point of view. The question of weak stability is usually answered by using a Fourier analysis. This method is also referred to as a von Neumann analysis. It is assumed that proof of weak stability using this method implies strong stability.

3.6.1 Fourier or von Neumann Analysis

Consider the finite-difference equation, Eq. (3.101). Let ϵ represent the error in the numerical solution due to round-off errors. The numerical solution actually computed may be written

$$N = D + \epsilon \tag{3.102}$$

This computed numerical solution must satisfy the difference equation. Substituting Eq. (3.102) into the difference equation, Eq. (3.101), yields

$$\frac{D_j^{n+1} + \epsilon_j^{n+1} - D_j^n - \epsilon_j^n}{\Delta t} = \alpha\left(\frac{D_{j+1}^n + \epsilon_{j+1}^n - 2D_j^n - 2\epsilon_j^n + D_{j-1}^n + \epsilon_{j-1}^n}{\Delta x^2}\right)$$

Since the exact solution D must satisfy the difference equation, the same is true of the error, i.e.,

$$\frac{\epsilon_j^{n+1} - \epsilon_j^n}{\Delta t} = \alpha\left(\frac{\epsilon_{j+1}^n - 2\epsilon_j^n + \epsilon_{j-1}^n}{\Delta x^2}\right) \tag{3.103}$$

In this case, the exact solution D and the error ϵ must both satisfy the same difference equation. This means that the numerical error and the exact numerical solution both possess the same growth property in time and either could be used

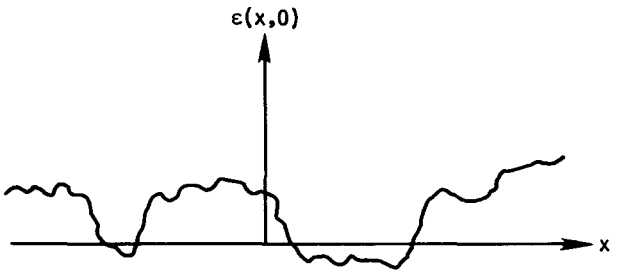

Figure 3.12 Initial error distribution.

to examine stability. Any perturbation of the input values at the nth time level will either be prevented from growing without bound for a stable system or will grow larger for an unstable system.

Consider a distribution of errors at any time in a mesh. We choose to view this distribution at time $t = 0$ for convenience. This error distribution is shown schematically in Fig. 3.12. We assume the error $\epsilon(x, t)$ can be written as a series of the form

$$\epsilon(x,t) = \sum_m b_m(t) e^{ik_m x} \tag{3.104}$$

where the period of the fundamental frequency ($m = 1$) is assumed to be $2L$. For the interval $2L$ units in length, the wave number may be written

$$k_m = \frac{2\pi m}{2L} \qquad m = 0, 1, 2, \ldots, M$$

where M is the number of increments Δx units long contained in length L. For instance, if an interval of length $2L$ is subdivided using five points, the value of M is 2, and the corresponding frequencies are

$$f_m = \frac{k_m}{2\pi} = \frac{m}{2L}$$

$$f_0 = 0 \qquad m = 0$$

$$f_1 = \frac{1}{2L} \qquad m = 1$$

$$f_2 = \frac{1}{L} \qquad m = 2$$

The frequency measures the number of wavelengths in each $2L$ units of length. The lowest frequency ($m = 0$, $f_0 = 0$) corresponds to a steady term in the assumed expansion. The highest frequency ($m = M$) has a wave number of $\pi/\Delta x$ and corresponds to the minimum number of points (3) required to approximately represent a sine or cosine wave between 0 and 2π.

Since the difference equation is linear, superposition may be used, and we may examine the behavior of a single term of the series given in Eq. (3.104). Consider the term

$$\epsilon_m(x,t) = b_m(t)e^{ik_m x}$$

We seek solutions of the form

$$z^n e^{ik_m x}$$

which reduces to $e^{ik_m x}$ when $t = 0$ ($n = 0$). Toward this end, let

$$z = e^{a\,\Delta t}$$

so that

$$\begin{aligned} z^n &= e^{an\,\Delta t} = e^{at} \\ \epsilon_m(x,t) &= e^{at}e^{ik_m x} \end{aligned} \tag{3.105}$$

where k_m is real but a may be complex.

If Eq. (3.105) is substituted into Eq. (3.103), we obtain

$$e^{a(t+\Delta t)}e^{ik_m x} - e^{at}e^{ik_m x} = r(e^{at}e^{ik_m(x+\Delta x)} - 2e^{at}e^{ik_m x} + e^{at}e^{ik_m(x-\Delta x)})$$

where $r = \alpha\,\Delta t/(\Delta x)^2$. If we divide by $e^{at}e^{ik_m x}$ and utilize the relation

$$\cos\beta = \frac{e^{i\beta} + e^{-i\beta}}{2}$$

the above expression becomes

$$e^{a\,\Delta t} = 1 + 2r(\cos\beta - 1)$$

where $\beta = k_m\,\Delta x$. Employing the trigonometric identity

$$\sin^2\frac{\beta}{2} = \frac{1 - \cos\beta}{2}$$

the final expression is

$$e^{a\,\Delta t} = 1 - 4r\sin^2\frac{\beta}{2} \tag{3.106}$$

Furthermore, since $\epsilon_j^{n+1} = e^{a\,\Delta t}\epsilon_j^n$ for each frequency present in the solution for the error, it is clear that if $|e^{a\,\Delta t}|$ is less than or equal to 1, a general component of the error will not grow from one time step to the next. This requires that

