The Numerical Solution of Second-Order Boundary Value Problems on Nonuniform Meshes

By Thomas A. Manteuffel and Andrew B. White, Jr.

Abstract. In this paper, we examine the solution of second-order, scalar boundary value problems on nonuniform meshes. We show that certain commonly used difference schemes yield second-order accurate solutions despite the fact that their truncation error is of lower order. This result illuminates a limitation of the standard stability, consistency proof of convergence for difference schemes defined on nonuniform meshes. A technique of reducing centered-difference approximations of first-order systems to discretizations of the underlying scalar equation is developed. We treat both vertex-centered and cell-centered difference schemes and indicate how these results apply to partial differential equations on Cartesian product grids.

1. Introduction. Much attention has been paid to the numerical solution of second-order differential equations on nonuniform meshes. To begin, we consider the solution of the linear, two-point boundary value problem

$$(1.1) \qquad y'' + a(x)y' + b(x)y = f(x), \qquad x \in (0,1),$$

$$(1.2a) \qquad b_{00}y(0) + b_{01}y'(0) = b_{02},$$

$$(1.2b) \qquad b_{10}y(1) + b_{11}y'(1) = b_{12},$$

through three-point (compact-as-possible in the sense of Kreiss [15]) difference schemes on a mesh $\{x_i\}_{i=0}^{i=N}$. The functions a, b, and f are assumed to be smooth. The standard mesh spacing will be written as $\Delta_i = x_i - x_{i-1}$, and we will denote a function evaluated at the mesh points by subscripts, y_i, and the vector with these entries as Y. The ith component of Y may also be written $(Y)_i$. We will make liberal use of the generic constant C and of $\Delta = \max\{\Delta_i\}$.

An example of such a difference scheme is that used by Pearson [21] in his classic work on singular perturbations and by Denny and Landis [5] and de Rivas [6] in their work on mesh-selection techniques. In this method, each term in (1.1) is approximated separately on the stencil $(x - \Delta_i, x, x + \Delta_{i+1})$. Approximations to the derivatives are

$$(1.3a) \qquad \left[\frac{2}{\Delta_i(\Delta_i + \Delta_{i+1})}, \frac{-2}{\Delta_i\Delta_{i+1}}, \frac{2}{(\Delta_i + \Delta_{i+1})\Delta_{i+1}} \right] \begin{bmatrix} y_{i-1} \\ y_i \\ y_{i+1} \end{bmatrix}$$

$$= y'' + \tfrac{1}{3}(\Delta_{i+1} - \Delta_i)y''' + O(\Delta^2),$$

$$(1.3b) \quad \left[\frac{-\Delta_{i+1}}{\Delta_i(\Delta_i + \Delta_{i+1})}, \frac{\Delta_{i+1} - \Delta_i}{\Delta_i\Delta_{i+1}}, \frac{\Delta_i}{\Delta_{i+1}(\Delta_i + \Delta_{i+1})} \right] \begin{bmatrix} y_{i-1} \\ y_i \\ y_{i+1} \end{bmatrix}$$

$$= y' + \tfrac{1}{6}\Delta_i\Delta_{i+1}y''' + O(\Delta^3),$$

where each term on the right-hand side of these equations is evaluated at x. Throughout, the "big O" notation is shorthand for terms bounded by $C\Delta_{max}^p$.

This difference scheme, applied to (1.1), can be written as

$$(1.4) \quad (L_hV)_i \equiv \frac{2 - a_i\Delta_{i+1}}{\Delta_i(\Delta_i + \Delta_{i+1})}(V)_{i-1} + \left[\frac{-2 + a_i(\Delta_{i+1} - \Delta_i)}{\Delta_i\Delta_{i+1}} + b_i \right](V)_i$$

$$+ \frac{2 + a_i\Delta_i}{\Delta_{i+1}(\Delta_i + \Delta_{i+1})}(V)_{i+1} = f_i$$

for $i = 1, \ldots, N - 1$. Note that this is a familiar, second-order accurate difference scheme on a uniform mesh (cf. Isaacson and Keller [11]). As usual, the truncation error is defined by replacing V in (1.4) with the exact solution Y,

$$(1.5) \qquad\qquad (L_hY)_i = f_i + (T)_i.$$

In this particular case, expanding in a Taylor series about x_i yields

$$(1.6) \qquad\qquad (T)_i = \tfrac{1}{3}(\Delta_{i+1} - \Delta_i)y_i''' + O(\Delta^2)$$

for $i = 1, \ldots, N - 1$. Notice that the truncation error is not second-order unless the mesh is almost uniform.

The standard proof of convergence uses stability and consistency to imply convergence. In other words, given stability, a sufficient condition for second-order accuracy is that the truncation error be second-order. This is not a necessary condition, but it is a convenient and powerful tool. Many attempts have been made to derive schemes that are second-order accurate, i.e., yield second-order convergence in the maximum mesh size, on nonuniform meshes by forcing the truncation error to be second-order accurate. Consider, for example, the use of smoothly varying mesh functions. The notion of slowly varying meshes is natural in solving problems of this sort (cf. Chong [2], Hoffman [10]). Here the idea is to restrict the change in the mesh size so that

$$(\Delta_{i+1} - \Delta_i) = O(\Delta^2).$$

Clearly, then, both terms in the truncation error (1.6) will be second-order. A closely related scheme assumes that there exists a smooth mesh function, say $g(s) \in C^2(0,1)$, such that

$$(1.7a) \qquad\qquad x_i = g(s_i), \quad i = 0, \ldots, N,$$
$$(1.7b) \qquad\qquad s_i = i/N, \quad i = 0, \ldots, N.$$

Such transformations have been discussed in various contexts by Ablow and Schechter [1], de Rivas [6], Davis and Flaherty [4], Hoffman [10], and White [24], among others. In our context, we need only note that

$$(\Delta_i)^2 = \left[\frac{x_i - x_{i-1}}{\Delta s} \right]^2 (\Delta s)^2 \approx [g'(s_i)]^2(\Delta s)^2,$$

and

$$\Delta_{i+1} - \Delta_i = \left[\frac{x_{i+1} - 2x_i + x_{i-1}}{\Delta s^2}\right](\Delta s)^2 \approx \left[g''(s_i)\right](\Delta s)^2.$$

Thus, the truncation error in (1.6) is now second order in Δs.

Another method of producing second-order truncation error is the use of implicit-difference schemes for both ordinary and partial differential equations. Such methods have been studied by Osborne [20], Lynch and Rice [18], Doedel [7], Swartz [22], Ciment, Leventhal, and Weinberg [3], among others. We will illustrate these methods with a simple example that is due to Doedel. Instead of a difference approximation like (1.4), an extra degree of freedom is added by looking at equations of the form

$$(1.8) \qquad \alpha_i v_{i-1} + \beta_i v_i + \gamma_i v_{i+1} = f(z_i),$$

where α_i, β_i, γ_i, and z_i are to be determined. Although the precise formulae are tedious to display, we note that if α_i, β_i, and γ_i correspond to the second divided difference and $z_i = (x_{i-1} + x_i + x_{i+1})/3$, then (1.8) approximates $y'' = f(x)$ with truncation error that is $O(\Delta^2)$ for uniform and nonuniform meshes.

We show in this paper that, for many common difference schemes, the accuracy is second-order in spite of first-order truncation error. The standard proof, although useful in its proper context, is inadequate to handle nonuniform meshes. We will develop new tools that are applicable to both uniform and nonuniform meshes. As a motivation, consider writing (1.1) as a pair of first-order equations,

$$(1.9a) \qquad\qquad\qquad y_1' = y_2,$$

$$(1.9b) \qquad\qquad y_2' = -a(x)y_2 - b(x)y_1 + f(x),$$

with appropriate boundary conditions. Keller [12] and Keller and White [13] have shown that centered-difference approximations of (1.9a,b) are second-order accurate under mild restrictions on the mesh. This approach has the additional advantages that its asymptotic error expansion proceeds in powers of $(\Delta_i)^2$ and that approximations to the derivative are just as accurate as those to the solution itself.

In Section 2, we begin our study by looking more closely at the solution of (1.9a,b) by centered differences. In the resulting linear system of equations, the unknowns Y_2 can be eliminated in such a way as to yield a compact-as-possible difference scheme for the unknowns Y_1. Like (1.6), the truncation error is first-order. However, we know that the accuracy is second-order because it is also the solution of the system (1.9a,b). The reduction procedure is examined to illuminate this apparent paradox. A simple result is proved that provides the basis for the remainder of the paper.

In Section 3, we prove that many difference approximations of (1.1) and (1.2a,b) yield second-order accurate solutions, even though their truncation errors are formally first-order on general nonuniform meshes. As an example, we look at (1.4) and show that no new requirements on the mesh, X, are necessary to ensure second-order accuracy. Several numerical examples will be given to illustrate this result. In Section 4, these results are extended to quasilinear problems. Section 5 examines these problems for cell-centered difference schemes, and we will discover

that these schemes are not nearly as forgiving on nonuniform meshes as are vertex-centered methods. Section 6 contains some concluding remarks.

