Chapter 3

FINITE-DIFFERENCE APPROXIMATIONS

In common with the equations governing unsteady fluid flow, our model equations contain partial derivatives with respect to both space and time. One can approximate these simultaneously and then solve the resulting difference equations. Alternatively, one can approximate the spatial derivatives first, thereby producing a system of ordinary differential equations. The time derivatives are approximated next, leading to a *time-marching method* which produces a set of difference equations. This is the approach emphasized here. In this chapter, the concept of finite-difference approximations to partial derivatives is presented. These can be applied either to spatial derivatives or time derivatives. Our emphasis in this chapter is on spatial derivatives; time derivatives are treated in Chapter 6. Strategies for applying these finite-difference approximations will be discussed in Chapter 4.

All of the material below is presented in a Cartesian system. We emphasize the fact that quite general classes of meshes expressed in general curvilinear coordinates in *physical space* can be transformed to a uniform Cartesian mesh with equispaced intervals in a so-called *computational space*, as shown in Figure 3.1. The *computational space* is uniform; all the geometric variation is absorbed into variable coefficients of the transformed equations. For this reason, in much of the following accuracy analysis, we use an equispaced Cartesian system without being unduly restrictive or losing practical application.

3.1 Meshes and Finite-Difference Notation

The simplest mesh involving both time and space is shown in Figure 3.2. Inspection of this figure permits us to define the terms and notation needed to describe finite-

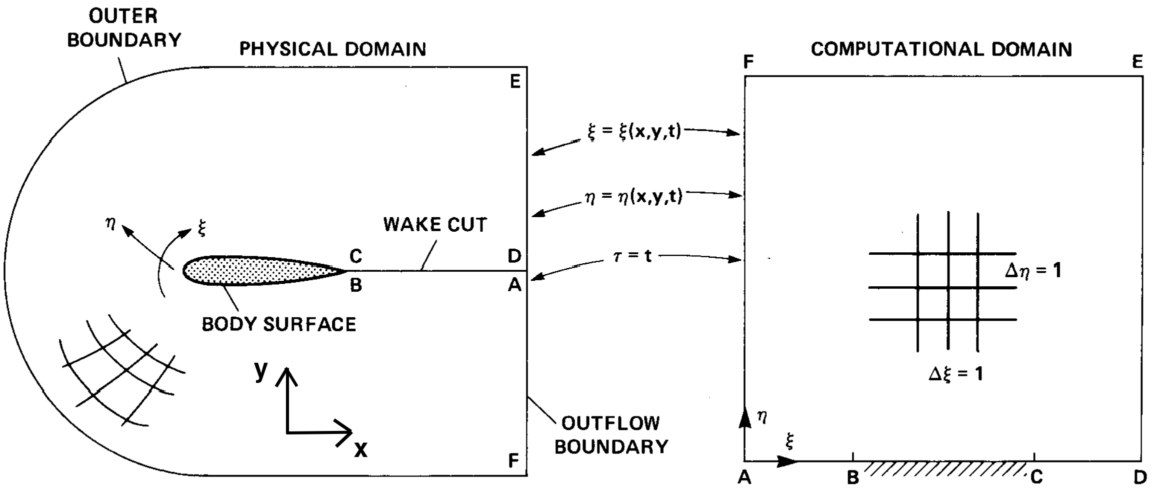

Figure 3.1: Physical and computational spaces.

difference approximations. In general, the dependent variables, u, for example, are functions of the independent variables t, and x, y, z. For the first several chapters we consider primarily the 1-D case $u = u(x, t)$. When only one variable is denoted, dependence on the other is assumed. The mesh index for x is *always* j, and that for t is *always* n. Then on an equispaced grid

$$x = x_j = j\Delta x \tag{3.1}$$
$$t = t_n = n\Delta t = nh \tag{3.2}$$

where Δx is the spacing in x and Δt the spacing in t, as shown in Figure 3.2. Note that $h = \Delta t$ throughout. Later k and l are used for y and z in a similar way. When n, j, k, l are used for other purposes (which is sometimes necessary), local context should make the meaning obvious.

The convention for subscript and superscript indexing is as follows:

$$\begin{aligned} u(t + kh) &= u([n + k]h) = u_{n+k} \\ u(x + m\Delta x) &= u([j + m]\Delta x) = u_{j+m} \\ u(x + m\Delta x, t + kh) &= u_{j+m}^{(n+k)} \end{aligned} \tag{3.3}$$

Notice that when used alone, both the time and space indices appear as a subscript, but when used together, time is always a superscript and is usually enclosed with parentheses to distinguish it from an exponent.

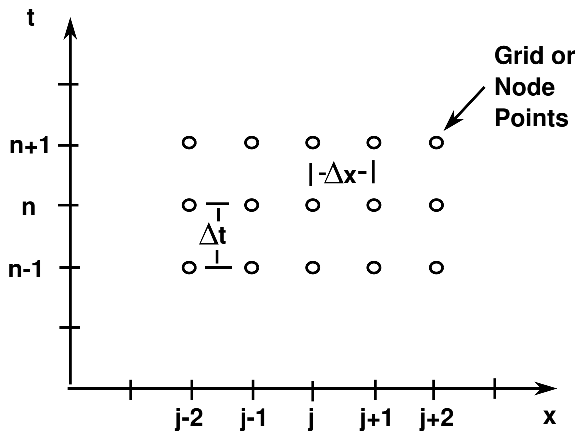

Figure 3.2: Space-time grid arrangement.

Derivatives are expressed according to the usual conventions. Thus for partial derivatives in space or time we use interchangeably

$$\partial_x u = \frac{\partial u}{\partial x}, \quad \partial_t u = \frac{\partial u}{\partial t}, \quad \partial_{xx} u = \frac{\partial^2 u}{\partial x^2}, \quad \text{etc.} \tag{3.4}$$

For the ordinary time derivative in the study of ODE's we use

$$u' = \frac{du}{dt} \tag{3.5}$$

In this text, subscripts on dependent variables are never used to express derivatives. Thus u_x will not be used to represent the first derivative of u with respect to x.

The notation for difference approximations follows the same philosophy, but (with one exception) it is not unique. By this we mean that the symbol δ is used to represent a difference approximation to a derivative such that, for example,

$$\delta_x \approx \partial_x, \qquad \delta_{xx} \approx \partial_{xx} \tag{3.6}$$

but the precise nature (and order) of the approximation is not carried in the symbol δ. Other ways are used to determine its precise meaning. The one exception is the symbol Δ, which is defined such that

$$\Delta t_n = t_{n+1} - t_n, \quad \Delta x_j = x_{j+1} - x_j, \quad \Delta u_n = u_{n+1} - u_n, \quad \text{etc.} \tag{3.7}$$

When there is no subscript on Δt or Δx, the spacing is uniform.

3.2 Space Derivative Approximations

A difference approximation can be generated or evaluated by means of a simple Taylor series expansion. For example, consider $u(x,t)$ with t fixed. Then, following the notation convention given in Eqs. 3.1 to 3.3, $x = j\Delta x$ and $u(x + k\Delta x) = u(j\Delta x + k\Delta x) = u_{j+k}$. Expanding the latter term about x gives[1]

$$u_{j+k} = u_j + (k\Delta x)\left(\frac{\partial u}{\partial x}\right)_j + \frac{1}{2}(k\Delta x)^2\left(\frac{\partial^2 u}{\partial x^2}\right)_j + \cdots + \frac{1}{n!}(k\Delta x)^n\left(\frac{\partial^n u}{\partial x^n}\right)_j + \cdots \quad (3.8)$$

Local difference approximations to a given partial derivative can be formed from linear combinations of u_j and u_{j+k} for $k = \pm 1, \pm 2, \cdots$.