$$\left|1 - 4r\sin^2\frac{\beta}{2}\right| \leqslant 1 \tag{3.107}$$

The factor $1 - 4r\sin^2\beta/2$ (representing $\epsilon_j^{n+1}/\epsilon_j^n$) is called the *amplification factor* and will be denoted by G. Clearly, the influence of boundary conditions is not included in this analysis. In general, the Fourier stability analysis assumes that we have imposed periodic boundary conditions.

In evaluating the inequality Eq. (3.107), two possible cases must be considered:

1. Suppose $(1 - 4r \sin^2 \beta/2) \geqslant 0$; then $4r \sin^2 \beta/2 \geqslant 0$.
2. Suppose $(1 - 4r \sin^2 \beta/2) < 0$; then $4r \sin^2 \beta/2 - 1 \leqslant 1$.

The first condition is always satisfied if $r \geqslant 0$. The second inequality is satisfied only if $r \leqslant \frac{1}{2}$, which is the stability requirement for this method. This numerically places a constraint on the size of the time step relative to the size of the mesh spacing. The reason for the physically implausible temperatures calculated in the example at the end of Section 3.3.4 is now very clear. The step size Δt selected was too large by a factor of 2, and the solution began to diverge immediately. The stability of the calculation with $\alpha(\Delta t/\Delta x^2) = \frac{1}{2}$ can easily be verified. It should be noted that the amplification factor given by Eq. (3.106) could have been deduced by substituting a general form given by Eq. (3.104) into the difference equation. The proof is left as an exercise for the reader.

Example 3.4 The simple implicit scheme applied to the heat equation is given by

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \frac{\alpha}{(\Delta x)^2}\left(u_{j+1}^{n+1} - 2u_j^{n+1} + u_{j-1}^{n+1}\right)$$

Determine the stability restrictions (if any) for this algorithm.

Solution After substituting Eq. (3.105) into this algorithm, we obtain

$$e^{a\,\Delta t}(1 + 2r - 2r \cos \beta) = 1$$

Using the trigonometric identity,

$$\sin^2 \frac{\beta}{2} = \frac{1 - \cos \beta}{2}$$

the amplification factor becomes

$$G = \frac{1}{1 + 4r \sin^2 \beta/2}$$

The condition for stability $|G| \leqslant 1$ is satisfied for all $r \geqslant 0$. Hence there is no upper limit on step size because of stability. However, there is a practical limit on step size because of T.E.

The application of the von Neumann or Fourier stability method is equally straightforward for hyperbolic equations. As an example, the first-order wave equation in one dimension is

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = 0 \tag{3.108}$$

where c is the wave speed. This equation has one characteristic given by a solution of $x_t = c$. The solution of Eq. (3.108) is given by

$$u(x - ct) = \text{const}$$

This solution requires the initial data prescribed at $t = 0$ to be propagated along the characteristics.

Lax (1954) proposed the following first-order method for solving equations of this form:

$$u_j^{n+1} = \frac{u_{j+1}^n + u_{j-1}^n}{2} - c\frac{\Delta t}{\Delta x}\left(\frac{u_{j+1}^n - u_{j-1}^n}{2}\right) \tag{3.109}$$

The first term on the right-hand side represents an average value of the unknown at the previous time level, while the second term is the difference form of the spatial derivative. If a term of the form

$$u_j^n = e^{at}e^{ik_m x}$$

is substituted into the difference equation, the amplification factor becomes

$$e^{a\Delta t} = \cos\beta - i\nu\sin\beta$$

The stability requirement is $|G| \leqslant 1$ or

$$|\cos\beta - i\nu\sin\beta| \leqslant 1$$

where $\nu = c\,\Delta t/\Delta x$ is called the Courant number. Since the square of the absolute value of a complex number is the sum of the squares of the real and imaginary parts, the method is stable if

$$|\nu| \leqslant 1 \tag{3.110}$$

Again, a conditional stability requirement must be placed on the time step and the spatial mesh spacing. This is called the Courant-Friedrichs-Lewy (CFL) condition and was discussed at length relative to the concepts of convergence and stability in an historically important paper by Courant et al. (1928). Some authorities consider this paper to be the starting point for the development of modern numerical methods for PDEs.