This work, although approached from a completely different direction, is, in part, a rediscovery of the results of Tikhonov and Samarskiǐ [23], whose paper seems to be little appreciated. We are indebted to David Levermore [17] for calling it to our attention. A related approach, generalized to compact schemes for nonselfadjoint or higher-order equations, invokes summation by parts—see, e.g., Kreiss, Manteuffel, Swartz, Wendroff, and White [16]. The work of Grigorieff [9] addresses similar issues. A Spijker norm is used and in this weaker norm, the truncation error is second-order. The difficulty in this approach comes in proving stability. The proofs presented in our paper are elementary and are motivated through the section on first-order systems. The techniques employed here have application to a wide variety of problems not previously addressed.

2. First-Order Systems of Equations. As we noted in Section 1, the differential equation (1.1) can be rewritten as a pair of first-order differential equations and approximated by centered differences. This system will have second-order truncation error, and thus second-order accuracy, on a nonuniform mesh. In order to examine this procedure in detail, it will be sufficient to consider the simple problem

$$(2.1) \qquad\qquad y'' = f(x).$$

One equivalent first-order system is given by the pair of equations

$$(2.2a) \qquad\qquad y_1'(x) = y_2(x),$$

$$(2.2b) \qquad\qquad y_2'(x) = f(x),$$

where $y_1(x) = y(x)$ and $y_2(x) = y'(x)$. Keller [12] has shown that it is easy to construct a second-order accurate solution of (2.2a,b) using centered-difference approximations,

$$(2.3a) \qquad \frac{1}{\Delta_i}[v_i - v_{i-1}] = \frac{1}{2}[w_i + w_{i-1}], \qquad i = 1, \ldots, N;$$

$$(2.3b) \qquad \frac{1}{\Delta_i}[w_i - w_{i-1}] = f(x_{i-1/2}) \equiv (F)_i, \qquad i = 1, \ldots, N.$$

We may rewrite these equations by defining an ($N \times N + 1$)-difference matrix

$$(2.4a) \qquad D_0 \equiv \begin{bmatrix} -\dfrac{1}{\Delta_1} & \dfrac{1}{\Delta_1} & & \\ & -\dfrac{1}{\Delta_2} & \dfrac{1}{\Delta_2} & \\ & & & \ddots \end{bmatrix},$$

and an ($N \times N + 1$)-average matrix

$$(2.4b) \qquad A_0 \equiv \begin{bmatrix} \dfrac{1}{2} & \dfrac{1}{2} & & \\ & \dfrac{1}{2} & \dfrac{1}{2} & \\ & & & \ddots \end{bmatrix}.$$

With these definitions, Eqs. (2.3a,b) become

$$\text{(2.5a)} \qquad D_0 V = A_0 W,$$

$$\text{(2.5b)} \qquad D_0 W = F.$$

For simplicity, consider Dirichlet boundary conditions appended to (2.1); the discrete approximations are given by

$$\text{(2.6)} \qquad v_0 = b_{02}, \qquad v_N = b_{12}.$$

Equations (2.5a,b) and (2.6) form a system of $(2N + 2)$ equations involving the $(2N + 2)$ unknowns, V and W.

Keller [12] proved the following theorem specialized to our example.

THEOREM 2.1. *Let* $f(x) \in C^2(a, b)$; *then for all quasi-uniform meshes,*

$$\text{(2.7)} \qquad \Delta_{max} \leqslant \lambda \Delta_{min},$$

with $\Delta_{max} \leqslant \Delta_0$, Δ_0 *sufficiently small, a unique solution of the discrete boundary value problem* (2.5a,b) *and* (2.6) *exists and has the property that*

$$\text{(2.8)} \qquad |y_1(x_i) - v_i| \leqslant C\Delta_{max}^2.$$

Proof. The proof is through the usual stability, consistency arguments. The truncation error for (2.3a,b) is found in the usual way by replacing v_i and w_i in (2.3a,b) with $y_1(x_i)$ and $y_2(x_i)$, respectively. In the form of (2.5a,b), we have

$$\text{(2.9a)} \qquad D_0 Y_1 = A_0 Y_2 + T_1,$$

$$\text{(2.9b)} \qquad D_0 Y_2 = F + T_2,$$

where a liberal dose of Taylor's series yields

$$\text{(2.10a)} \qquad (T_1)_i = -\tfrac{1}{12}\Delta_i^2 y_1'''(x_{i-1/2}) + O(\Delta^3),$$

$$\text{(2.10b)} \qquad (T_2)_i = -\tfrac{1}{24}\Delta_i^2 y_1''''(x_{i-1/2}) + O(\Delta^3).$$

Stability will now yield Theorem 2.1. □

However, we can reduce the $(2N + 2)$ equations (2.5a,b) and (2.6) to $(N + 1)$ equations by eliminating W, thus deriving a difference approximation to the original second-order equation, (2.1). If the reduction is done carefully, the new difference scheme will be a three-point scheme and the boundary conditions will remain unaffected. To accomplish this, we define another difference matrix,

$$\text{(2.11a)} \qquad D_1 \equiv \begin{bmatrix} -\dfrac{2}{\Delta_1 + \Delta_2} & \dfrac{2}{\Delta_1 + \Delta_2} & & \\ & -\dfrac{2}{\Delta_2 + \Delta_3} & \dfrac{2}{\Delta_2 + \Delta_3} & \\ & & \ddots & \end{bmatrix},$$

and another average matrix,

$$\text{(2.11b)} \qquad A_1 \equiv \begin{bmatrix} \dfrac{\Delta_1}{\Delta_1 + \Delta_2} & \dfrac{\Delta_2}{\Delta_1 + \Delta_2} & & \\ & \dfrac{\Delta_2}{\Delta_2 + \Delta_3} & \dfrac{\Delta_3}{\Delta_2 + \Delta_3} & \\ & & \ddots & \end{bmatrix},$$

where both matrices are $(N - 1) \times N$. Recalling (2.4a,b), it can be shown that

$$(2.12) \qquad\qquad A_1 D_0 = D_1 A_0.$$

If we left-multiply (2.5a) by D_1, left-multiply (2.5b) by A_1, envoke (2.12), and add the equations, we have

$$(2.13) \qquad\qquad D_1 D_0 V = A_1 F,$$

where $D_1 D_0 \equiv D_2$ is the second divided-difference operator as defined in (1.3a).

The truncation error associated with (2.13) can be derived in two equivalent ways. Replacing V in (2.13) by the exact solution, Y_1, in the usual fashion, we get

$$(2.14) \qquad\qquad D_1 D_0 Y_1 = A_1 F + T,$$

where a Taylor's series expansion yields

$$(2.15) \qquad\qquad (T)_i = -\tfrac{1}{6}(\Delta_{i+1} - \Delta_i) y'''(x_i) + O(\Delta^2).$$

However, we may also derive the truncation error by repeating the reduction shown in Eqs. (2.12) and (2.13), and eliminating Y_2 from (2.9a,b). This procedure yields

$$(2.16) \qquad\qquad D_1 D_0 Y_1 = A_1 F + D_1 T_1 + A_1 T_2.$$

Thus, the truncation error, T, in (2.14) has the representation

$$T = D_1 T_1 + A_1 T_2.$$

Recalling the expressions (2.10a,b), we see that $A_1 T_2$ is $O(\Delta^2)$ because A_1 is an averaging matrix. The term $D_1 T_1$ provides the troublesome $(\Delta_{i+1} - \Delta_i)$-term in (2.15).

The error equation is found by subtracting (2.13) from (2.16), which yields

$$(2.17) \qquad\qquad D_1 D_0 E = D_1 T_1 + A_1 T_2,$$

where $(E)_i \equiv y(x_i) - v_i$. Assuming stability of the difference scheme and ignoring boundary conditions, Eq. (2.17) yields an error estimate

$$\| E \|_\infty \leqslant C \| D_1 T_1 \|_\infty + C \| A_1 T_2 \|_\infty.$$

Recalling (2.10a,b), we have

$$(2.18a) \qquad\qquad \| D_1 T_1 \|_\infty \leqslant C \max_i |\Delta_{i+1} - \Delta_i| = O(\Delta),$$

$$(2.18b) \qquad\qquad \| A_1 T_2 \|_\infty \leqslant C \max_i |\Delta_i^2| = O(\Delta^2).$$

From (2.17) we see that the error can be divided into two terms, $E = E_1 + E_2$, such that

$$(2.19a) \qquad\qquad D_1 D_0 E_1 = D_1 T_1,$$

$$(2.19b) \qquad\qquad D_1 D_0 E_2 = A_1 T_2.$$

The usual arguments applied to (2.19b) yield second-order accuracy. However, these same techniques applied to (2.19a), using (2.18a), give a poor approximation, i.e., E_1 is $O(\Delta)$, yet we know from Theorem 2.1 that the scheme is second-order accurate. A careful accounting of E_1 must produce an estimate similar to (2.8). The following lemma examines the solution of (2.19a) and provides a key to many of the results in the following sections.