For example, consider the Taylor series expansion for u_{j+1}:

$$u_{j+1} = u_j + (\Delta x)\left(\frac{\partial u}{\partial x}\right)_j + \frac{1}{2}(\Delta x)^2\left(\frac{\partial^2 u}{\partial x^2}\right)_j + \cdots + \frac{1}{n!}(\Delta x)^n\left(\frac{\partial^n u}{\partial x^n}\right)_j + \cdots \quad (3.9)$$

Now subtract u_j and divide by Δx to obtain

$$\frac{u_{j+1} - u_j}{\Delta x} = \left(\frac{\partial u}{\partial x}\right)_j + \frac{1}{2}(\Delta x)\left(\frac{\partial^2 u}{\partial x^2}\right)_j + \cdots \quad (3.10)$$

Thus the expression $(u_{j+1} - u_j)/\Delta x$ is a reasonable approximation for $\left(\frac{\partial u}{\partial x}\right)_j$ as long as Δx is small relative to some pertinent length scale. Next consider the space difference approximation $(u_{j+1} - u_{j-1})/(2\Delta x)$. Expand the terms in the numerator about j and regroup the result to form the following equation

$$\frac{u_{j+1} - u_{j-1}}{2\Delta x} - \left(\frac{\partial u}{\partial x}\right)_j = \frac{1}{6}\Delta x^2\left(\frac{\partial^3 u}{\partial x^3}\right)_j + \frac{1}{120}\Delta x^4\left(\frac{\partial^5 u}{\partial x^5}\right)_j \ldots \quad (3.11)$$

When expressed in this manner, it is clear that the discrete terms on the left side of the equation represent a first derivative with a certain amount of error which appears on the right side of the equal sign. It is also clear that the error depends on the grid spacing to a certain order. The error term containing the grid spacing to the lowest power gives the order of the method. From Eq. 3.10, we see that the expression $(u_{j+1} - u_j)/\Delta x$ is a first-order approximation to $\left(\frac{\partial u}{\partial x}\right)_j$. Similarly, Eq. 3.11 shows that $(u_{j+1} - u_{j-1})/(2\Delta x)$ is a second-order approximation to a first derivative. The latter is referred to as the three-point centered difference approximation, and one often sees the summary result presented in the form

$$\left(\frac{\partial u}{\partial x}\right)_j = \frac{u_{j+1} - u_{j-1}}{2\Delta x} + O(\Delta x^2) \quad (3.12)$$

[1]We assume that $u(x,t)$ is continuously differentiable.

3.3 Finite-Difference Operators

3.3.1 Point Difference Operators

Perhaps the most common examples of finite-difference formulas are the three-point centered-difference approximations for the first and second derivatives:[2]

$$\left(\frac{\partial u}{\partial x}\right)_j = \frac{1}{2\Delta x}(u_{j+1} - u_{j-1}) + O(\Delta x^2) \tag{3.13}$$

$$\left(\frac{\partial^2 u}{\partial x^2}\right)_j = \frac{1}{\Delta x^2}(u_{j+1} - 2u_j + u_{j-1}) + O(\Delta x^2) \tag{3.14}$$

These are the basis for *point difference operators* since they give an approximation to a derivative at one discrete point in a mesh in terms of surrounding points. However, neither of these expressions tells us how other points in the mesh are differenced or how boundary conditions are enforced. Such additional information requires a more sophisticated formulation.

3.3.2 Matrix Difference Operators

Consider the relation

$$(\delta_{xx} u)_j = \frac{1}{\Delta x^2}(u_{j+1} - 2u_j + u_{j-1}) \tag{3.15}$$

which is a point difference approximation to a second derivative. Now let us derive a *matrix* operator representation for the same approximation. Consider the four point mesh with boundary points at a and b shown below. Notice that when we speak of "the number of points in a mesh", we mean the number of *interior* points excluding the boundaries.

	a	1	2	3	4	b
$x =$	0	–	–	–	–	π
$j =$		1	·	·	M	

Four point mesh. $\Delta x = \pi/(M+1)$

Now impose Dirichlet boundary conditions, $u(0) = u_a,\ u(\pi) = u_b$ and use the centered difference approximation given by Eq. 3.15 at every point in the mesh. We

[2]We will derive the second derivative operator shortly.

arrive at the four equations:

$$\begin{aligned}
(\delta_{xx}u)_1 &= \frac{1}{\Delta x^2}(u_a - 2u_1 + u_2) \\
(\delta_{xx}u)_2 &= \frac{1}{\Delta x^2}(u_1 - 2u_2 + u_3) \\
(\delta_{xx}u)_3 &= \frac{1}{\Delta x^2}(u_2 - 2u_3 + u_4) \\
(\delta_{xx}u)_4 &= \frac{1}{\Delta x^2}(u_3 - 2u_4 + u_b)
\end{aligned} \tag{3.16}$$

Writing these equations in the more suggestive form

$$\begin{array}{lclcccccccl}
(\delta_{xx}u)_1 & = & (& u_a & -2u_1 & +u_2 & & & & &)/\Delta x^2 \\
(\delta_{xx}u)_2 & = & (& & u_1 & -2u_2 & +u_3 & & & &)/\Delta x^2 \\
(\delta_{xx}u)_3 & = & (& & & u_2 & -2u_3 & +u_4 & & &)/\Delta x^2 \\
(\delta_{xx}u)_4 & = & (& & & & u_3 & -2u_4 & +u_b & &)/\Delta x^2
\end{array} \tag{3.17}$$

it is clear that we can express them in a vector-matrix form, and further, that the resulting matrix has a very special form. Introducing

$$\vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix}, \qquad \left(\vec{bc}\right) = \frac{1}{\Delta x^2}\begin{bmatrix} u_a \\ 0 \\ 0 \\ u_b \end{bmatrix} \tag{3.18}$$

and

$$A = \frac{1}{\Delta x^2}\begin{bmatrix} -2 & 1 & & \\ 1 & -2 & 1 & \\ & 1 & -2 & 1 \\ & & 1 & -2 \end{bmatrix} \tag{3.19}$$

we can rewrite Eq. 3.17 as

$$\delta_{xx}\vec{u} = A\vec{u} + \left(\vec{bc}\right) \tag{3.20}$$

This example illustrates a matrix difference operator. Each line of a matrix difference operator is based on a point difference operator, but the point operators used from line to line are not necessarily the same. For example, boundary conditions may dictate that the lines at or near the bottom or top of the matrix be modified. In the extreme case of the matrix difference operator representing a spectral method, none of the lines is the same. The matrix operators representing the three-point central-difference approximations for a first and second derivative with Dirichlet boundary conditions on a four-point mesh are

$$\delta_x = \frac{1}{2\Delta x}\begin{bmatrix} 0 & 1 & & \\ -1 & 0 & 1 & \\ & -1 & 0 & 1 \\ & & -1 & 0 \end{bmatrix}, \quad \delta_{xx} = \frac{1}{\Delta x^2}\begin{bmatrix} -2 & 1 & & \\ 1 & -2 & 1 & \\ & 1 & -2 & 1 \\ & & 1 & -2 \end{bmatrix} \tag{3.21}$$

As a further example, replace the fourth line in Eq. 3.16 by the following point operator for a Neumann boundary condition (See Section 3.6.):

$$(\delta_{xx}u)_4 = \frac{2}{3}\frac{1}{\Delta x}\left(\frac{\partial u}{\partial x}\right)_b - \frac{2}{3}\frac{1}{\Delta x^2}(u_4 - u_3) \tag{3.22}$$

where the boundary condition is

$$\left(\frac{\partial u}{\partial x}\right)_{x=\pi} = \left(\frac{\partial u}{\partial x}\right)_b \tag{3.23}$$

Then the matrix operator for a three-point central-differencing scheme at interior points and *a second-order approximation* for a Neumann condition on the right is given by

$$\delta_{xx} = \frac{1}{\Delta x^2}\begin{bmatrix} -2 & 1 & & \\ 1 & -2 & 1 & \\ & 1 & -2 & 1 \\ & & 2/3 & -2/3 \end{bmatrix} \tag{3.24}$$

Each of these matrix difference operators is a square matrix with elements that are all zeros except for those along bands which are clustered around the central diagonal. We call such a matrix a *banded matrix* and introduce the notation

$$B(M : a, b, c) = \begin{array}{c} \begin{bmatrix} b & c & & & \\ a & b & c & & \\ & & \ddots & & \\ & & a & b & c \\ & & & a & b \end{bmatrix} \end{array} \begin{array}{c} 1 \\ \\ \\ \vdots \\ M \end{array} \tag{3.25}$$

where the matrix dimensions are $M \times M$. Use of M in the argument is optional, and the illustration is given for a simple *tridiagonal* matrix although any number of bands is a possibility. A tridiagonal matrix without constants along the bands can be expressed as $B(\vec{a}, \vec{b}, \vec{c})$. The arguments for a banded matrix are always odd in number and the central one *always* refers to the central diagonal.