The amplification factor or growth factor for a particular numerical method depends upon mesh size and wave number or frequency. The amplification factor for the Lax finite-difference method may be written

$$G = \cos\beta - i\nu\sin\beta = |G|e^{i\phi} = \sqrt{\cos^2\beta + \nu^2\sin^2\beta}\; e^{i\tan^{-1}(-\nu\tan\beta)} \tag{3.111}$$

where ϕ is the phase angle. Clearly, the magnitude of G changes with Courant number ν and frequency parameter β, which varies between 0 and π. A good understanding of the amplification factor can be obtained from a polar plot. Figure 3.13 is a plot of Eq. (3.111) for several different Courant numbers. Several interesting results can be deduced by a careful examination of this plot. The phase angle for the Lax method varies from 0 for the low frequencies to $-\pi$ for the high frequencies. This may be seen by computing the phase for both cases. For a Courant number of 1, all frequency components are propagated

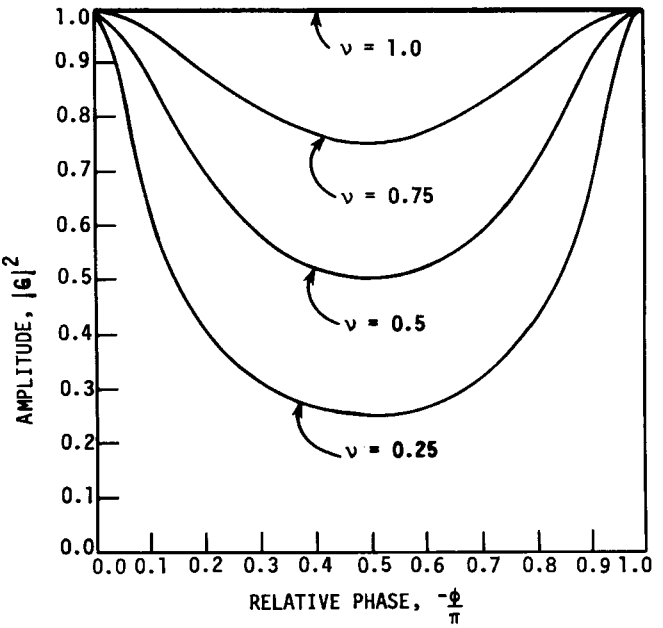

Figure 3.13 Amplitude-phase plot for the amplification factor of the Lax scheme.

without attenuation in the mesh. For Courant numbers less than 1, the low- and high-frequency components are only mildly altered, while the midrange frequency signal content is severely attenuated. The phase is also shown, and we can determine the phase error for any frequency from these curves.

A physical interpretation of the results provided by Eq. (3.110) for hyperbolic equations is important. Consider the second-order wave equation:

$$u_{tt} - c^2 u_{xx} = 0 \tag{3.112}$$

This equation has characteristics

$$x + ct = \text{const} = c_1$$
$$x - ct = \text{const} = c_2$$

A solution at a point (x, t) depends upon data contained between the characteristics that intersect that point, as sketched in Fig. 3.14. The analytic solution at (x, t) is influenced only by information contained between c_1 and c_2.

The numerical stability requirement for many explicit numerical methods for solving hyperbolic PDEs is the CFL condition, which, for the wave equation, is

$$\left| c \frac{\Delta t}{\Delta x} \right| \leqslant 1$$

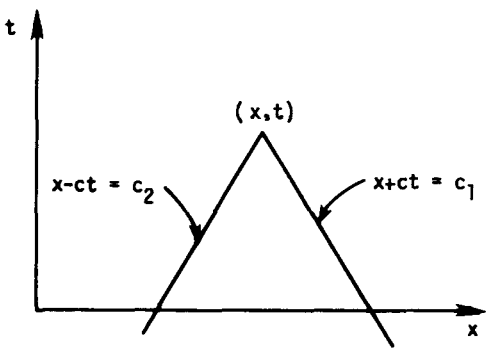

Figure 3.14 Characteristics of the second-order wave equation.

This is the same as given in Eq. (3.110) and may be written as

$$\left(\frac{\Delta t}{\Delta x}\right)^2 \leqslant \frac{1}{c^2}$$

The characteristic slopes are given by $dt/dx = \pm 1/c$. The CFL condition requires that the analytic domain of influence lie within the numerical domain of influence. The numerical domain may include more than, but not less than, the analytical zone. Another interpretation is that the slope of the lines connecting $(j \pm 1, n)$ and $(j, n+1)$ must be smaller in absolute value (flatter) than the characteristics. The CFL requirement makes sense from a physical point of view. One would also expect the numerical solution to be degraded if too much unnecessary information is included by allowing $c(\Delta t/\Delta x)$ to become greatly different from unity. This is, in fact, what occurs numerically. The best results for hyperbolic systems using the most common explicit methods are obtained with Courant numbers near unity. This is consistent with our observations about attenuation associated with the Lax method, as shown in Fig. 3.13.

Before we begin our study of stability for systems of equations, an example demonstrating the application of the von Neumann method to higher dimensional problems is in order.

Example 3.5 A solution of the 2-D heat equation

$$\frac{\partial u}{\partial t} = \alpha\frac{\partial^2 u}{\partial x^2} + \alpha\frac{\partial^2 u}{\partial y^2}$$

is desired using the simple explicit scheme. What is the stability requirement for the method?