LEMMA 2.2. *On any mesh, define an $(N + 1)$-vector, E_1, by*

(2.20a)
$$(E_1)_0 = 0,$$

(2.20b)
$$(E_1)_i = \sum_{j=1}^{i} \Delta_j(T_1)_j, \qquad i = 1, \ldots, N.$$

Then E_1 is a solution of (2.19a). *Further,*

(2.21a)
$$\|E_1\|_\infty \leqslant \|T_1\|_\infty,$$

(2.21b)
$$\|D_0 E_1\|_\infty = \|T_1\|_\infty.$$

Proof. Let us write (2.19a) as

$$D_1(D_0 E_1 - T_1) = 0.$$

The quantity in parentheses must be in the null space of D_1, that is, a constant vector that we shall denote as $\lambda \mathbf{1}$. The vector E_1 must satisfy

(2.22)
$$D_0 E_1 = T_1 + \lambda \mathbf{1}.$$

The components of this equation are

$$(E_1)_i = (E_1)_{i-1} + \Delta_i(T_1)_i + \lambda \Delta_i, \qquad i = 1, \ldots, N-1.$$

Summing the right-hand side yields

$$(E_1)_i = (E_1)_0 + \sum_{j=1}^{i} \Delta_j(T_1)_j + \lambda x_i.$$

We are only interested in a particular solution of (2.19a) and so shall choose $(E_1)_0 = \lambda = 0$, which yields (2.20a,b). The bound (2.21a) is derived by taking absolute values in (2.20b) to get

$$|(E_1)_i| \leqslant \left| \sum_{j=1}^{i} \Delta_j \right| \|T_1\|_\infty \leqslant \|T_1\|_\infty.$$

Equality (2.21b) follows immediately from (2.22) with $\lambda = 0$. □

In the next section, we show that many truncation error expressions, like (2.15) and (1.6), can be rewritten in the form

$$T = D_1 T_1 + T_2,$$

where T_1 and T_2 are both second-order. Lemma 2.2 will then provide the means for showing that the solution error is second-order accurate in such cases.

3. Linear Second-Order Equations. In Section 2, we saw that at least one difference scheme approximating the equation

$$y'' = f(x)$$

with Dirichlet boundary conditions yields second-order accurate solutions on non-uniform meshes. In this section, we extend our discussion to linear second-order equations of the form (1.1) and (1.2a,b). Second-order accurate difference schemes for (1.1) can be constructed by forming an associated first-order system, discretizing with centered differences, and carefully eliminating the auxiliary unknowns as described in Section 2. More will be said about this and higher-order equations in a

subsequent paper. In this paper, we prove a result that will help to identify those three-point difference approximations to (1.1) and (1.2a,b) that are second-order accurate.

We use the notation of Section 1, namely,

(3.1) $$(L_h V)_i = (F)_i, \qquad i = 1, \ldots, N - 1,$$

(3.2a) $$B_0 V = b_{02},$$

(3.2b) $$B_1 V = b_{12}.$$

The truncation error will be defined by replacing V with Y in (3.1) and (3.2a,b).

Many difference schemes have the property that the truncation error can be divided into two parts,

(3.3) $$T = \tilde{T}_1 + \tilde{T}_2,$$

where

(3.4a) $$(\tilde{T}_1)_0 = (\tilde{T}_1)_N = 0,$$

(3.4b) $$(\tilde{T}_1)_i = (\Delta_{i+1} - \Delta_i) p(x_i), \qquad i = 1, \ldots, N - 1,$$

for some function $p(x) \in C^1(0, 1)$, and

(3.4c) $$\|\tilde{T}_2\|_\infty \leqslant C\Delta^2_{\max}.$$

The next lemma will show that difference schemes of this type may be second-order accurate.

LEMMA 3.1. *Let the truncation error for a difference scheme (3.1) and (3.2a,b) satisfy (3.3) and (3.4a,b,c); then*

(3.5a) $$(T)_0 = (\tilde{T}_2)_0,$$

(3.5b) $$(T)_N = (\tilde{T}_2)_N,$$

(3.5c) $$(T)_i = (D_1 T_1 + T_2)_i, \qquad i = 1, \ldots, N - 1,$$

where T_1 is an N-vector, T_2 is an $(N + 1)$-vector, D_1 is defined in (2.11a), and

$$\|T_1\|_\infty \leqslant C\Delta^2_{\max}, \qquad \|T_2\|_\infty \leqslant C\Delta^2_{\max}.$$

Proof. From (3.4b), note that

$$(\tilde{T}_1)_i = \frac{\Delta^2_{i+1}}{\Delta_{i+1} + \Delta_i} p(x_i) - \frac{\Delta^2_i}{\Delta_{i+1} + \Delta_i} p(x_i).$$

The function $p(x)$ is differentiable, so we can rewrite this expression as

$$(\tilde{T}_1)_i = \frac{1}{\Delta_{i+1} + \Delta_i} \left[\Delta^2_{i+1} p(x_{i+1/2}) - \Delta^2_i p(x_{i-1/2}) \right]$$

$$+ \frac{\Delta^3_{i+1}}{2(\Delta_{i+1} + \Delta_i)} p'(\xi_1) + \frac{\Delta^3_i}{2(\Delta_{i+1} + \Delta_i)} p'(\xi_2),$$

for some ξ_1 and $\xi_2 \in (x_{i-1}, x_{i+1})$. Now, identify

(3.6) $$(T_1)_i \equiv \tfrac{1}{2} \Delta^2_i p(x_{i-1/2}), \qquad i = 1, \ldots, N,$$

and

$$(3.7) \qquad (T_2)_i \equiv (\tilde{T}_2)_i + \frac{1}{2(\Delta_{i+1} + \Delta_i)} \left[\Delta_{i+1}^3 p'(\xi_1) + \Delta_i^3 p'(\xi_2) \right],$$

$$i = 1, \ldots, N - 1.$$

With these definitions, we have proved (3.5a,b,c). If we recall that $p(x)$ is continuously differentiable on $[0, 1]$, then (3.6) implies that

$$\| T_1 \|_\infty \leqslant \max_{[0,1]} \left\{ | p(x) | \right\} \Delta_{\max}^2,$$

with no constraints on the mesh. To complete the definition of T_2, we take

$$(T)_0 \equiv (\tilde{T}_2)_0 \equiv (T_2)_0, \qquad (T)_N \equiv (\tilde{T}_2)_N \equiv (T_2)_N.$$

A bound for (3.7) independent of the mesh is derived by noting that

$$\frac{1}{2(\Delta_{i+1} + \Delta_i)} \left[\Delta_{i+1}^3 p'(\xi_1) + \Delta_i^3 p'(\xi_2) \right] \leqslant \max_{[0,1]} \left\{ | p'(x) | \right\} \Delta_{\max}^2.$$

Thus, again without any constraints on the mesh, we have

$$\| T_2 \|_\infty \leqslant C \Delta_{\max}^2. \qquad \square$$

The hypothesis of Lemma 3.1, Eqs. (3.4a,b,c), is satisfied by many difference schemes. The second divided difference applied to y'' yields

$$(T)_i = \tfrac{1}{3}(\Delta_{i+1} - \Delta_i) y''' + O(\Delta^2), \qquad i = 1, \ldots, N - 1.$$

Consider several reasonable methods of approximating y'. If we take the difference of values at x_{i-1} and x_{i+1}, we have

$$\left[-\frac{1}{\Delta_{i+1} + \Delta_i}, 0, \frac{1}{\Delta_{i+1} + \Delta_i} \right] \begin{bmatrix} y_{i-1} \\ y_i \\ y_{i+1} \end{bmatrix} = y' + \frac{1}{2}(\Delta_{i+1} - \Delta_i) y'' + O(\Delta^2).$$

If we average function values on the left and right subintervals and take the difference of the averages, we have

$$\left[-\frac{1}{2\Delta_i}, \frac{\Delta_{i+1} - \Delta_i}{2\Delta_i \Delta_{i+1}}, \frac{1}{2\Delta_{i+1}} \right] \begin{bmatrix} y_{i-1} \\ y_i \\ y_{i+1} \end{bmatrix} = y' + \frac{1}{4}(\Delta_{i+1} - \Delta_i) y'' + O(\Delta^2).$$

Recall that the approximation (1.3b) used in Pearson [21] is

$$\left[\frac{-\Delta_{i+1}}{\Delta_i(\Delta_{i+1} + \Delta_i)}, \frac{\Delta_{i+1} - \Delta_i}{\Delta_{i+1}\Delta_i}, \frac{\Delta_i}{\Delta_{i+1}(\Delta_{i+1} + \Delta_i)} \right] \begin{bmatrix} y_{i-1} \\ y_i \\ y_{i+1} \end{bmatrix} = y' + O(\Delta^2).$$

Each of the functions on the right-hand sides of these truncation errors is evaluated at x_i. Note that each scheme has a truncation error of the form (3.3), and that each of these approximations is the same on a uniform mesh. With Lemmas 2.2 and 3.1 in hand, we now examine the question of second-order accuracy for more general difference schemes.