We can now generalize our previous examples. Defining $\vec{u}$ as[3]

$$\vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \\ u_M \end{bmatrix} \tag{3.26}$$

we can approximate the second derivative of $\vec{u}$ by

$$\delta_{xx}\vec{u} = \frac{1}{\Delta x^2} B(1, -2, 1)\vec{u} + \left(\vec{bc}\right) \tag{3.27}$$

where $\left(\vec{bc}\right)$ stands for the vector holding the Dirichlet boundary conditions on the left and right sides:

$$\left(\vec{bc}\right) = \frac{1}{\Delta x^2}[u_a, 0, \cdots, 0, u_b]^T \tag{3.28}$$

If we prescribe Neumann boundary conditions on the right side, as in Eqs. 3.24 and 3.22, we find

$$\delta_{xx}\vec{u} \;=\; \frac{1}{\Delta x^2} B(\vec{a}, \vec{b}, 1)\vec{u} + \left(\vec{bc}\right) \tag{3.29}$$

where

$$\begin{aligned} \vec{a} &= [1, 1, \cdots, 2/3]^T \\ \vec{b} &= [-2, -2, -2, \cdots, -2/3]^T \\ \left(\vec{bc}\right) &= \frac{1}{\Delta x^2}\left[u_a, 0, 0, \cdots, \frac{2\Delta x}{3}\left(\frac{\partial u}{\partial x}\right)_b\right]^T \end{aligned}$$

Notice that the matrix operators given by Eqs. 3.27 and 3.29 carry more information than the point operator given by Eq. 3.15. In Eqs. 3.27 and 3.29, the boundary conditions have been uniquely specified and it is clear that the same point operator has been applied at every point in the field except at the boundaries. The ability to specify in the matrix derivative operator the exact nature of the approximation at the

[3]Note that $\vec{u}$ is a function of time only since each element corresponds to one specific spatial location.

various points in the field including the boundaries permits the use of quite general constructions which will be useful later in considerations of stability.

Since we make considerable use of both matrix and point operators, it is important to establish a relation between them. A point operator is generally written for some derivative at the reference point j in terms of neighboring values of the function. For example

$$(\delta_x u)_j = a_2 u_{j-2} + a_1 u_{j-1} + b u_j + c_1 u_{j+1} \tag{3.30}$$

might be the point operator for a first derivative. The corresponding matrix operator has for its arguments the coefficients giving the weights to the values of the function at the various locations. A j-shift in the point operator corresponds to a diagonal shift in the matrix operator. Thus the matrix equivalent of Eq. 3.30 is

$$\delta_x \vec{u} = B(a_2, a_1, b, c_1, 0)\vec{u} \tag{3.31}$$

Note the addition of a zero in the fifth element which makes it clear that b is the coefficient of u_j.

3.3.3 Periodic Matrices

The above illustrated cases in which the boundary conditions are fixed. If the boundary conditions are *periodic*, the form of the matrix operator changes. Consider the eight-point periodic mesh shown below. This can either be presented on a linear mesh with repeated entries, or more suggestively on a circular mesh as in Figure 3.3. When the mesh is laid out on the perimeter of a circle, it doesn't really matter where the numbering starts as long as it "ends" at the point just preceding its starting location.

$\cdots$	7	8	1	2	3	4	5	6	7	8	1	2	$\cdots$
$x =$	–	–	0	–	–	–	–	–	–	–	2π	–	
$j =$			0	1	·	·	·	·	·	·	M		

Eight points on a linear periodic mesh. $\Delta x = 2\pi/M$

The matrix that represents differencing schemes for scalar equations on a periodic mesh is referred to as a *periodic* matrix. A typical periodic tridiagonal matrix operator

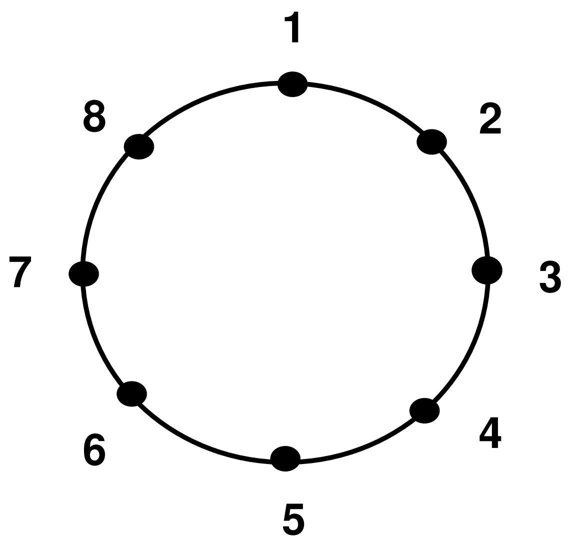

Figure 3.3: Eight points on a circular mesh.

with nonuniform entries is given for a 6-point mesh by

$$B_p(6 : \vec{a}, \vec{b}, \vec{c}) = \begin{bmatrix} b_1 & c_2 & & & & a_6 \\ a_1 & b_2 & c_3 & & & \\ & a_2 & b_3 & c_4 & & \\ & & a_3 & b_4 & c_5 & \\ & & & a_4 & b_5 & c_6 \\ c_1 & & & & a_5 & b_6 \end{bmatrix} \tag{3.32}$$

3.3.4 Circulant Matrices

In general, as shown in the example, the elements along the diagonals of a periodic matrix are not constant. However, a special subset of a periodic matrix is the circulant matrix, formed when the elements along the various bands are constant. Circulant matrices play a vital role in our analysis. We will have much more to say about them later. The most general circulant matrix of order 4 is

$$\begin{bmatrix} b_0 & b_1 & b_2 & b_3 \\ b_3 & b_0 & b_1 & b_2 \\ b_2 & b_3 & b_0 & b_1 \\ b_1 & b_2 & b_3 & b_0 \end{bmatrix} \tag{3.33}$$

Notice that each row of a circulant matrix is shifted (see Figure 3.3) one element to the right of the one above it. The special case of a tridiagonal circulant matrix is given by

$$B_p(M : a, b, c) = \begin{array}{c@{\;}c} \begin{bmatrix} b & c & & & a \\ a & b & c & & \\ & & \ddots & & \\ & & a & b & c \\ c & & & a & b \end{bmatrix} & \begin{array}{c} 1 \\ \\ \\ \vdots \\ M \end{array} \end{array} \tag{3.34}$$

When the standard three-point central-differencing approximations for a first and second derivative, see Eq. 3.21, are used with periodic boundary conditions, they take the form

$$(\delta_x)_p = \frac{1}{2\Delta x}\begin{bmatrix} 0 & 1 & & -1 \\ -1 & 0 & 1 & \\ & -1 & 0 & 1 \\ 1 & & -1 & 0 \end{bmatrix} = \frac{1}{2\Delta x} B_p(-1, 0, 1)$$

and

$$(\delta_{xx})_p = \frac{1}{\Delta x^2}\begin{bmatrix} -2 & 1 & & 1 \\ 1 & -2 & 1 & \\ & 1 & -2 & 1 \\ 1 & & 1 & -2 \end{bmatrix} = \frac{1}{\Delta x^2} B_p(1, -2, 1) \tag{3.35}$$

Clearly, these special cases of periodic operators are also circulant operators. Later on we take advantage of this special property. Notice that there are no boundary condition vectors since this information is all interior to the matrices themselves.

3.4 Constructing Differencing Schemes of Any Order

3.4.1 Taylor Tables

The Taylor series expansion of functions about a fixed point provides a means for constructing finite-difference point-operators of any order. A simple and straightforward way to carry this out is to construct a "Taylor table," which makes extensive use of the expansion given by Eq. 3.8. As an example, consider Table 3.1, which represents a Taylor table for an approximation of a second derivative using three values of the function centered about the point at which the derivative is to be evaluated.

$$\left(\frac{\partial^2 u}{\partial x^2}\right)_j - \frac{1}{\Delta x^2}(a\,u_{j-1} + b\,u_j + c\,u_{j+1}) = ?$$

	u_j	$\Delta x \cdot \left(\frac{\partial u}{\partial x}\right)_j$	$\Delta x^2 \cdot \left(\frac{\partial^2 u}{\partial x^2}\right)_j$	$\Delta x^3 \cdot \left(\frac{\partial^3 u}{\partial x^3}\right)_j$	$\Delta x^4 \cdot \left(\frac{\partial^4 u}{\partial x^4}\right)_j$
$\Delta x^2 \cdot \left(\frac{\partial^2 u}{\partial x^2}\right)_j$			1		
$-a \cdot u_{j-1}$	$-a$	$-a \cdot (\text{-}1) \cdot \frac{1}{1}$	$-a \cdot (\text{-}1)^2 \cdot \frac{1}{2}$	$-a \cdot (\text{-}1)^3 \cdot \frac{1}{6}$	$-a \cdot (\text{-}1)^4 \cdot \frac{1}{24}$
$-b \cdot u_j$	$-b$				
$-c \cdot u_{j+1}$	$-c$	$-c \cdot (1) \cdot \frac{1}{1}$	$-c \cdot (1)^2 \cdot \frac{1}{2}$	$-c \cdot (1)^3 \cdot \frac{1}{6}$	$-c \cdot (1)^4 \cdot \frac{1}{24}$

Table 3.1. Taylor table for centered 3-point Lagrangian approximation to a second derivative.