Solution The finite-difference equation for this problem is

$$u_{j,k}^{n+1} = u_{j,k}^n + r_x(u_{j+1,k}^n - 2u_{j,k}^n + u_{j-1,k}^n) + r_y(u_{j,k+1}^n - 2u_{j,k}^n + u_{j,k-1}^n)$$

where $r_x = \alpha[\Delta t/(\Delta x)^2]$ and $r_y = \alpha[\Delta t/(\Delta y)^2]$. In this case, a Fourier component of the form

$$u^n_{j,k} = e^{at}e^{ik_x x}e^{ik_y y}$$

is assumed. If $\beta_1 = k_x \Delta x$ and $\beta_2 = k_y \Delta y$, we obtain

$$e^{a\Delta t} = 1 + 2r_x(\cos \beta_1 - 1) + 2r_y(\cos \beta_2 - 1)$$

If the identity $\sin^2(\beta/2) = (1 - \cos \beta)/2$ is used, the amplification factor is

$$G = 1 - 4r_x \sin^2 \frac{\beta_1}{2} - 4r_y \sin^2 \frac{\beta_2}{2}$$

Thus for stability, $|1 - 4r_x \sin^2(\beta_1/2) - 4r_y \sin^2(\beta_2/2)| \leqslant 1$, which is true only if $(4r_x \sin^2 \beta_1/2 + 4r_y \sin^2 \beta_2/2) \leqslant 2$. The stability requirement is then $(r_x + r_y) \leqslant \frac{1}{2}$ or $\alpha \Delta t[1/(\Delta x)^2 + 1/(\Delta y)^2] \leqslant \frac{1}{2}$. This is similar to the analysis of the same method for the 1-D case but shows that the effective time step in two dimensions is reduced. This example was easily completed, but in general, a stability analysis in more than a single space dimension and time is difficult. Frequently, the stability must be determined by computing the magnitude of the amplification factor for different values of r_x and r_y.

3.6.2 Stability Analysis for Systems of Equations

The previous discussion illustrates how the von Neumann analysis can be used to evaluate stability for a single equation. The basic idea used in this technique also provides a useful method of viewing stability for systems of equations. Systems of equations encountered in fluid mechanics and heat transfer can often be written in the form

$$\frac{\partial \mathbf{E}}{\partial t} + \frac{\partial \mathbf{F}}{\partial x} = 0 \tag{3.113}$$

where $\mathbf{E}$ and $\mathbf{F}$ are vectors and $\mathbf{F} = \mathbf{F}(\mathbf{E})$. In general, this system of equations is nonlinear. In order to perform a linear stability analysis, we rewrite the system as

$$\frac{\partial \mathbf{E}}{\partial t} + \left[\frac{\partial \mathbf{F}}{\partial \mathbf{E}}\right]\frac{\partial \mathbf{E}}{\partial x} = 0 \tag{3.114}$$

or

$$\frac{\partial \mathbf{E}}{\partial t} + [A]\frac{\partial \mathbf{E}}{\partial x} = 0$$

where $[A]$ is the Jacobian matrix $[\partial \mathbf{F}/\partial \mathbf{E}]$. We locally linearize the system by holding $[A]$ constant while the $\mathbf{E}$ vector is advanced through a single time step. A similar linearization is used for a single nonlinear equation, permitting the application of the von Neumann method of the previous section.

For the sake of discussion, let us apply the Lax method to this system. The result is

$$\mathbf{E}_j^{n+1} = \frac{1}{2}\left([I] + \frac{\Delta t}{\Delta x}[A]^n\right)\mathbf{E}_{j-1}^n + \frac{1}{2}\left([I] - \frac{\Delta t}{\Delta x}[A]^n\right)\mathbf{E}_{j+1}^n \quad (3.115)$$

where the notation is as previously defined and $[I]$ is the identity matrix. The stability of the difference equation can again be evaluated by applying the Fourier or von Neumann method. If a typical term of a Fourier series is substituted into Eq. (3.115), the following expression is obtained,

$$\mathbf{e}^{n+1}(k) = [G(\Delta t, k)]\mathbf{e}^n(k) \quad (3.116)$$

where

$$[G] = [I]\cos\beta - i\frac{\Delta t}{\Delta x}[A]\sin\beta \quad (3.117)$$

and $\mathbf{e}^n$ represents the Fourier coefficients of the typical term. The $[G]$ matrix is called the amplification matrix. This matrix is now dependent upon step size and frequency or wave number, i.e., $[G] = [G(\Delta t, k)]$. For a stable finite-difference calculation, the largest eigenvalue of $[G]$, $\sigma_{\max}$, must obey

$$|\sigma_{\max}| \leqslant 1 \quad (3.118)$$

This leads to the requirement that

$$\left|\lambda_{\max}\frac{\Delta t}{\Delta x}\right| \leqslant 1 \quad (3.119)$$

where $\lambda_{\max}$ is the largest eigenvalue of the $[A]$ matrix, i.e., the Jacobian matrix of the system. A simple example to demonstrate this is of value.

Example 3.6 Determine the stability requirement necessary for solving the system of first-order equations

$$\frac{\partial u}{\partial t} + c\frac{\partial v}{\partial x} = 0$$

$$\frac{\partial v}{\partial t} + c\frac{\partial u}{\partial x} = 0$$

using the Lax method.