THEOREM 3.2. *Let the difference scheme* (3.1) *and* (3.2a,b) *have the following properties*:

(i) *the difference scheme is stable for all meshes in some class M with* $\Delta_{\max} \leqslant \Delta_0$;

(ii) $L_h = D_2 + L_1$, where D_2 is the second divided difference; and

(iii) the truncation errors satisfy the hypothesis of Lemma 3.1 (Eqs. (3.3) and (3.4a,b,c)).

Then, the error at the mesh points is bounded by

(3.8)
$$\|E\|_\infty \leqslant C[\|T_1\|_\infty + \|T_2\|_\infty + \|RE_1\|_\infty],$$

where E_1 is defined as in Lemma 2.2, T_1 and T_2 are defined in Lemma 3.1, and R is an $(N+1) \times (N+1)$-matrix,

(3.9a)
$$(RV)_0 = (B_0 V)_0,$$

(3.9b)
$$(RV)_i = (L_1 V)_i, \qquad i = 1, \ldots, N-1,$$

(3.9c)
$$(RV)_N = (B_1 V)_N.$$

In other words, R is the difference operator without the second divided difference.

Proof. The equations for the pointwise error, E, are

$$(B_0 E)_0 = (T)_0,$$
$$([D_2 + L_1]E)_i = (T)_i, \qquad i = 1, \ldots, N-1,$$
$$(B_1 E)_N = (T)_N.$$

Recalling that $D_2 = D_1 D_0$, and using Lemma 3.1, we have

$$(B_0 E)_0 = (T_2)_0,$$
$$([D_1 D_0 + L_1]E)_i = (D_1 T_1)_i + (T_2)_i, \qquad i = 1, \ldots, N-1,$$
$$(B_1 E)_N = (T_2)_N.$$

Let us define

$$E = E_1 + E_2,$$

where E_1 is as given in (2.20a,b). Then, because E_1 satisfies $D_1 D_0 E_1 = D_1 T_1$, we have

$$(B_0 E_2)_0 = (T_2)_0 - (RE_1)_0,$$
$$(L_h E_2)_i = (T_2)_i - (RE_1)_i, \qquad i = 1, \ldots, N-1,$$
$$(B_N E_2)_N = (T_2)_N - (RE_1)_N.$$

By hypothesis, the difference scheme is stable, which yields the estimate

$$\|E_2\|_\infty \leqslant C(\|T_2\|_\infty + \|RE_1\|_\infty).$$

Combining this with the bound for E_1 found in Lemma 2.2 yields

$$\|E\|_\infty \leqslant \|E_1\|_\infty + \|E_2\|_\infty \leqslant C(\|T_1\|_\infty + \|T_2\|_\infty + \|RE_1\|_\infty). \quad \square$$

The shortcoming of this general result is that we do not have bounds on the last term in (3.8). Before we prove a more practical result, let us examine what this term represents. In most schemes, L_1 will be a consistent approximation to a first-order differential operator, say

(3.10a)
$$g(x) \frac{d}{dx} + h(x),$$

and B_0, B_1 will be consistent approximations to the boundary conditions

(3.10b)
$$b_{01}\frac{d}{dx} + b_{00},$$

(3.10c)
$$b_{11}\frac{d}{dx} + b_{10}.$$

Taking liberties with notation, the error E_1, recall (2.20a,b), can be written as

$$E_1 \equiv \int_0^x T_1\,dx.$$

Applying (3.10a,b,c) to this expression yields

(3.11a)
$$g(x)T_1 + h(x)\int_0^x T_1\,dx,$$

(3.11b)
$$b_{01}(T_1)_0,$$

(3.11c)
$$b_{11}(T_1)_N + b_{10}\int_0^1 T_1\,dx.$$

Clearly, all terms can be bounded by $C\|T_1\|_\infty$. Tikhonov and Samarskiĭ [23] remark that their result illustrates the global nature of the error. Expressions (3.11a,b,c) show that there is both a global and a local component to the error if $g(x) \neq 0$.

Corollary 3.3 will show that the heuristic argument developed above is correct.

COROLLARY 3.3. *Let the discrete boundary value problem satisfy all the conditions of Theorem 3.2. Further, let B_0, L_1, B_1 be consistent, three-point difference approximations*:

(3.12a)
$$(B_0V)_0 = \alpha_0 v_0 + \beta_0 v_1 + \gamma_0 v_2,$$

(3.12b)
$$(L_1V)_i = \alpha_i v_{i-1} + \beta_i v_i + \gamma_i v_{i+1}, \qquad i = 1,\ldots, N-1,$$

(3.12c)
$$(B_1V)_N = \alpha_N v_{N-2} + \beta_N v_{N-1} + \gamma_N v_N$$

to $b_{00} + b_{01}\,d/dx$, $a(x)d/dx + b(x)$, *and* $b_{10} + b_{11}\,d/dx$, *respectively; let*

(3.13)
$$\max_i\{|\gamma_i\Delta_{i+1}|, |\alpha_i\Delta_i|\} \leqslant C, \qquad i = 1,\ldots, N-1;$$

and let $\max\{|\gamma_0\Delta_2|, |\alpha_0\Delta_1|, |\gamma_N\Delta_N|, |\alpha_N\Delta_{N-1}|\} \leqslant C$. *Then,*

(3.14)
$$\|Y - V\|_\infty \leqslant C\Delta_{\max}^2.$$

Proof. Theorem 3.2 and Lemma 3.1 yield the result

$$\|E\|_\infty = \|Y - V\|_\infty \leqslant C\Delta_{\max}^2 + C\|RE_1\|_\infty,$$

so it is this last term we need to estimate. Looking at (3.12b), we have

$$(RE_1)_i = (\alpha_i + \beta_i + \gamma_i)(E_1)_i + [\gamma_i\Delta_{i+1}(T_1)_{i+1} - \alpha_i\Delta_i(T_1)_i],$$

and, consequently,

$$|(RE_1)_i| \leqslant |\alpha_i + \beta_i + \gamma_i|\,\|E_1\|_\infty + 2\max_i\{|\gamma_i\Delta_{i+1}|, |\alpha_i\Delta_i|\}\|T_1\|_\infty.$$

Clearly, consistency requires that $|(\alpha_i + \beta_i + \gamma_i)| \leqslant C$ for all meshes with $\Delta_{\max} \leqslant \Delta_0$. From Lemma 2.2, we have

$$\|E_1\|_\infty \leqslant C\|T_1\|_\infty;$$

thus,

$$\left|(RE_1)_i\right| \leqslant C\|T_1\|_\infty.$$

The boundary conditions (3.12a,c) are handled in a similar fashion. Coupled with Lemma 3.1, we now have

$$\|E\|_\infty \leqslant C\Delta_{\max}^2. \quad \square$$

Condition (3.13) is, at worst, a constraint on the local mesh ratio. However, if we look at the difference scheme (1.4), we note that

$$|\gamma_i \Delta_{i+1}| = \left| a(x_i) \frac{\Delta_i}{\Delta_{i+1} + \Delta_i} \right| \leqslant \max_{[0,1]} \{|a(x)|\},$$

and

$$|\alpha_i \Delta_i| = \left| a(x_i) \frac{\Delta_{i+1}}{\Delta_{i+1} + \Delta_i} \right| \leqslant \max_{[0,1]} \{|a(x)|\},$$

so that (3.13) is satisfied with no mesh restrictions at all. For the boundary conditions, we might use the second-order accurate scheme at $x = 1$. This yields

$$b_{10}v_N + b_{11}\left[\frac{\Delta_N}{\Delta_{N-1}(\Delta_N + \Delta_{N-1})}v_{N-2} \right.$$
$$\left. - \frac{\Delta_N + \Delta_{N-1}}{\Delta_N \Delta_{N-1}}v_{N-1} + \frac{2\Delta_N + \Delta_{N-1}}{\Delta_N(\Delta_N + \Delta_{N-1})}v_N \right] = b_{12}.$$

Thus, we have

$$|\gamma_N \Delta_N| \leqslant |b_{10}|\Delta_N + |b_{11}|\frac{2\Delta_N + \Delta_{N-1}}{\Delta_N + \Delta_{N-1}} \leqslant C,$$

$$|\alpha_N \Delta_{N-1}| \leqslant |b_{11}|\frac{\Delta_N}{\Delta_N + \Delta_{N-1}} \leqslant C.$$

This also does not require any special restrictions on the mesh. The same thing will be true for this boundary condition applied at $x = 0$. Thus, we have shown that Pearson's scheme, if stable for a class of meshes with $\max_i\{\Delta_i\} \leqslant \Delta_0$, will be second-order accurate for those same meshes. The same is true for many other commonly used difference schemes.