The table is constructed so that some of the algebra is simplified. At the top of the table we see an expression with a question mark. This represents one of the questions that a study of this table can answer; namely, what is the local error caused by the use of this approximation? Notice that all of the terms in the equation appear in a column at the left of the table (although, in this case, Δx^2 has been multiplied into each term in order to simplify the terms to be put into the table). Then notice that at the head of each column there appears the common factor that occurs in the expansion of each term about the point j, that is,

$$\Delta x^k \cdot \left(\frac{\partial^k u}{\partial x^k}\right)_j \qquad k = 0, 1, 2, \cdots$$

The columns to the right of the leftmost one, under the headings, make up the Taylor table. Each entry is the coefficient of the term at the top of the corresponding column in the Taylor series expansion of the term to the left of the corresponding row. For example, the last row in the table corresponds to the Taylor series expansion of $-c\,u_{j+1}$:

$$\begin{aligned} -c\,u_{j+1} &= -c\,u_j - c \cdot (1) \cdot \frac{1}{1}\Delta x \cdot \left(\frac{\partial u}{\partial x}\right)_j - c \cdot (1)^2 \cdot \frac{1}{2}\Delta x^2 \cdot \left(\frac{\partial^2 u}{\partial x^2}\right)_j \\ &\quad -c \cdot (1)^3 \cdot \frac{1}{6}\Delta x^3 \cdot \left(\frac{\partial^3 u}{\partial x^3}\right)_j - c \cdot (1)^4 \cdot \frac{1}{24}\Delta x^4 \cdot \left(\frac{\partial^4 u}{\partial x^4}\right)_j - \cdots \end{aligned} \tag{3.36}$$

A Taylor table is simply a convenient way of forming linear combinations of Taylor series on a term by term basis.

Consider the sum of each of these columns. To maximize the order of accuracy of the method, we proceed from left to right and force, by the proper choice of a, b, and c, these sums to be zero. One can easily show that the sums of the first three columns are zero if we satisfy the equation

$$\begin{bmatrix} -1 & -1 & -1 \\ 1 & 0 & -1 \\ -1 & 0 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -2 \end{bmatrix}$$

The solution is given by $[a, b, c] = [1, -2, 1]$.

The columns that do not sum to zero constitute the error.

> We designate the first nonvanishing sum to be er_t, and refer to it as the Taylor series error.

In this case er_t occurs at the fifth column in the table (for this example all even columns will vanish by symmetry) and one finds

$$er_t = \frac{1}{\Delta x^2} \left[\frac{-a}{24} + \frac{-c}{24} \right] \Delta x^4 \left(\frac{\partial^4 u}{\partial x^4} \right)_j = \frac{-\Delta x^2}{12} \left(\frac{\partial^4 u}{\partial x^4} \right)_j \tag{3.37}$$

Note that Δx^2 has been divided through to make the error term consistent. We have just derived the familiar 3-point central-differencing point operator for a second derivative

$$\left(\frac{\partial^2 u}{\partial x^2} \right)_j - \frac{1}{\Delta x^2}(u_{j-1} - 2u_j + u_{j+1}) = O(\Delta x^2) \tag{3.38}$$

The Taylor table for a 3-point backward-differencing operator representing a first derivative is shown in Table 3.2.

$$\left(\frac{\partial u}{\partial x}\right)_j - \frac{1}{\Delta x}(a_2 u_{j-2} + a_1 u_{j-1} + b\,u_j) = ?$$

	u_j	$\Delta x \cdot \left(\frac{\partial u}{\partial x}\right)_j$	$\Delta x^2 \cdot \left(\frac{\partial^2 u}{\partial x^2}\right)_j$	$\Delta x^3 \cdot \left(\frac{\partial^3 u}{\partial x^3}\right)_j$	$\Delta x^4 \cdot \left(\frac{\partial^4 u}{\partial x^4}\right)_j$
$\Delta x \cdot \left(\frac{\partial u}{\partial x}\right)_j$		1			
$-a_2 \cdot u_{j-2}$	$-a_2$	$-a_2 \cdot (\text{-}2) \cdot \frac{1}{1}$	$-a_2 \cdot (\text{-}2)^2 \cdot \frac{1}{2}$	$-a_2 \cdot (\text{-}2)^3 \cdot \frac{1}{6}$	$-a_2 \cdot (\text{-}2)^4 \cdot \frac{1}{24}$
$-a_1 \cdot u_{j-1}$	$-a_1$	$-a_1 \cdot (\text{-}1) \cdot \frac{1}{1}$	$-a_1 \cdot (\text{-}1)^2 \cdot \frac{1}{2}$	$-a_1 \cdot (\text{-}1)^3 \cdot \frac{1}{6}$	$-a_1 \cdot (\text{-}1)^4 \cdot \frac{1}{24}$
$-b \cdot u_j$	$-b$				

Table 3.2. Taylor table for backward 3-point Lagrangian approximation to a first derivative.

This time the first three columns sum to zero if

$$\begin{bmatrix} -1 & -1 & -1 \\ 2 & 1 & 0 \\ -4 & -1 & 0 \end{bmatrix} \begin{bmatrix} a_2 \\ a_1 \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}$$

which gives $[a_2, a_1, b] = \frac{1}{2}[1, -4, 3]$. In this case the fourth column provides the leading truncation error term:

$$er_t = \frac{1}{\Delta x}\left[\frac{8a_2}{6} + \frac{a_1}{6}\right]\Delta x^3 \left(\frac{\partial^3 u}{\partial x^3}\right)_j = \frac{\Delta x^2}{3}\left(\frac{\partial^3 u}{\partial x^3}\right)_j \tag{3.39}$$

Thus we have derived a second-order backward-difference approximation of a first derivative:

$$\left(\frac{\partial u}{\partial x}\right)_j - \frac{1}{2\Delta x}(u_{j-2} - 4u_{j-1} + 3u_j) = O(\Delta x^2) \tag{3.40}$$

3.4.2 Generalization of Difference Formulas

In general, a difference approximation to the mth derivative at grid point j can be cast in terms of $q + p + 1$ neighboring points as

$$\left(\frac{\partial^m u}{\partial x^m}\right)_j - \sum_{i=-p}^{q} a_i u_{j+i} = er_t \tag{3.41}$$

where the a_i are coefficients to be determined through the use of Taylor tables to produce approximations of a given order. Clearly this process can be used to find forward, backward, skewed, or central point operators of any order for any derivative. It could be computer automated and extended to higher dimensions. More important, however, is the fact that it can be further generalized. In order to do this, let us approach the subject in a slightly different way, that is from the point of view of interpolation formulas. These formulas are discussed in many texts on numerical analysis.

3.4.3 Lagrange and Hermite Interpolation Polynomials

The Lagrangian interpolation polynomial is given by

$$u(x) = \sum_{k=0}^{K} a_k(x) u_k \tag{3.42}$$

where $a_k(x)$ are polynomials in x of degree K. The construction of the $a_k(x)$ can be taken from the simple Lagrangian formula for quadratic interpolation (or extrapolation) with non-equispaced points

$$\begin{aligned} u(x) = u_0 \frac{(x_1 - x)(x_2 - x)}{(x_1 - x_0)(x_2 - x_0)} &+ u_1 \frac{(x_0 - x)(x_2 - x)}{(x_0 - x_1)(x_2 - x_1)} \\ &+ u_2 \frac{(x_0 - x)(x_1 - x)}{(x_0 - x_2)(x_1 - x_2)} \end{aligned} \tag{3.43}$$

Notice that the coefficient of each u_k is one when $x = x_k$, and zero when x takes any other discrete value in the set. If we take the first or second derivative of $u(x)$, impose an equispaced mesh, and evaluate these derivatives at the appropriate discrete point, we rederive the finite-difference approximations just presented. Finite-difference schemes that can be derived from Eq. 3.42 are referred to as Lagrangian approximations.