Solution In this problem

$$\mathbf{E} = \begin{bmatrix} u \\ v \end{bmatrix}$$

and

$$\frac{\partial \mathbf{E}}{\partial t} + [A]\frac{\partial \mathbf{E}}{\partial x} = 0$$

where

$$[A] = \begin{bmatrix} 0 & c \\ c & 0 \end{bmatrix}$$

Thus, the maximum eigenvalue of [A] is c, and the stability requirement is the usual CFL condition

$$\left| c \frac{\Delta t}{\Delta x} \right| \leqslant 1$$

It should be noted that the stability analysis presented above does not include the effect of boundary conditions even though a matrix notation for the system is used. The influence of boundary conditions is easily included for systems of difference equations.

Equation (3.116) shows that the stability of a finite-difference operator is related to the amplification matrix. We may also write Eq. (3.116) as

$$\mathbf{e}^{n+1}(k) = [G(\Delta t, k)]^n [\mathbf{e}^1(k)] \tag{3.120}$$

The stability condition (Richtmyer and Morton, 1967) requires that for some positive τ, the matrices $[G(\Delta t, k)]^n$ be uniformly bounded for

$$0 < \Delta t < \tau$$
$$0 \leqslant n \Delta t \leqslant T$$

for all k, where T is the maximum time. This leads to the *von Neumann necessary condition* for stability, which is

$$|\sigma_i(\Delta t, k)| \leqslant 1 + O(\Delta t) \qquad 0 < \Delta t < \tau \tag{3.121}$$

for each eigenvalue and wave number, where σ_i represents the eigenvalues of $[G(\Delta t, k)]$. For a scalar equation, Eq. (3.121) reduces to

$$|G| \leqslant 1 + O(\Delta t)$$

The stability requirement used in previous examples required that the maximum eigenvalue have a modulus less than or equal to 1. Clearly, that requirement is more stringent than Eq. (3.121). The von Neumann necessary condition provides that local growth $c\,\Delta t$ can be acceptable and, in fact, must be possible in many physical problems. The classical example illustrating this point is the heat equation with a source term.

Example 3.7 Suppose we wish to solve the heat equation with a source term

$$\frac{\partial u}{\partial t} = \alpha \frac{\partial^2 u}{\partial x^2} + cu$$

using the simple explicit finite-difference method. Determine the stability requirement.

Solution If a Fourier stability analysis is performed, the amplification factor is

$$G = 1 - 4r \sin^2 \frac{\beta}{2} + c\,\Delta t$$

This shows that the solution of the difference equation may grow with time and still satisfy the von Neumann necessary condition. Physical insight must be used when the stability of a finite-difference method is investigated. One must

recognize that for hyperbolic systems the strict condition less than or equal to 1 should be used. Hyperbolic equations are wave-like and do not possess solutions that increase exponentially with time.

We have investigated stability of various finite-difference methods by using the von Neumann method. If the influence of boundary conditions on stability is desired, we must use the *matrix method.* This is most easily demonstrated by applying the Lax method to solve the 1-D linear wave equation:

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = 0$$

Assume that an array of m points is used to solve this problem and that the boundary conditions are periodic, i.e.,

$$u^n_{m+1} = u^n_1 \tag{3.122}$$

If the Lax method is applied to this problem, a system of algebraic equations is generated that has the form

$$\mathbf{u}^{n+1} = [X]\mathbf{u}^n \tag{3.123}$$

where

$$\mathbf{u}^n = [u_1, u_2, \ldots, u_m]^T \tag{3.124}$$

and

$$[X] = \begin{bmatrix} 0 & \frac{1-\nu}{2} & 0 & \cdots & \frac{1+\nu}{2} \\ \frac{1+\nu}{2} & 0 & \frac{1-\nu}{2} & \cdots & 0 \\ 0 & \frac{1+\nu}{2} & & & \vdots \\ \vdots & 0 & & & \vdots \\ & \vdots & & & \frac{1-\nu}{2} \\ \frac{1-\nu}{2} & 0 & \cdots & \frac{1+\nu}{2} & 0 \end{bmatrix} \tag{3.125}$$

The stability of the finite-difference calculation in Eq. (3.123) is governed by the eigenvalue structure of $[X]$. Since $[X]$ was formed assuming periodic boundary conditions, only the three diagonals noted in Eq. (3.125) and the two corner elements contribute to the calculation. This matrix is called an aperiodic matrix.

For matrices of the form

$$\begin{bmatrix} a_1 & a_2 & 0 & \cdot & \cdot & \cdot & a_0 \\ a_0 & a_1 & a_2 & 0 & \cdot & \cdot & 0 \\ 0 & a_0 & \ddots & \ddots & \ddots & & \cdot \\ \cdot & \cdot & \ddots & \ddots & \ddots & & \cdot \\ \cdot & \cdot & & \ddots & \ddots & \ddots & \cdot \\ \cdot & \cdot & & & \ddots & \ddots & a_2 \\ a_2 & 0 & \cdot & \cdot & \cdot & a_0 & a_1 \end{bmatrix} \tag{3.126}$$

the eigenvalues are given by

$$\lambda_j = a_1 + (a_0 + a_2)\cos\frac{2\pi}{m}(j-1) + i(a_0 - a_2)\sin\frac{2\pi}{m}(j-1)$$

In this case a_0, a_1, and a_2 have the values

$$a_0 = \frac{1+\nu}{2} \qquad a_1 = 0 \qquad a_2 = \frac{1-\nu}{2}$$

and the eigenvalues are

$$\lambda_j = \cos\frac{2\pi}{m}(j-1) + i\nu\sin\frac{2\pi}{m}(j-1) \tag{3.127}$$

The numerical method is thus stable if $|\nu| \leqslant 1$, i.e., if the CFL condition is satisfied. This shows that an analysis based upon the matrix operator associated with the Lax method yields the same stability requirement as previously derived for the simple wave equation. For periodic boundary conditions, the Fourier and matrix method yield virtually identical results. Another example is needed in order to demonstrate the effect of boundary conditions and the discreteness of the mesh.