4. Quasilinear, Second-Order Equations. In this section, we extend the basic result for linear boundary value problems to the solution of quasilinear boundary value problems. These will be of the form

(4.1) $$y'' = f(y', y, x), \qquad x \in (0,1),$$

(4.2a) $$b_0(y(0), y'(0)) = 0,$$

(4.2b) $$b_1(y(1), y'(1)) = 0.$$

Throughout we will assume that (4.1) and (4.2a,b) have an isolated solution, see Keller and White [13], $y(x) \in C^4[0,1]$. The difference equations approximating this boundary value problem will be denoted by

(4.3) $$(N_h(V))_i = 0, \qquad i = 1, \ldots, N-1;$$

(4.4a) $B_0(V) = 0;$

(4.4b) $B_1(V) = 0.$

Together, we write these equations as

$$\text{(4.5a)} \qquad N(V) = \begin{bmatrix} B_0(V) \\ N_h(V) \\ B_1(V) \end{bmatrix},$$

and once again assuming that y'' is approximated by the second divided difference, let

$$\text{(4.5b)} \qquad R(V) = N(V) - \begin{bmatrix} 0 \\ D_2V \\ 0 \end{bmatrix} = \begin{bmatrix} B_0(V) \\ N_1(V) \\ B_1(V) \end{bmatrix}.$$

The truncation error, T, will be defined in the usual way,

$$\text{(4.6)} \qquad N(Y) = T.$$

Stability of the difference scheme for quasilinear problems is a local phenomenon (cf. Keller and White [13], Doedel [8]). We will define a cylinder around the solution by

$$S_1(Y, \rho) = \left\{ V \mid \|V - Y\|_\infty \leqslant \rho,\ \|D_0V - Y'\|_\infty \leqslant \rho \right\},$$

where D_0V is an N-vector consisting of first divided differences and $(Y')_i = y'(x_{i-1/2})$, $i = 1, \ldots, N$. Stability of the difference scheme (4.3) and (4.4a,b) is taken to mean there exists C, and some $\rho > 0$, such that

$$\text{(4.7)} \qquad \|V - W\| \leqslant C\|N(V) - N(W)\|,$$

for every $V, W \in S_1(Y, \rho)$. With these definitions, we have the following generalization of Theorem 3.2.

THEOREM 4.1. *Let the difference scheme* (4.3) *and* (4.4a,b) *have the following properties*:

   (i) *the difference scheme is stable, in the sense of* (4.7), *for all meshes in some class M with $\Delta_{\max} \leqslant \Delta_0$*;
   (ii) $N_h(V) = D_2V + N_1(V)$, *where D_2 is the second divided difference*; *and*
   (iii) *the truncation errors satisfy Eqs.* (3.3) *and* (3.4a,b,c).

Then, if $N(V) = 0$, $V \in S_1(Y, \rho)$, the error at the mesh points satisfies

$$\text{(4.8)} \qquad \|E\|_\infty \leqslant C\left[\|T_1\|_\infty + \|T_2\|_\infty + \|R(Y - E_1) - R(Y)\|_\infty\right],$$

where, as before, E_1 is defined in Lemma 2.2, T_1 and T_2 are defined in Lemma 3.1, and R is defined in (4.5b).

Proof. Let us write the error, E, in the following way:

$$E = E_1 + E_2.$$

Since $E = Y - V$, we have the expression

$$E_2 = (Y - E_1) - V.$$

By assumption, $V \in S_1(Y, \rho)$, and by Lemma 2.2, we can choose Δ_0 sufficiently small so that $Y - E_1 \in S_1(Y, \rho)$. Applying the stability bound (4.7) yields

$$\|E_2\|_\infty \leqslant C \|N(Y - E_1)\|_\infty.$$

Because $N(Y) = T$, where T is the truncation error, we may write

$$\|E_2\|_\infty \leqslant C \|N(Y - E_1) - N(Y) + T\|_\infty.$$

Let us denote the right-hand side as

$$F = N(Y - E_1) - N(Y) + T.$$

Consider the components of this equation. We have for the boundary conditions

(4.9a) $$(F)_0 = B_0(Y - E_1) - B_0(Y) + (T_2)_0,$$

(4.9b) $$(F)_N = B_1(Y - E_1) - B_1(Y) + (T_2)_N,$$

and for the differential equation, we have

$$(F)_i = (D_2 Y - D_2 E_1 + N_1(Y - E_1) - D_2 Y - N_1(Y))_i + D_1(T_1)_i + (T_2)_i$$

for $i = 1, \ldots, N - 1$. Recalling the definition of E_1, this becomes

(4.9c) $$(F)_i = (T_2)_i + (N_1(Y - E_1) - N_1(Y))_i.$$

Combining (4.9a,b,c), we have

$$\|E_2\|_\infty \leqslant C \big[\|T_2\|_\infty + \|R(Y - E_1) - R(Y)\|_\infty\big].$$

From Lemma 2.2, we know that $\|E_1\|_\infty \leqslant C\|T_1\|_\infty$, which yields the result:

$$\|E\|_\infty \leqslant C \big[\|T_1\|_\infty + \|T_2\|_\infty + \|R(Y - E_1) - R(Y)\|_\infty\big]. \qquad \square$$

This error estimate reduces to (3.8) if R is linear. However, an extra hypothesis is required for Theorem 4.1 because it is a local result; that is, we assume that $V \in S_1(Y, \rho)$. This condition can be removed in many instances by further restricting the maximum allowable mesh size, but we will not duplicate the work of Keller [14], Doedel [8], Keller and White [13], and others here.

A result very similar to Corollary 3.3 requires that we look closely at the linearized difference equations (4.3) and (4.4a,b). These linear operators are given below:

(4.10) $$D_2 V + \frac{\partial N_1}{\partial V}(Y)V,$$

(4.11a) $$\frac{\partial B_0}{\partial V}(Y)V,$$

(4.11b) $$\frac{\partial B_1}{\partial V}(Y)V,$$

where the i, jth element of $\partial N_1 / \partial V$ is $\partial (N_1)_i / \partial v_j$.

COROLLARY 4.2. *Let the discrete boundary value problem satisfy the conditions of Theorem 4.1. Further, let the linearized, discrete operators $\partial R(Y)/\partial V$ be consistent, three-point difference approximations to the first-order portion of the linearized boundary value problem; let $\partial R / \partial V$ be uniformly continuous in $S_1(\rho, Y)$; let*

(4.12) $$\max_i \left\{ \left| \frac{\partial (R)_i}{\partial v_{i+1}} \Delta_{i+1} \right|, \left| \frac{\partial (R)_i}{\partial v_{i-1}} \Delta_i \right| \right\} \leqslant C, \qquad i = 1, \ldots, N - 1;$$

and let

$$\left\{ \left| \frac{\partial B_0}{\partial v_2} \Delta_2 \right|, \left| \frac{\partial B_0}{\partial v_0} \Delta_1 \right|, \left| \frac{\partial B_1}{\partial v_N} \Delta_N \right|, \left| \frac{\partial B_1}{\partial v_{N-2}} \Delta_{N-1} \right| \right\} \leqslant C.$$

Then,

$$\| Y - V \|_\infty \leqslant C \Delta_{\max}^2.$$

Proof. The proof proceeds in very much the same way as Corollary 3.3. Theorem 4.1 and Lemma 3.1 yield the result:

(4.13) $$\| E \|_\infty \leqslant C \Delta_{\max}^2 + C \| R(Y - E_1) - R(Y) \|_\infty.$$

To estimate the last term on the right-hand side of (4.13), we write

$$R(Y - E_1) - R(Y) = - \int_0^1 \frac{\partial R}{\partial V} (Y - sE_1) E_1 \, ds.$$

Now, we define

(4.14) $$[\alpha_i(Z), \beta_i(Z), \gamma_i(Z)] \equiv \left[\frac{\partial (R)_i}{\partial v_{i-1}}(Z), \frac{\partial (R)_i}{\partial v_i}(Z), \frac{\partial (R)_i}{\partial v_{i+1}}(Z) \right],$$

so that we have

$$(R(Y - E_1) - R(Y))_i$$
$$= - \int_0^1 \left(\alpha_i (Y - sE_1)(E_1)_{i-1} + \beta_i (Y - sE_1)(E_1)_i + \gamma_i (Y - sE_1)(E_1)_{i+1} \right) ds$$

for $i = 1, \ldots, N - 1$. Recalling that $(E_1)_i = \sum_{j=0}^i \Delta_j (T_1)_j$, we see that

$$(R(Y - E_1) - R(Y))_i$$
$$= - \int_0^1 \left(\alpha_i (Y - sE_1) + \beta_i (Y - sE_1) + \gamma_i (Y - sE_1) \right)(E_1)_i \, ds$$
$$- \int_0^1 \left(\gamma_i (Y - sE_1)\Delta_{i+1}(T_1)_{i+1} - \alpha_i (Y - sE_1)\Delta_i (T_1)_i \right) ds$$

for $i = 1, \ldots, N - 1$. Because the difference operator (4.14) with $Z = Y$ is a consistent approximation to a first-order differential operator, and because $\partial R(Z)/\partial V$ is uniformly continuous on $S_1(Y, \rho)$, we have

$$[\alpha_i (Y - sE_1) + \beta_i (Y - sE_1) + \gamma_i (Y - sE_1)] \leqslant C$$

for all meshes with Δ_0 sufficiently small. Likewise, we have assumed that

$$\max_i \left\{ |\gamma_i(Y)\Delta_{i+1}|, |\alpha_i(Y)\Delta_i| \right\} \leqslant C$$

so that

$$\max_i \left\{ |\gamma_i(Y - sE_1)\Delta_{i+1}|, |\alpha_i(Y - sE_1)\Delta_i| \right\} \leqslant C$$

for all meshes with Δ_0 sufficiently small. The boundary conditions can be handled in a similar fashion; thus, from Eq. (4.14), we have

$$\| R(Y - E_1) - R(Y) \|_\infty \leqslant C \left[\| E_1 \|_\infty + \| T_1 \|_\infty \right],$$

and finally from (4.13), we have

$$\| E \|_\infty \leqslant C\Delta^2_{\max}$$

for all meshes with Δ_0 sufficiently small. □

These results show that the nonlinearity in the boundary value problem (4.1) and (4.2a,b) poses no essential difficulty in extending the results of Section 3.