A generalization of the Lagrangian approach is brought about by using Hermitian interpolation. To construct a polynomial for $u(x)$, Hermite formulas use values of the function *and its derivative(s)* at given points in space. Our illustration is for the case in which discrete values of the function and its first derivative are used, producing the expression

$$u(x) = \sum a_k(x) u_k + \sum b_k(x) \left(\frac{\partial u}{\partial x} \right)_k \tag{3.44}$$

Obviously higher-order derivatives could be included as the problems dictate. A complete discussion of these polynomials can be found in many references on numerical methods, but here we need only the concept.

The previous examples of a Taylor table constructed *explicit* point difference operators from Lagrangian interpolation formulas. Consider next the Taylor table for an *implicit* space differencing scheme for a first derivative arising from the use of a Hermite interpolation formula. A generalization of Eq. 3.41 can include derivatives at neighboring points, i.e.,

$$\sum_{i=-r}^{s} b_i \left(\frac{\partial^m u}{\partial x^m}\right)_{j+i} + \sum_{i=-p}^{q} a_i u_{j+i} = er_t \tag{3.45}$$

analogous to Eq. 3.44. An example formula is illustrated at the top of Table 3.3. Here not only is the derivative at point j represented, but also included are derivatives at points $j-1$ and $j+1$, which also must be expanded using Taylor series about point j. This requires the following generalization of the Taylor series expansion given in Eq. 3.8:

$$\left(\frac{\partial^m u}{\partial x^m}\right)_{j+k} = \left\{\left[\sum_{n=0}^{\infty} \frac{1}{n!}(k\Delta x)^n \frac{\partial^n}{\partial x^n}\right]\left(\frac{\partial^m u}{\partial x^m}\right)\right\}_j \tag{3.46}$$

The derivative terms now have coefficients (the coefficient on the j point is taken as one to simplify the algebra) which must be determined using the Taylor table approach as outlined below.

$$d\left(\frac{\partial u}{\partial x}\right)_{j-1} + \left(\frac{\partial u}{\partial x}\right)_j + e\left(\frac{\partial u}{\partial x}\right)_{j+1} - \frac{1}{\Delta x}(au_{j-1} + bu_j + cu_{j+1}) = \,?$$

	u_j	$\Delta x \cdot \left(\frac{\partial u}{\partial x}\right)_j$	$\Delta x^2 \cdot \left(\frac{\partial^2 u}{\partial x^2}\right)_j$	$\Delta x^3 \cdot \left(\frac{\partial^3 u}{\partial x^3}\right)_j$	$\Delta x^4 \cdot \left(\frac{\partial^4 u}{\partial x^4}\right)_j$	$\Delta x^5 \cdot \left(\frac{\partial^5 u}{\partial x^5}\right)_j$
$\Delta x \cdot d\left(\frac{\partial u}{\partial x}\right)_{j-1}$		d	$d \cdot (\text{-}1) \cdot \frac{1}{1}$	$d \cdot (\text{-}1)^2 \cdot \frac{1}{2}$	$d \cdot (\text{-}1)^3 \cdot \frac{1}{6}$	$d \cdot (\text{-}1)^4 \cdot \frac{1}{24}$
$\Delta x \cdot \left(\frac{\partial u}{\partial x}\right)_j$		1				
$\Delta x \cdot e\left(\frac{\partial u}{\partial x}\right)_{j+1}$		e	$e \cdot (1) \cdot \frac{1}{1}$	$e \cdot (1)^2 \cdot \frac{1}{2}$	$e \cdot (1)^3 \cdot \frac{1}{6}$	$e \cdot (1)^4 \cdot \frac{1}{24}$
$-a \cdot u_{j-1}$	$-a$	$-a \cdot (\text{-}1) \cdot \frac{1}{1}$	$-a \cdot (\text{-}1)^2 \cdot \frac{1}{2}$	$-a \cdot (\text{-}1)^3 \cdot \frac{1}{6}$	$-a \cdot (\text{-}1)^4 \cdot \frac{1}{24}$	$-a \cdot (\text{-}1)^5 \cdot \frac{1}{120}$
$-b \cdot u_j$	$-b$					
$-c \cdot u_{j+1}$	$-c$	$-c \cdot (1) \cdot \frac{1}{1}$	$-c \cdot (1)^2 \cdot \frac{1}{2}$	$-c \cdot (1)^3 \cdot \frac{1}{6}$	$-c \cdot (1)^4 \cdot \frac{1}{24}$	$-c \cdot (1)^5 \cdot \frac{1}{120}$

Table 3.3. Taylor table for central 3-point Hermitian approximation to a first derivative.

To maximize the order of accuracy, we must satisfy the relation

$$\begin{bmatrix} -1 & -1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 1 & 1 \\ -1 & 0 & -1 & -2 & 2 \\ 1 & 0 & -1 & 3 & 3 \\ -1 & 0 & -1 & -4 & 4 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

having the solution $[a, b, c, d, e] = \frac{1}{4}[-3, 0, 3, 1, 1]$. Under these conditions the sixth column sums to

$$er_t = \frac{\Delta x^4}{120}\left(\frac{\partial^5 u}{\partial x^5}\right)_j \tag{3.47}$$

and the method can be expressed as

$$\left(\frac{\partial u}{\partial x}\right)_{j-1} + 4\left(\frac{\partial u}{\partial x}\right)_j + \left(\frac{\partial u}{\partial x}\right)_{j+1} - \frac{3}{\Delta x}(-u_{j-1} + u_{j+1}) = O(\Delta x^4) \tag{3.48}$$

This is also referred to as a *Padé* formula.

3.4.4 Practical Application of Padé Formulas

It is one thing to construct methods using the Hermitian concept and quite another to implement them in a computer code. In the form of a point operator it is probably not evident at first just how Eq. 3.48 can be applied. However, the situation is quite easy to comprehend if we express the same method in the form of a matrix operator. A banded matrix notation for Eq. 3.48 is

$$\frac{1}{6}B(1,4,1)\delta_x\vec{u} = \frac{1}{2\Delta x}B(-1,0,1)\vec{u} + \left(\vec{bc}\right) \tag{3.49}$$

in which Dirichlet boundary conditions have been imposed.[4] Mathematically this is equivalent to

$$\delta_x\vec{u} = 6[B(1,4,1)]^{-1}\left[\frac{1}{2\Delta x}B(-1,0,1)\vec{u} + \left(\vec{bc}\right)\right] \tag{3.50}$$

which can be reexpressed by the "predictor-corrector" sequence

$$\begin{aligned} \vec{\tilde{u}} &= \frac{1}{2\Delta x}B(-1,0,1)\vec{u} + \left(\vec{bc}\right) \\ \delta_x\vec{u} &= 6\,[B(1,4,1)]^{-1}\vec{\tilde{u}} \end{aligned} \tag{3.51}$$

[4]In this case the vector containing the boundary conditions would include values of both u and $\partial u/\partial x$ at both boundaries.

With respect to practical implementation, the meaning of the predictor in this sequence should be clear. It simply says – take the vector array $\vec{u}$, difference it, add the boundary conditions, and store the result in the intermediate array $\vec{\tilde{u}}$. The meaning of the second row is more subtle, since it is demanding the evaluation of an inverse operator, but it still can be given a simple interpretation. An inverse matrix operator implies the solution of a coupled set of linear equations. These operators are very common in finite difference applications. They appear in the form of banded matrices having a small bandwidth, in this case a tridiagonal. The evaluation of $[B(1,4,1)]^{-1}$ is found by means of a tridiagonal "solver", which is simple to code, efficient to run, and widely used. In general, Hermitian or Padé approximations can be practical when they can be implemented by the use of efficient banded solvers.