Example 3.8 As in the previous example, assume that the Lax method is used to solve the first-order linear wave equation. If a four-point mesh is used, special treatment is needed to enforce the boundary conditions at the first and fourth points. For simplicity we set u at the first point equal to a constant value for all time, so the equation for the first point reads

$$u_1^{n+1} = u_1^n$$

Since we are computing a solution to the wave equation, the value of u_4 cannot be arbitrarily chosen. It must be consistent with the way the solution is propagated. We elect to set

$$u_4^{n+1} = u_3^n$$

which determines the boundary value from the interior solution.

Solution For the present boundary condition treatment the $[X]$ matrix becomes

$$[X] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \dfrac{1+\nu}{2} & 0 & \dfrac{1-\nu}{2} & 0 \\ 0 & \dfrac{1+\nu}{2} & 0 & \dfrac{1-\nu}{2} \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

The eigenvalues are easily computed and are

$$\lambda_1 = 1$$

$$\lambda_2 = 0$$

$$\lambda_{3,4} = \pm\tfrac{1}{2}\sqrt{(1-\nu)(3+\nu)}$$

Using the requirement that $|\lambda| \leqslant 1$ for stability, the restriction on ν is not the usual CFL condition but is

$$(-\sqrt{8} - 1) \leqslant \nu \leqslant (\sqrt{8} - 1)$$

The CFL condition is altered by the boundary conditions in this example, as is normally the case.

It is clear that the boundary conditions on the mesh are included in the matrix method. This means that the influence of boundary conditions on stability is automatically included if the matrix analysis is used. Unfortunately, a closed-form solution for the eigenvalues is usually not available for arbitrary end boundary conditions.

The treatment of stability presented in this section has included the Fourier (von Neumann) method and the matrix method of analysis. These two techniques are probably the most widely used to determine the stability of numerical schemes. Other methods of analyzing stability have been devised and are frequently very convenient to use. The works of Hirt (1968) and Warming and Hyett (1974) are typical of these techniques. A more comprehensive mathematical analysis of stability including many theorems and proofs is contained in the book by Richtmyer and Morton (1967).

PROBLEMS

3.1 Verify that

$$\left.\frac{\partial^3 u}{\partial x^3}\right)_{i,j} = \frac{\Delta_x^3 u_{i,j}}{(\Delta x)^3} + O(\Delta x)$$

3.2 Consider the function $f(x) = e^x$. Using a mesh increment $\Delta x = 0.1$, determine $f'(x)$ at $x = 2$ with the forward-difference formula, Eq. (3.26), the central-difference formula, Eq. (3.28), and the second-order three-point formula, Eq. (3.29). Compare the results with the exact value. Repeat the comparisons for $\Delta x = 0.2$. Have the order estimates for truncation errors been a reliable guide? Discuss this point.

3.3 Verify whether or not the following difference representation for the continuity equation for a 2-D steady incompressible flow has the conservation property:

$$\frac{(u_{i+1,j} + u_{i+1,j-1} - u_{i,j} - u_{i,j-1})}{2\,\Delta x} + \frac{(v_{i+1,j} - v_{i+1,j-1})}{\Delta y} = 0$$

where u and v are the x and y components of velocity, respectively.

3.4 Repeat Prob. 3.3, for the following difference representation for the continuity equation:

$$\frac{(u_{i+1,j} - u_{i-1,j})}{2\,\Delta x} + \frac{(v_{i,j+1} - v_{i,j-1})}{2\,\Delta y} = 0$$

3.5 Consider the nonlinear equation

$$u\frac{\partial u}{\partial x} = \mu\frac{\partial^2 u}{\partial y^2}$$

where μ is a constant.

(*a*) Is this equation in conservative form? If not, can you suggest a conservative form for the equation?

(*b*) Develop a finite-difference formulation for this equation using the integral approach.

3.6 Verify the approximation to $\partial^2 u/\partial x\,\partial y$ given by Eq. (3.50) in Table 3.2.

3.7 Verify the approximation to $\partial^2 u/\partial x^2$ given by Eq. (3.40) in Table 3.1.

3.8 Verify Eq. (3.79) in Table 3.3.

3.9 Verify Eq. (3.80) in Table 3.3.

3.10 Verify the following finite-difference approximation for use in two dimensions at the point (i, j). Assume $\Delta x = \Delta y = h$.