5. Cell-Centered Schemes. In many fluid mechanics calculations, it is convenient to approximate some physical variables at cell-centers rather than at nodes or vertices of the mesh. In one dimension, a cell-centered mesh is constructed by placing the nodes x_1 ($> x_0$) and x_{N-1} ($< x_N$) on the interval boundaries. Quantities of interest are then approximated at $x_{i-1/2} = \frac{1}{2}(x_i + x_{i-1})$, $i = 1, \ldots, N$. Actually, we have been dealing with cell-centered quantities all along; it would just have been confusing to point them out as such. For example, all the truncation errors in the discussion of first-order systems in Section 2 are cell-centered. In this section, V will still denote a vector, but the components may be given as $(V)_i \equiv v_i$, as before, if it is a vertex-centered quantity, or $(V)_{i-1/2} \equiv v_{i-1/2}$ if it is a cell-centered quantity. In each case, the subscript denotes the variable's association with the mesh point x_i or the cell-centered point $x_{i-1/2}$. The cell-centered mesh distances will be written as $\Delta_{i-1/2} \equiv \frac{1}{2}(\Delta_i + \Delta_{i-1}) = (x_{i-1/2} - x_{i-3/2})$.
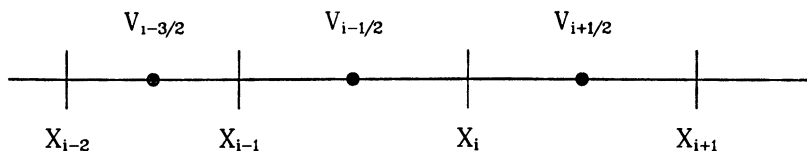

FIGURE 5.1
Portion of a cell-centered mesh.

We will examine the solution of the linear, second-order boundary value problem (1.1) and (1.2a,b). For cell-centered schemes, the usual approximation to y'' can be explained using Figure 5.1. Assume that the left difference $[(v_{i-1/2} - v_{i-3/2})/\Delta_{i-1/2}]$ is an accurate representation of y' at x_{i-1} and, similarly, $[(v_{i+1/2} - v_{i-1/2})/\Delta_{i+1/2}]$ at x_i. Now, the difference of these first differences gives us an approximation to y'' at $x_{i-1/2}$:

$$(5.1) \qquad \frac{1}{\Delta_i}\left[\frac{v_{i+1/2} - v_{i-1/2}}{\Delta_{i+1/2}} - \frac{v_{i-1/2} - v_{i-3/2}}{\Delta_{i-1/2}} \right] \equiv (D_2^{cc}V)_{i-1/2},$$

where superscript cc denotes cell-centered. Many people have noted (see [23]) that (5.1) is, in general, an inconsistent approximation to the second derivative on a nonuniform mesh. That is,

$$
\begin{aligned}
(D_2^{cc}Y)_{i-1/2} = {} & y''(x_{i-1/2}) + \frac{\Delta_{i+1} - 2\Delta_i + \Delta_{i-1}}{4\Delta_i}y''(x_{i-1/2}) \\
& + \frac{\Delta_{i+1}^2 + 2\Delta_{i+1}\Delta_i - 2\Delta_i\Delta_{i-1} - \Delta_{i-1}^2}{24\Delta_i}y'''(x_{i-1/2}) + \cdots,
\end{aligned}
$$

(5.2)

and the approximation is inconsistent. We will show that some difference schemes still yield $O(\Delta^2)$ solution errors in spite of this truncation error.

We will employ first-order systems, in an unorthodox way, to help us find the correct way to analyze these methods. The major complications will arise from the $O(1)$ truncation error term in (5.2). Of itself, it is easily treated, but its effect on boundary condition approximations and on difference approximations to $ay' + by$ will be considerable.

First, we notice that (just as for the vertex-centered scheme) the cell-centered approximation to y'' is given by multiplying the matrices D_0 and D_1 together but in the reverse order. That is,

$$(5.3) \qquad (D_2^{cc}V)_{i-1/2} \equiv (D_0D_1V)_{i-1/2}, \qquad i = 2, \ldots, N-1,$$

where D_0 and D_1 are precisely as defined in (2.4a) and (2.11a), except that the first and last rows of D_0 are deleted.

Consider the solution of the constant-coefficient equation,

$$y'' + ay' + by = f.$$

Written as a first-order system, we can approximate this equation by

$$(5.4a) \qquad\qquad D_1V = \bar{A}_1W,$$

$$(5.4b) \qquad\qquad D_1W = -a\bar{A}_1W - b\bar{A}_1V + F,$$

where $(F)_i = f_i$ and

$$(5.5) \qquad \bar{A}_1 \equiv \begin{bmatrix} \dfrac{\Delta_2}{\Delta_1 + \Delta_2} & \dfrac{\Delta_1}{\Delta_1 + \Delta_2} & & \\ & \dfrac{\Delta_3}{\Delta_2 + \Delta_3} & \dfrac{\Delta_2}{\Delta_2 + \Delta_3} & \\ & & \ddots & \end{bmatrix}.$$

Replacing V and W with Y and Y' (exact derivative), we get

$$(5.6a) \qquad\qquad D_1Y = \bar{A}_1Y' + D_1T_1 + \bar{T}_2,$$

$$(5.6b) \qquad\qquad D_1Y' = -a\bar{A}_1Y' - b\bar{A}_1Y + F + D_1\bar{T}_3 + \bar{T}_4,$$

where

$$(5.7a) \qquad\qquad (T_1)_{i-1/2} = \tfrac{1}{8}\Delta_i^2 y''_{i-1/2},$$

$$(5.7b) \qquad (\bar{T}_2)_i = -\tfrac{1}{24}\left(2\Delta_{i+1}^2 + \Delta_{i+1}\Delta_i + 2\Delta_i^2\right)y_i''' + O(\Delta^3),$$

$$(5.7c) \qquad\qquad (\bar{T}_3)_{i-1/2} = \tfrac{1}{8}\Delta_i^2 y'''_{i-1/2},$$

$$(5.7d) \qquad (\bar{T}_4)_i = \tfrac{1}{24}\left(2\Delta_{i+1}^2 + \Delta_{i+1}\Delta_i + 2\Delta_i^2\right)\left(ay_i''' + by_i''\right) + O(\Delta^3).$$

We can find appropriate commuting pairs of matrices,

$$(5.8) \qquad\qquad A_0D_1 = D_0\bar{A}_1,$$

which allow us to combine (5.6a,b) into

$$(5.9) \qquad \begin{aligned} &D_0D_1Y + aD_0\bar{A}_1Y + bA_0\bar{A}_1Y \\ &= A_0F + D_0D_1T_1 + D_0\left(\bar{T}_2 + \bar{A}_1\bar{T}_3 + a\bar{A}_1T_1\right) + A_0\left(\bar{T}_4 + a\bar{T}_2\right). \end{aligned}$$

Note that (5.9) defines a three-point, cell-centered difference scheme approximating this constant-coefficient differential equation.

528   THOMAS A. MANTEUFFEL AND ANDREW B. WHITE, JR.

Now we will examine the pointwise error equation implied by (5.9), namely,

$$(5.10) \quad \begin{aligned} D_0 D_1 E &+ a D_0 \overline{A}_1 E + b A_0 \overline{A}_1 E \\ &= D_0 D_1 T_1 + D_0 (\overline{T}_2 + \overline{A}_1 \overline{T}_3 + a \overline{A}_1 T_1) + A_0 (\overline{T}_4 + a \overline{T}_2), \end{aligned}$$

in order to see how the analysis should proceed in a more general case. In the course of this examination, we will point out the purpose of each of the four lemmas in Appendix A leading to the main result of this section, Theorem 5.1. (For Appendix A, see the supplements section at the end of this issue.) The first question to consider is what sort of truncation error can be put in a form similar to (5.9). This question will be answered in Lemma A.1.