3.4.5 Other Higher-Order Schemes

Hermitian forms of the second derivative can also be easily derived by means of a Taylor table. For example

$$\delta_{xx}\vec{u} = 12\,[B(1,10,1)]^{-1}\left[\frac{1}{\Delta x^2}B(1,-2,1)\vec{u} + \left(\vec{bc}\right)\right] \tag{3.52}$$

is $O(\Delta x^4)$ and makes use of only tridiagonal operations. It should be mentioned that the spline approximation is one form of a Padé matrix difference operator. It is given by

$$\delta_{xx}\vec{u} = 6\,[B(1,4,1)]^{-1}\left[\frac{1}{\Delta x^2}B(1,-2,1)\vec{u} + \left(\vec{bc}\right)\right] \tag{3.53}$$

but its order of accuracy is only $O(\Delta x^2)$. How much this reduction in accuracy is offset by the increased global continuity built into a spline fit is not known. We note that the spline fit of a first derivative is identical to any of the expressions in Eqs. 3.48 to 3.51.

A final word on Hermitian approximations. Clearly they have an advantage over 3-point Lagrangian schemes because of their increased accuracy. However, a more subtle point is that they get this increase in accuracy using information that is still local to the point where the derivatives are being evaluated. In application, this can be advantageous at boundaries and in the vicinity of steep gradients. It is obvious, of course, that five point schemes using Lagrangian approximations can be derived that have the same order of accuracy as the methods given in Eqs. 3.48 and 3.52, but they will have a wider spread of space indices. In particular, two Lagrangian schemes with the same order of accuracy are (here we ignore the problem created by the boundary conditions, although this is one of the principal issues in applying these schemes):

$$\frac{\partial\vec{u}}{\partial x} - \frac{1}{12\Delta x}B_p(1,-8,0,8,-1)\vec{u} = O\left(\Delta x^4\right) \tag{3.54}$$

$$\frac{\partial^2 \vec{u}}{\partial x^2} - \frac{1}{12\Delta x^2} B_p(-1, 16, -30, 16, -1)\vec{u} = O\Big(\Delta x^4\Big) \tag{3.55}$$

3.5 Fourier Error Analysis

In order to select a finite-difference scheme for a given application one must be able to assess the accuracy of the candidate schemes. The accuracy of an operator is often expressed in terms of the order of the leading error term determined from a Taylor table. While this is a useful measure, it provides a fairly limited description. Further information about the error behavior of a finite-difference scheme can be obtained using Fourier error analysis.

3.5.1 Application to a Spatial Operator

An arbitrary periodic function can be decomposed into its Fourier components, which are in the form $e^{i\kappa x}$, where κ is the wavenumber. It is therefore of interest to examine how well a given finite-difference operator approximates derivatives of $e^{i\kappa x}$. We will concentrate here on first derivative approximations, although the analysis is equally applicable to higher derivatives.

The exact first derivative of $e^{i\kappa x}$ is

$$\frac{\partial e^{i\kappa x}}{\partial x} = i\kappa e^{i\kappa x} \tag{3.56}$$

If we apply, for example, a second-order centered difference operator to $u_j = e^{i\kappa x_j}$, where $x_j = j\Delta x$, we get

$$\begin{aligned}
(\delta_x u)_j &= \frac{u_{j+1} - u_{j-1}}{2\Delta x} \\
&= \frac{e^{i\kappa\Delta x(j+1)} - e^{i\kappa\Delta x(j-1)}}{2\Delta x} \\
&= \frac{(e^{i\kappa\Delta x} - e^{-i\kappa\Delta x})e^{i\kappa x_j}}{2\Delta x} \\
&= \frac{1}{2\Delta x}[(\cos\kappa\Delta x + i\sin\kappa\Delta x) - (\cos\kappa\Delta x - i\sin\kappa\Delta x)]e^{i\kappa x_j} \\
&= i\frac{\sin\kappa\Delta x}{\Delta x}e^{i\kappa x_j} \\
&= i\kappa^* e^{i\kappa x_j}
\end{aligned} \tag{3.57}$$

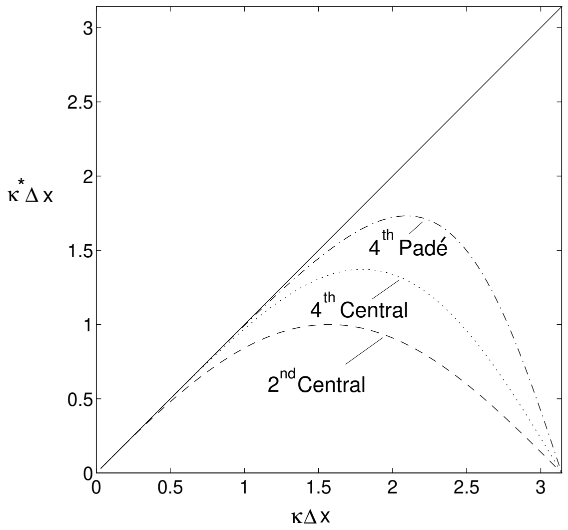

Figure 3.4: Modified wavenumber for various schemes.

where κ^* is the modified wavenumber. The modified wavenumber is so named because it appears where the wavenumber, κ, appears in the exact expression. Thus the degree to which the modified wavenumber approximates the actual wavenumber is a measure of the accuracy of the approximation.

For the second-order centered difference operator the modified wavenumber is given by

$$\kappa^* = \frac{\sin \kappa \Delta x}{\Delta x} \tag{3.58}$$

Note that κ^* approximates κ to second-order accuracy, as is to be expected, since

$$\frac{\sin \kappa \Delta x}{\Delta x} = \kappa - \frac{\kappa^3 \Delta x^2}{6} + \ldots$$

Equation 3.58 is plotted in Figure 3.4, along with similar relations for the standard fourth-order centered difference scheme and the fourth-order Padé scheme. The expression for the modified wavenumber provides the accuracy with which a given wavenumber component of the solution is resolved for the entire wavenumber range available in a mesh of a given size, $0 \leq \kappa \Delta x \leq \pi$.

In general, finite-difference operators can be written in the form

$$(\delta_x)_j = (\delta_x^a)_j + (\delta_x^s)_j$$

where $(\delta_x^a)_j$ is an antisymmetric operator and $(\delta_x^s)_j$ is a symmetric operator.[5] If we restrict our interest to schemes extending from $j-3$ to $j+3$, then

$$(\delta_x^a u)_j = \frac{1}{\Delta x}[a_1(u_{j+1} - u_{j-1}) + a_2(u_{j+2} - u_{j-2}) + a_3(u_{j+3} - u_{j-3})]$$

and

$$(\delta_x^s u)_j = \frac{1}{\Delta x}[d_0 u_j + d_1(u_{j+1} + u_{j-1}) + d_2(u_{j+2} + u_{j-2}) + d_3(u_{j+3} + u_{j-3})]$$

The corresponding modified wavenumber is

$$\begin{aligned} i\kappa^* &= \frac{1}{\Delta x}[d_0 + 2(d_1 \cos\kappa\Delta x + d_2 \cos 2\kappa\Delta x + d_3 \cos 3\kappa\Delta x) \\ &+ 2i(a_1 \sin\kappa\Delta x + a_2 \sin 2\kappa\Delta x + a_3 \sin 3\kappa\Delta x) \end{aligned} \tag{3.59}$$

When the finite-difference operator is antisymmetric (centered), the modified wavenumber is purely real. When the operator includes a symmetric component, the modified wavenumber is complex, with the imaginary component being entirely error. The fourth-order Padé scheme is given by

$$(\delta_x u)_{j-1} + 4(\delta_x u)_j + (\delta_x u)_{j+1} = \frac{3}{\Delta x}(u_{j+1} - u_{j-1})$$

The modified wavenumber for this scheme satisfies[6]

$$i\kappa^* e^{-i\kappa\Delta x} + 4i\kappa^* + i\kappa^* e^{i\kappa\Delta x} = \frac{3}{\Delta x}(e^{i\kappa\Delta x} - e^{-i\kappa\Delta x})$$

which gives

$$i\kappa^* = \frac{3i\sin\kappa\Delta x}{(2+\cos\kappa\Delta x)\Delta x}$$

The modified wavenumber provides a useful tool for assessing difference approximations. In the context of the linear convection equation, the errors can be given a physical interpretation. Consider once again the linear convection equation in the form

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

[5]In terms of a circulant matrix operator A, the antisymmetric part is obtained from $(A - A^T)/2$ and the symmetric part from $(A + A^T)/2$.