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \frac{u_{i+1,j-1} + u_{i+1,j+1} + u_{i-1,j-1} + u_{i-1,j+1} - 4u_{i,j}}{2h^2} + O(h^2)$$

3.11 Develop a finite-difference approximation with T.E. of $O(\Delta y)$ for $\partial^2 u/\partial y^2$ at point (i, j) using $u_{i,j}, u_{i,j+1}, u_{i,j-1}$ when the grid spacing is *not* uniform. Use the Taylor-series method. Can you devise a three-point scheme with second-order accuracy with unequal spacing? Before you draw your final conclusions, consider the use of compact implicit representations.

3.12 Establish the T.E. of the following finite-difference approximation to $\partial u/\partial y$ at the point (i, j) for a uniform mesh:

$$\frac{\partial u}{\partial y} \simeq \frac{-3u_{i,j} + 4u_{i,j+1} - u_{i,j+2}}{2\,\Delta y}$$

What is the order of the T.E.?

3.13 Investigate the T.E. of the following finite-difference approximation for a uniform mesh:

$$\left.\frac{\partial u}{\partial x}\right)_{i,j} \simeq \frac{1}{2h}\frac{\bar{\delta}_x u_{i,j}}{1 + \delta_x^2/6}$$

3.14 Utilize Taylor-series expansions about the point $(n + \frac{1}{2}, j)$ to determine the T.E. of the Crank-Nicolson representation of the heat equation, Eq. (3.71*a*). Compare these results with the T.E. obtained from Taylor-series expansions about point n, j.

3.15 Develop a finite-difference approximation with T.E. of $O(\Delta y)^2$ for $\partial T/\partial y$ at point (i, j) using $T_{i,j}$, $T_{i,j+1}$, and $T_{i,j+2}$ when the grid spacing is *not* uniform.

3.16 Determine the T.E. of the following finite-difference approximation for $\partial u/\partial x$ at point (i, j) when the grid spacing is *not* uniform:

$$\left.\frac{\partial u}{\partial x}\right)_{i,j} \simeq \frac{u_{i+1,j} - (\Delta x_+/\Delta x_-)^2 u_{i-1,j} - [1 - (\Delta x_+/\Delta x_-)^2]u_{i,j}}{\Delta x_-(\Delta x_+/\Delta x_-)^2 + \Delta x_+}$$

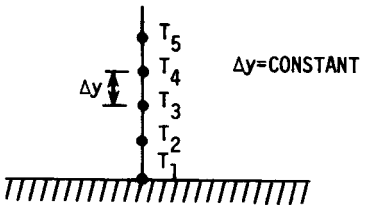

Figure P3.1

3.17 Suppose that a finite-difference solution has been obtained for the temperature T, near but not at an adiabatic boundary (i.e., $\partial T/\partial y = 0$ at the boundary) (Fig. P3.1). In most instances, it would be necessary or desirable to evaluate the temperature at the boundary point itself. For this case of an adiabatic boundary, develop expressions for the temperature at the boundary T_1, in terms of temperatures at neighboring points T_2, T_3, etc., by assuming that the temperature distribution in the neighborhood of the boundary is

(*a*) a straight line

(*b*) a second-degree polynomial

(*c*) a cubic polynomial (you only need to indicate how you would derive this one).

Indicate the order of the T.E. in each of the above approximations used to evaluate T_1.

3.18 Consider a steady-state conduction problem governed by Laplace's equation with convective boundary conditions (see Fig. P3.2). The formal statement of the boundary condition is $-k\,\partial T/\partial y)_{\text{bdy}} = h(T_w - T_\infty)$, which can be readily cast into finite-difference form as $-k[(T_0 - T_\infty)/\Delta y] + O(\Delta y) = h(T_0 - T_\infty)$. Use the control-volume approach to develop an expression for the boundary condition at point 0. Evaluate the T.E. in this expression assuming that Laplace's equation applies at the boundary point.

3.19 Consider a heat conduction problem governed by $\partial T/\partial t = \alpha(\partial^2 T/\partial x^2)$. Develop a finite-difference representation for this equation by the control-volume approach. Do not assume that the grid is uniform.

3.20 For 2-D steady-state conduction in a solid, apply the control-volume method to derive an appropriate difference expression for the boundary temperature in control volume B in Fig. 3.7 for *adiabatic wall* boundary conditions.

3.21 Solve the 1-D heat equation using forward-time centered-space differences with $\alpha(\Delta t/\Delta x^2) = \frac{1}{2}$. Let the grid consist of five points, including three interior and two boundary points. Assume a constant unity wall temperature and a zero initial temperature on the interior. Complete this calculation for 10 integration steps. Compare your results with those obtained in the example of Section 3.3.4.

3.22 Refer to Fig. 3.10. Following the methodology illustrated in the text material associated with Fig. 3.10, develop an appropriate finite-volume expression for the heat flux across the boundary from *c* to *d*.