The leading term in the truncation error is easily removed by defining

$$E = T_1 + E_1.$$

Substituting this into (5.10), we have

$$\begin{aligned} D_0 D_1 E_1 &+ a D_0 \overline{A}_1 E_1 + b A_0 \overline{A}_1 E_1 \\ &= D_0 (\overline{T}_2 + \overline{A}_1 \overline{T}_3 + a \overline{A}_1 T_1) + A_0 (\overline{T}_4 + a \overline{T}_2) - (a D_0 \overline{A}_1 + b A_0 \overline{A}_1) T_1. \end{aligned}$$

We recognize the last term on the right-hand side of this equation to be $-L_1 T_1$, where L_1 is a difference approximation to $ay' + by$. Rewriting this equation gives us

$$(5.11) \quad (D_0 D_1 + L_1) E_1 = D_0 (\overline{T}_2 + \overline{A}_1 \overline{T}_3 + a \overline{A}_1 T_1) + A_0 (\overline{T}_4 + a \overline{T}_2) - L_1 T_1.$$

Note that this error equation has the same form as those encountered in Sections 2 and 3 if we can rewrite $L_1 T_1$ as $D_0 \overline{T}_5 + \overline{T}_6$, where $\overline{T}_5$ and $\overline{T}_6$ are $O(\Delta^2)$. Lemma A.2 will derive conditions sufficient for (5.11) to be written as

$$(5.12) \quad (D_0 D_1 + L_1) E_1 = D_0 T_2 + A_0 T_4 + \overline{T}_6,$$

where, in our particular case, $T_2 = \overline{T}_2 + \overline{A}_1 \overline{T}_3 + a \overline{A}_1 T_1 + \overline{T}_5$, $T_4 = \overline{T}_4 + a \overline{T}_2$, and $\overline{T}_6$ are $O(\Delta^2)$.

The remainder of the proof proceeds exactly like that in Section 3. Let

$$(5.13a) \quad E_1 = E_2 + E_3,$$

where

$$(5.13b) \quad D_1 E_2 = T_2.$$

Lemma A.3 will prove that $E_2 = O(\Delta^2)$. Substituting (5.13a) into (5.12) yields the following equation for E_3,

$$(D_0 D_1 + L_1) E_3 = A_0 \overline{T}_4 + \overline{T}_6 - L_1 E_2.$$

Finally, Lemma A.4 will show that under mild assumptions on L_1, we have $L_1 E_2 = O(\Delta^2)$, and thus stability (plus boundary condition stuff) will yield second-order accuracy for these schemes. These four lemmas are proven in Appendix A, but in some cases are tedious and ugly and should be passed over except by the most resolute readers.

The main result of this section utilizes the results in Lemmas A.2–A.4. This theorem corresponds to Corollary 3.3 for vertex-centered meshes. The discrete boundary conditions will be written as

$$(5.14a) \quad B_0 V = \alpha_1 v_{1/2} + \beta_1 v_{3/2} + \gamma_1 v_{5/2},$$

and

$$(5.14b) \quad B_1 V = \alpha_N v_{N-5/2} + \beta_N v_{N-3/2} + \gamma_N v_{N-1/2}.$$

The discrete approximation to the differential equation is given by

(5.15) $([D_0 D_1 + L_1]V)_{i-1/2} = (F)_{i-1/2}, \qquad i = 2, \ldots, N - 1,$

where $(F)_{i-1/2}$ may stand for a linear combination of the function $f(x)$ evaluated at certain points.

THEOREM 5.1. *Let the difference scheme* (5.14a,b) *and* (5.15) *satisfy the hypotheses of Lemmas A.2 and A.4 and be stable for all meshes in some class M. Further, let the difference scheme have the following properties*:
   (i) *the truncation error for the differential equation approximation* (5.15) *has the form*

(5.16) $(D_0 D_1 + L_1)Y = F + D_0 D_1 T_1 + D_0 \overline{T}_2 + \overline{T}_3,$

where $T_1 \overline{T}_2 \overline{T}_3$ *are all* $O(\Delta^2)$;
   (ii) *the truncation error for the boundary condition approximations has the form*,

(5.17a) $B_0(Y - T_1) = b_{02} + (\overline{T}_3)_{1/2},$

(5.17b) $B_1(Y - T_1) = b_{12} + (\overline{T}_3)_{N-1/2},$

where $(\overline{T}_3)_{1/2}$ *and* $(\overline{T}_3)_{N-1/2}$ *are* $O(\Delta^2)$; *and*
   (iii) *the discrete boundary conditions satisfy*

(5.18a) $|\alpha_1 + \beta_1 + \gamma_1|, |\Delta_{1/2}\alpha_1|, |\Delta_{3/2}\gamma_1| \leqslant C,$

(5.18b) $|\alpha_N + \beta_N + \gamma_N|, |\Delta_{N-3/2}\alpha_N|, |\Delta_{N-1/2}\gamma_N| \leqslant C.$

Then,

$$\|Y - V\|_\infty \leqslant C\Delta_{\max}^2$$

for all meshes in M.

Proof. From (5.16), we can derive the pointwise error equation for the differential equation,

$$(D_0 D_1 + L_1)E = D_0 D_1 T_1 + D_0 \overline{T}_2 + \overline{T}_3.$$

Let $E = T_1 + E_1$, and we get

$$(D_0 D_1 + L_1)E_1 = D_0 \overline{T}_2 + \overline{T}_3 - L_1 T_1.$$

However, Lemma A.2 allows us to write

$$L_1 T_1 = D_0 \overline{T}_1 - \overline{T}_6,$$

where $\overline{T}_1, \overline{T}_6 = O(\Delta^2)$. Defining $\overline{T}_2 - \overline{T}_1 = T_2$, we have

$$(D_0 D_1 + L_1)E_1 = D_0 T_2 + (\overline{T}_3 + \overline{T}_6).$$

Let $E_1 = E_2 + E_3$, where E_2 is defined in Lemma A.3; then

$$(D_0 D_1 + L_1)E_3 = \overline{T}_3 + \overline{T}_6 - L_1 E_2.$$

Finally, Lemma A.4 shows that $L_1 E_2 = O(\Delta^2)$; thus we can write

(5.19) $([D_0 D_1 + L_1]E_3)_{i-1/2} = (T_3)_{i-1/2}, \qquad i = 2, \ldots, N - 1,$

where $(T_3)_{i-1/2} = (\overline{T}_3 + \overline{T}_6 - L_1 E_2)_{i-1/2}.$

The pointwise error equations for the boundary conditions are

$$B_0 E = B_0 T_1 + (\overline{T}_3)_{1/2}, \qquad B_1 E = B_1 T_1 + (\overline{T}_3)_{N-1/2}.$$

Substituting $E = T_1 + E_2 + E_3$ (unraveling the substitutions made for the discrete differential equations), we have

$$B_0 E_3 = (\overline{T}_3)_{1/2} - B_0 E_2, \qquad B_1 E_3 = (\overline{T}_3)_{N-1/2} - B_1 E_2.$$

Property (iii) in the hypothesis is sufficient (see the proof of Lemma A.4) to show that

$$|B_0 E_2|, |B_1 E_2| \leqslant C\Delta_{\max}^2.$$

Thus, we define

$$(T_3)_{1/2} = (\overline{T}_3)_{1/2} - B_0 E_2, \qquad (T_3)_{N-1/2} = (\overline{T}_3)_{N-1/2} - B_1 E_2,$$

and we know that

$$\|T_3\|_\infty \leqslant C\Delta_{\max}^2.$$

Pulling all this together, we have

$$B_0 E_3 = (T_3)_{1/2},$$

$$([D_0 D_1 + L_1] E_3)_{i-1/2} = (T_3)_{i-1/2}, \qquad i = 2, \ldots, N-1,$$

$$B_1 E_3 = (T_3)_{N-1/2},$$

and stability of the difference scheme yields

$$\|E_3\| \leqslant C\Delta_{\max}^2.$$

The error is given by

$$E = Y - V = T_1 + E_2 + E_3,$$

where we already know that T_1, E_2 are $O(\Delta^2)$, so

$$\|Y - V\|_\infty \leqslant C\Delta_{\max}^2. \quad \square$$

This theorem points out that the boundary conditions for cell-centered schemes must be handled carefully in order to preserve second-order accuracy in the solution. For purely Dirichlet boundary conditions ($b_{01} = b_{11} = 0$), the condition that

$$B_i(Y - T_1) - b_{i2} = O(\Delta^2)$$

will not alter the 'usual' approximation because $B_i T_1 = O(\Delta^2)$. However, for mixed boundary conditions ($b_{01} \neq 0$ or $b_{11} \neq 0$), failure to satisfy this condition will result in a solution that is first-order accurate. As an example, we will derive the correct boundary conditions at $x = 0$ for the difference scheme (5.4a,b). Recall from (5.7a) that $(T_1)_{i-1/2} = \frac{1}{8}\Delta_i^2 y_{i-1/2}''$. Thus, we have

$$\alpha_1\left(y_{1/2} - \tfrac{1}{8}\Delta_1^2 y_{1/2}''\right) + \beta_1\left(y_{3/2} - \tfrac{1}{8}\Delta_2^2 y_{3/2}''\right) + \gamma_1\left(y_{5/2} - \tfrac{1}{8}\Delta_3^2 y_{5/2}''\right) - b_{02} = O(\Delta^2).$$

Expanding in Taylor series about $x = 0$ gives us

$$(\alpha_1 + \beta_1 + \gamma_1)y_1 + \left(-\tfrac{1}{2}\Delta_1\alpha_1 + \tfrac{1}{2}\Delta_2\beta_1 + \left[\tfrac{1}{2}\Delta_2 + \Delta_{3/2}\right]\gamma_1\right)y_1'$$
$$+ (\Delta_2\Delta_{5/2})\gamma_1 y_1'' = b_{00}y_1 + b_{01}y_1' O(\Delta^2).$$

A discrete boundary condition satisfying this equation is

$$(5.20\text{a}) \qquad b_{00}\left[\frac{\Delta_1 v_{3/2} + \Delta_2 v_{1/2}}{\Delta_2 + \Delta_1}\right] + b_{01}\left[\frac{v_{3/2} - v_{1/2}}{\Delta_{3/2}}\right] = b_{02}.$$

At $x = 1$, we have a corresponding condition

$$(5.20\text{b}) \quad b_{10}\left[\frac{\Delta_{N-1} v_{N-1/2} + \Delta_N v_{N-3/2}}{\Delta_N + \Delta_{N-1}}\right] + b_{11}\left[\frac{v_{N-1/2} - v_{N-3/2}}{\Delta_{N-1/2}}\right] = b_{12}.$$

Numerical evidence leads us to believe that if property (iii) in Theorem 5.1 is violated, then the solution will not be second-order accurate. That is, this is a genuine constraint and not an artifact of the proof.