[6]Note that terms such as $(\delta_x u)_{j-1}$ are handled by letting $(\delta_x u)_j = ik^* e^{i\kappa j\Delta x}$ and evaluating the shift in j.

on a domain extending from $-\infty$ to ∞. Recall from Section 2.3.2 that a solution initiated by a harmonic function with wavenumber κ is

$$u(x,t) = f(t)e^{i\kappa x} \tag{3.60}$$

where $f(t)$ satisfies the ODE

$$\frac{df}{dt} = -ia\kappa f$$

Solving for $f(t)$ and substituting into Eq. 3.60 gives the exact solution as

$$u(x,t) = f(0)e^{i\kappa(x-at)}$$

If second-order centered differences are applied to the spatial term, the following ODE is obtained for $f(t)$:

$$\frac{df}{dt} = -ia\left[\frac{\sin\kappa\Delta x}{\Delta x}\right]f = -ia\kappa^* f \tag{3.61}$$

Solving this ODE exactly (since we are considering the error from the spatial approximation only) and substituting into Eq. 3.60, we obtain

$$u_{\text{numerical}}(x,t) = f(0)e^{i\kappa(x-a^*t)} \tag{3.62}$$

where a^* is the numerical (or modified) phase speed, which is related to the modified wavenumber by

$$\frac{a^*}{a} = \frac{\kappa^*}{\kappa}$$

For the above example,

$$\frac{a^*}{a} = \frac{\sin\kappa\Delta x}{\kappa\Delta x}$$

The numerical phase speed is the speed at which a harmonic function is propagated numerically. Since $a^*/a \leq 1$ for this example, the numerical solution propagates too slowly. Since a^* is a function of the wavenumber, the numerical approximation introduces dispersion, although the original PDE is nondispersive. As a result, a waveform consisting of many different wavenumber components eventually loses its original form.

Figure 3.5 shows the numerical phase speed for the schemes considered previously. The number of *points per wavelength* (PPW) by which a given wave is resolved is given by $2\pi/\kappa\Delta x$. The resolving efficiency of a scheme can be expressed in terms of the PPW required to produce errors below a specified level. For example, the second-order centered difference scheme requires 80 PPW to produce an error in phase speed of less than 0.1 percent. The 5-point fourth-order centered scheme and

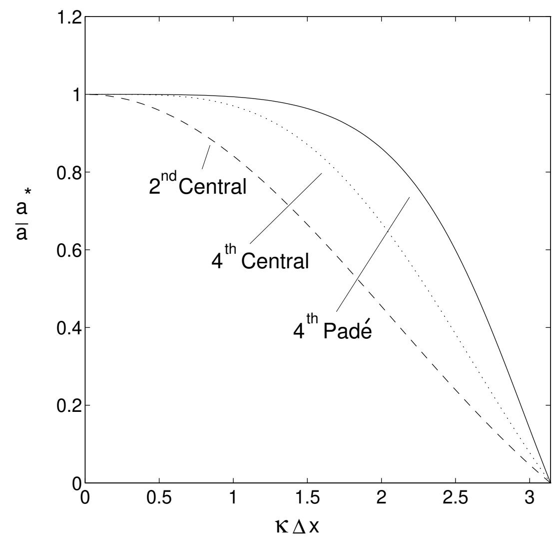

Figure 3.5: Numerical phase speed for various schemes.

the fourth-order Padé scheme require 15 and 10 PPW respectively to achieve the same error level.

For our example using second-order centered differences, the modified wavenumber is purely real, but in the general case it can include an imaginary component as well, as shown in Eq. 3.59. In that case, the error in the phase speed is determined from the real part of the modified wavenumber, while the imaginary part leads to an error in the amplitude of the solution, as can be seen by inspecting Eq. 3.61. Thus the antisymmetric portion of the spatial difference operator determines the error in speed and the symmetric portion the error in amplitude. This will be further discussed in Section 11.2.

3.6 Difference Operators at Boundaries

As discussed in Section 3.3.2, a matrix difference operator incorporates both the difference approximation in the interior of the domain and that at the boundaries. In this section, we consider the boundary operators needed for our model equations for convection and diffusion. In the case of periodic boundary conditions, no special boundary operators are required.

3.6.1 The Linear Convection Equation

Referring to Section 2.3, the boundary conditions for the linear convection equation can be either periodic or of inflow-outflow type. In the latter case, a Dirichlet boundary condition is given at the inflow boundary, while no condition is specified at the outflow boundary. Here we assume that the wave speed a is positive; thus the left-hand boundary is the inflow boundary, and the right-hand boundary is the outflow boundary. Thus the vector of unknowns is

$$\vec{u} = [u_1,\ u_2,\ \ldots,\ u_M]^T \tag{3.63}$$

and u_0 is specified.

Consider first the inflow boundary. It is clear that as long as the interior difference approximation does not extend beyond u_{j-1}, then no special treatment is required for this boundary. For example, with second-order centered differences we obtain

$$\delta_x \vec{u} = A\vec{u} + \left(\vec{bc}\right) \tag{3.64}$$

with

$$A = \frac{1}{2\Delta x}\begin{bmatrix} 0 & 1 & & & \\ -1 & 0 & 1 & & \\ & -1 & 0 & 1 & \\ & & & \ddots & \end{bmatrix}, \qquad \left(\vec{bc}\right) = \frac{1}{2\Delta x}\begin{bmatrix} -u_0 \\ 0 \\ 0 \\ \vdots \end{bmatrix} \tag{3.65}$$

However, if we use the fourth-order interior operator given in Eq. 3.54, then the approximation at $j = 1$ requires a value of u_{j-2}, which is outside the domain. Hence, a different operator is required at $j = 1$ which extends only to $j - 1$, while having the appropriate order of accuracy. Such an operator, known as a *numerical boundary scheme*, can have an order of accuracy which is *one order lower* than that of the interior scheme, and the global accuracy will equal that of the interior scheme.[7] For example, with fourth-order centered differences, we can use the following third-order operator at $j = 1$:

$$(\delta_x u)_1 = \frac{1}{6\Delta x}(-2u_0 - 3u_1 + 6u_2 - u_3) \tag{3.66}$$

which is easily derived using a Taylor table. The resulting difference operator has the form of Eq. 3.64 with

$$A = \frac{1}{12\Delta x}\begin{bmatrix} -6 & 12 & -2 & & & \\ -8 & 0 & 8 & -1 & & \\ 1 & -8 & 0 & 8 & -1 & \\ & & & \ddots & & \end{bmatrix}, \qquad \left(\vec{bc}\right) = \frac{1}{12\Delta x}\begin{bmatrix} -4u_0 \\ u_0 \\ 0 \\ \vdots \end{bmatrix} \tag{3.67}$$

[7] Proof of this theorem is beyond the scope of this book; the interested reader should consult the literature for further details.

This approximation is globally fourth-order accurate.

At the outflow boundary, no boundary condition is specified. We must approximate $\partial u/\partial x$ at node M with no information about u_{M+1}. Thus the second-order centered-difference operator, which requires u_{j+1}, cannot be used at $j = M$. A backward-difference formula must be used. With a second-order interior operator, the following first-order backward formula can be used:

$$(\delta_x u)_M = \frac{1}{\Delta x}(u_M - u_{M-1}) \tag{3.68}$$

This produces a difference operator with

$$A = \frac{1}{2\Delta x}\begin{bmatrix} 0 & 1 & & & & \\ -1 & 0 & 1 & & & \\ & -1 & 0 & 1 & & \\ & & & \ddots & & \\ & & & -1 & 0 & 1 \\ & & & & -2 & 2 \end{bmatrix}, \quad \left(\vec{bc}\right) = \frac{1}{2\Delta x}\begin{bmatrix} -u_0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{3.69}$$

In the case of a fourth-order centered interior operator the last two rows of A require modification.

Another approach to the development of boundary schemes is in terms of space extrapolation. The following formula allows u_{M+1} to be extrapolated from the interior data to arbitrary order on an equispaced grid:

$$(1 - E^{-1})^p u_{M+1} = 0 \tag{3.70}$$

where E is the shift operator defined by $Eu_j = u_{j+1}$ and the order of the approximation is $p - 1$. For example, with $p = 2$ we obtain

$$(1 - 2E^{-1} + E^{-2})u_{M+1} = u_{M+1} - 2u_M + u_{M-1} = 0 \tag{3.71}$$

which gives the following first-order approximation to u_{M+1}:

$$u_{M+1} = 2u_M - u_{M-1} \tag{3.72}$$

Substituting this into the second-order centered-difference operator applied at node M gives

$$\begin{aligned}(\delta_x u)_M = \frac{1}{2\Delta x}(u_{M+1} - u_{M-1}) &= \frac{1}{2\Delta x}(2u_M - u_{M-1} - u_{M-1}) \\ &= \frac{1}{\Delta x}(u_M - u_{M-1})\end{aligned} \tag{3.73}$$

which is identical to Eq. 3.68.