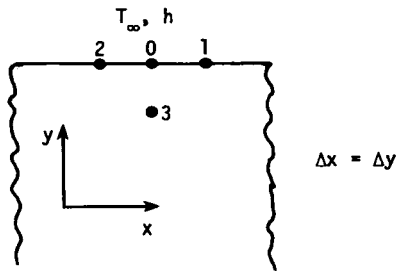

Figure P3.2

3.23 Refer to Fig. 3.10 and the associated text material. In an example in Section 3.5, an expression was developed for the heat flux across boundary *a-b* for control volume B (Eq. 3.100). Following a similar methodology, develop an appropriate finite-volume expression for the heat flow into volume B across the boundary from *d* to *a*. If the difference scheme is to be conservative, this heat flow should be equal in magnitude but opposite in direction to the inflow computed for volume A across boundary *d-a*. Check to see if this is true.

3.24 Show that the amplification factor derived for the finite-difference solution of the heat equation, Eq. (3.101), could be obtained by direct substitution of a solution of the form

$$u_j^n = \sum_{-\infty}^{+\infty} C_m g_m^n e^{ik_x x}$$

In this form C_m represent the Fourier coefficients of the initial error distribution and g_m is the amplification factor. Identify g_m with Eq. (3.106). Discuss the convergence of the solution and relate your conclusions to the Lax equivalence theorem.

3.25 Use a von Neumann stability analysis to show for the wave equation that a simple explicit Euler predictor using central differencing in space is unstable. The difference equation is

$$u_j^{n+1} = u_j^n - c\frac{\Delta t}{\Delta x}\left(\frac{u_{j+1}^n - u_{j-1}^n}{2}\right)$$

Now show that the same difference method is stable when written as the implicit formula

$$u_j^{n+1} = u_j^n - c\frac{\Delta t}{\Delta x}\left(\frac{u_{j+1}^{n+1} - u_{j-1}^{n+1}}{2}\right)$$

3.26 The DuFort-Frankel method for solving the heat equation requires solution of the difference equation

$$\frac{u_j^{n+1} - u_j^{n-1}}{2\,\Delta t} = \frac{\alpha}{(\Delta x)^2}(u_{j+1}^n - u_j^{n+1} - u_j^{n-1} + u_{j-1}^n)$$

Develop the stability requirements necessary for the solution of this equation.

3.27 Prove that the CFL condition is the stability requirement when the Lax-Wendroff method is applied to solve the simple 1-D wave equation. The difference equation is of the form

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

3.28 An implicit scheme for solving the heat equation is given by

$$u_j^{n+1} = u_j^n + \frac{\alpha\,\Delta t}{(\Delta x)^2}\left[\frac{1}{3}(u_{j+1}^{n+1} - 2u_j^{n+1} + u_{j-1}^{n+1}) + \frac{2}{3}(u_{j+1}^n - 2u_j^n + u_{j-1}^n)\right]$$

Apply the Fourier stability analysis to this scheme and determine the stability restrictions, if any.

3.29 An implicit scheme for solving the first-order wave equation is given by

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

Apply the Fourier stability analysis to this scheme and determine the stability restrictions, if any.

3.30 The leap frog method for solving the 1-D wave equation is given by

$$\frac{u_j^{n+1} - u_j^{n-1}}{2\,\Delta t} + c\frac{u_{j+1}^n - u_{j-1}^n}{2\,\Delta x} = 0$$

Apply the Fourier stability analysis to this method, and determine the stability restrictions, if any.

3.31 Determine the stability requirement necessary to solve the 1-D heat equation with a source term

$$\frac{\partial u}{\partial t} = \alpha \frac{\partial^2 u}{\partial x^2} + ku$$

Use the central-space, forward-time difference method. Does the von Neumann necessary condition, Eq. (3.121), make physical sense for this type of computational problem?

3.32 Use the matrix method to determine the stability of the Lax method used to solve the first-order wave equation on a mesh with two interior points and two boundary points. Assume the boundaries are held at constant values $u_{\text{left}} = 1$, $u_{\text{right}} = 0$.

3.33 Use the matrix method and evaluate the stability of the numerical method used in Prob. 3.21 for the heat equation using a five-point mesh. How many frequencies must one be concerned with in this case?

3.34 In attempting to solve a simple PDE, a system of finite-difference equations of the form $u_j^{n+1} = [A]u_j^n$ has evolved, where

$$[A] = \begin{bmatrix} 1+\nu & \nu & 0 \\ 0 & 1+\nu & \nu \\ -\nu & 0 & 1+\nu \end{bmatrix}$$

Investigate the stability of this scheme.

3.35 The application of a finite-difference scheme to the heat equation on a three-point grid results in the following system of equations:

$$u_j^{n+1} = [A]u_j^n$$

where

$$[A] = \begin{bmatrix} 1 & 0 & 0 \\ r & 1-2r & r \\ 0 & 0 & 1 \end{bmatrix}$$

and $r = \alpha \Delta t/(\Delta x)^2$. Determine the stability of this scheme.

3.36 The upstream scheme

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + c\frac{u_j^n - u_{j-1}^n}{\Delta x} = 0$$

is used to solve the wave equation on a four-point grid for the boundary conditions

$$u_1 = 1 \qquad u_4^{n+1} = u_3^n$$

and the initial conditions ($n = 1$)

$$u_1^1 = 1 \qquad u_2^1 = u_3^1 = u_4^1 = 0$$

Use the matrix method to determine the stability restrictions for this method.