3.6.2 The Diffusion Equation

In solving the diffusion equation, we must consider Dirichlet and Neumann boundary conditions. The treatment of a Dirichlet boundary condition proceeds along the same lines as the inflow boundary for the convection equation discussed above. With the second-order centered interior operator

$$(\delta_{xx}u)_j = \frac{1}{\Delta x^2}(u_{j+1} - 2u_j + u_{j-1}) \tag{3.74}$$

no modifications are required near boundaries, leading to the difference operator given in Eq. 3.27.

For a Neumann boundary condition, we assume that $\partial u/\partial x$ is specified at $j = M+1$, that is

$$\left(\frac{\partial u}{\partial x}\right)_{M+1} = \left(\frac{\partial u}{\partial x}\right)_b \tag{3.75}$$

Thus we design an operator at node M which is in the following form:

$$(\delta_{xx}u)_M = \frac{1}{\Delta x^2}(au_{M-1} + bu_M) + \frac{c}{\Delta x}\left(\frac{\partial u}{\partial x}\right)_{M+1} \tag{3.76}$$

where a, b, and c are constants which can easily be determined using a Taylor table, as shown in Table 3.4.

$$\left(\frac{\partial^2 u}{\partial x^2}\right)_j - \left[\frac{1}{\Delta x^2}(au_{j-1} + bu_j) + \frac{c}{\Delta x}\left(\frac{\partial u}{\partial x}\right)_{j+1}\right] = ?$$

	u_j	$\Delta x \cdot \left(\frac{\partial u}{\partial x}\right)_j$	$\Delta x^2 \cdot \left(\frac{\partial^2 u}{\partial x^2}\right)_j$	$\Delta x^3 \cdot \left(\frac{\partial^3 u}{\partial x^3}\right)_j$	$\Delta x^4 \cdot \left(\frac{\partial^4 u}{\partial x^4}\right)_j$
$\Delta x^2 \cdot \left(\frac{\partial^2 u}{\partial x^2}\right)_j$			1		
$-a \cdot u_{j-1}$	$-a$	$-a \cdot (\text{-}1) \cdot \frac{1}{1}$	$-a \cdot (\text{-}1)^2 \cdot \frac{1}{2}$	$-a \cdot (\text{-}1)^3 \cdot \frac{1}{6}$	$-a \cdot (\text{-}1)^4 \cdot \frac{1}{24}$
$-b \cdot u_j$	$-b$				
$-\Delta x \cdot c \cdot \left(\frac{\partial u}{\partial x}\right)_{j+1}$		$-c$	$-c \cdot (1) \cdot \frac{1}{1}$	$-c \cdot (1)^2 \cdot \frac{1}{2}$	$-c \cdot (1)^3 \cdot \frac{1}{6}$

Table 3.4. Taylor table for Neumann boundary condition.

Solving for a, b, and c, we obtain the following operator:

$$(\delta_{xx}u)_M = \frac{1}{3\Delta x^2}(2u_{M-1} - 2u_M) + \frac{2}{3\Delta x}\left(\frac{\partial u}{\partial x}\right)_{M+1} \tag{3.77}$$

which produces the difference operator given in Eq. 3.29. Notice that this operator is second-order accurate. In the case of a numerical approximation to a Neumann boundary condition, this is necessary to obtain a globally second-order accurate formulation. This contrasts with the numerical boundary schemes described previously which can be one order lower than the interior scheme.

We can also obtain the operator in Eq. 3.77 using the space extrapolation idea. Consider a second-order backward-difference approximation applied at node $M+1$:

$$\left(\frac{\partial u}{\partial x}\right)_{M+1} = \frac{1}{2\Delta x}(u_{M-1} - 4u_M + 3u_{M+1}) + O(\Delta x^2) \tag{3.78}$$

Solving for u_{M+1} gives

$$u_{M+1} = \frac{1}{3}\left[4u_M - u_{M-1} + 2\Delta x \left(\frac{\partial u}{\partial x}\right)_{M+1}\right] + O(\Delta x^3) \tag{3.79}$$

Substituting this into the second-order centered difference operator for a second derivative applied at node M gives

$$\begin{aligned}(\delta_{xx}u)_M &= \frac{1}{\Delta x^2}(u_{M+1} - 2u_M + u_{M-1}) &(3.80)\\ &= \frac{1}{3\Delta x^2}\left[3u_{M-1} - 6u_M + 4u_M - u_{M-1} + 2\Delta x\left(\frac{\partial u}{\partial x}\right)_{M+1}\right] \\ &= \frac{1}{3\Delta x^2}(2u_{M-1} - 2u_M) + \frac{2}{3\Delta x}\left(\frac{\partial u}{\partial x}\right)_{M+1} &(3.81)\end{aligned}$$

which is identical to Eq. 3.77.

3.7 Problems

1. Derive a third-order finite-difference approximation to a first derivative in the form

$$(\delta_x u)_j = \frac{1}{\Delta x}(au_{j-2} + bu_{j-1} + cu_j + du_{j+1})$$

   Find the leading error term.

2. Derive a finite-difference approximation to a first derivative in the form

$$a(\delta_x u)_{j-1} + (\delta_x u)_j = \frac{1}{\Delta x}(bu_{j-1} + cu_j + du_{j+1})$$

   Find the leading error term.

3. Using a 4 (interior) point mesh, write out the 4×4 matrices and the boundary-condition vector formed by using the scheme derived in question 2 when both u and $\partial u/\partial x$ are given at $j = 0$ and u is given at $j = 5$.

4. Repeat question 2 with $d = 0$.

5. Derive a finite-difference approximation to a third derivative in the form

$$(\delta_{xxx}u)_j = \frac{1}{\Delta x^3}(au_{j-2} + bu_{j-1} + cu_j + du_{j+1} + eu_{j+2})$$

   Find the leading error term.

6. Derive a compact (or Padé) finite-difference approximation to a second derivative in the form

$$d(\delta_{xx}u)_{j-1} + (\delta_{xx}u)_j + e(\delta_{xx}u)_{j+1} = \frac{1}{\Delta x^2}(au_{j-1} + bu_j + cu_{j+1})$$

   Find the leading error term.

7. Find the modified wavenumber for the operator derived in question 1. Plot the real and imaginary parts of $\kappa^*\Delta x$ vs. $\kappa\Delta x$ for $0 \leq \kappa\Delta x \leq \pi$. Compare the real part with that obtained from the fourth-order centered operator (Eq. 3.54).

8. Application of the second-derivative operator to the function $e^{i\kappa x}$ gives

$$\frac{\partial^2 e^{i\kappa x}}{\partial x^2} = -\kappa^2 e^{i\kappa x}$$

   Application of a difference operator for the second derivative gives

$$(\delta_{xx}e^{i\kappa j\Delta x})_j = -\kappa^{*2} e^{i\kappa x}$$

   thus defining the modified wavenumber κ^* for a second derivative approximation. Find the modified wavenumber for the second-order centered difference operator for a second derivative, the noncompact fourth-order operator (Eq. 3.55), and the compact fourth-order operator derived in question 6. Plot $(\kappa^*\Delta x)^2$ vs. $(\kappa\Delta x)^2$ for $0 \leq \kappa\Delta x \leq \pi$.

9. Find the grid-points-per-wavelength (PPW) requirement to achieve a phase speed error less than 0.1 percent for sixth-order noncompact and compact centered approximations to a first derivative.

10. Consider the following one-sided differencing schemes, which are first-, second-, and third-order, respectively:

$$(\delta_x u)_j = (u_j - u_{j-1})\,/\Delta x$$

$$(\delta_x u)_j = (3u_j - 4u_{j-1} + u_{j-2})\,/(2\Delta x)$$

$$(\delta_x u)_j = (11u_j - 18u_{j-1} + 9u_{j-2} - 2u_{j-3})\,/(6\Delta x)$$

Find the modified wavenumber for each of these schemes. Plot the real and imaginary parts of $\kappa^* \Delta x$ vs. $\kappa \Delta x$ for $0 \leq \kappa \Delta x \leq \pi$. Derive the two leading terms in the truncation error for each scheme.