6. Summary. We have shown that most reasonable, vertex-centered difference schemes for linear and quasilinear boundary value problems yield second-order accurate solutions on nonuniform meshes. This result is derived in spite of the fact that the truncation error for these compact schemes is only first-order. Each of these results can be extended to systems of second-order equations without substantial difficulty.

The numerical results shown in Figures 6.1–6.4, following, are solutions to the boundary value problem

$$(6.1) \qquad (3x + 1)^2 y'' + 9y' + 9\left[\tfrac{1}{4}(3x + 1)^2 - 1\right] y = 0,$$
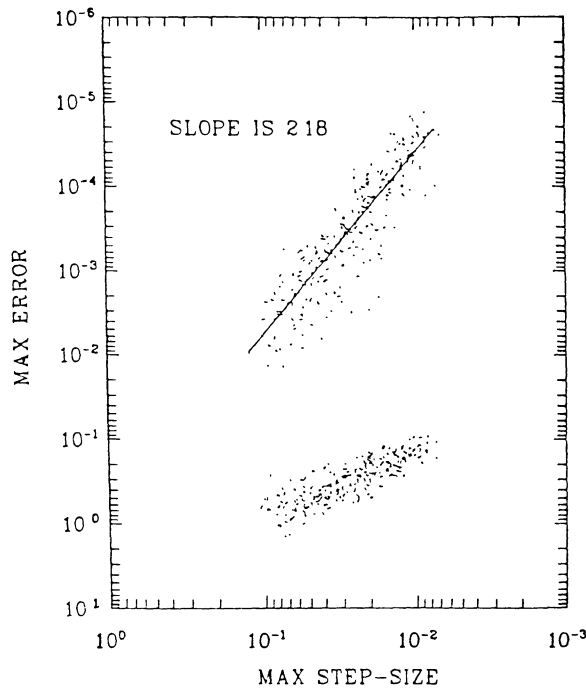
$$(6.2) \qquad\qquad y'(0),\; y(1) \quad given.$$

FIGURE 6.1
First-order system.
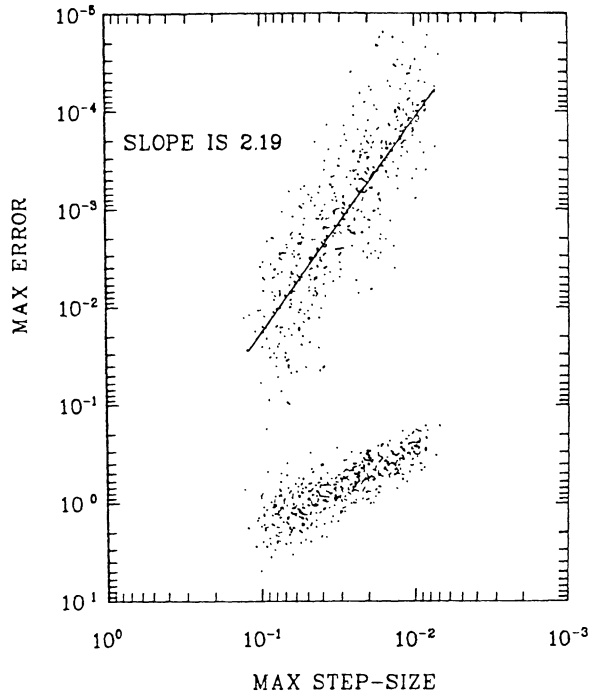
FIGURE 6.2
Pearson's scheme.

In each figure, the differential equation is approximated on 700 random meshes ($N - 1$ points placed in $(0, 1)$ at random) and the maximum error is plotted versus the maximum step size; N ranges from 100 to 800.

Figure 6.1 is the vertex-centered difference scheme derived from a first-order system (see (2.14)). The actual solution error for each mesh appears in the upper portion of this plot; the straight line "through" these points is the least-squares fit to the solution error. In the lower portion of this plot is the truncation error for each mesh. The slope of the least-squares fit to these points is 0.78. Figure 6.2 is the vertex-centered difference scheme used by Pearson [21] and displays similar information as Figure 6.1.

In both these figures, the results of Theorem 3.2 and Corollary 3.3 are borne out. That is, the truncation error is of lower order than the solution error. For these meshes, the local mesh ratio ($\Delta_{max}/\Delta_{min}$) may be as large as 10^7, so this is a very stringent test.

We have also shown that some cell-centered schemes for linear boundary value problems (quasilinear equations should pose no essential difficulty) yield second-order accurate solutions. This is in spite of the fact that these schemes are inconsistent. Most important, this result reveals that mixed boundary conditions must be approximated carefully in order to preserve this accuracy.

Figures 6.3 and 6.4 show the solution error in solving (6.1) by the cell-centered scheme (5.1) and (A.7). In Figure 6.3, the boundary conditions are approximated correctly using (5.20a,b), and the solution error is second order. In Figure 6.4, the
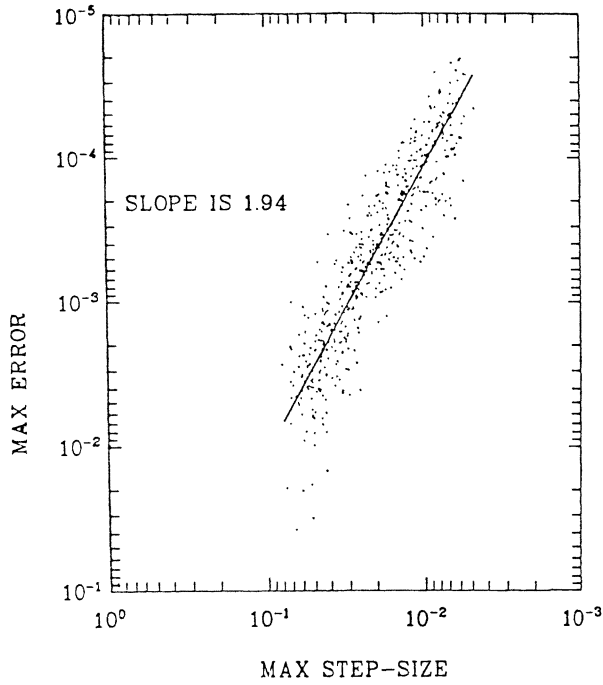
FIGURE 6.3

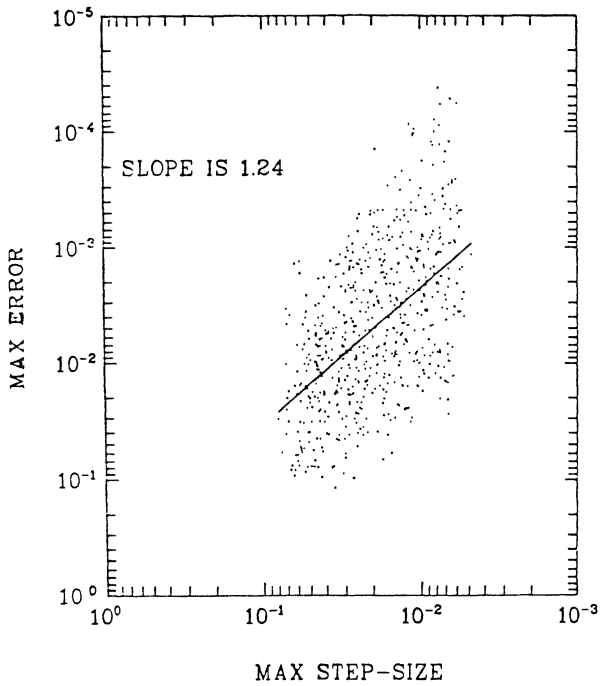*Correct cell-centered boundary conditions.*

FIGURE 6.4

Second-order cell-centered boundary conditions.

boundary conditions are approximated correct to second order:

$$B_0 Y = b_{00} y(0) + b_{01} y'(0) + O(\Delta^2),$$

$$B_1 Y = b_{10} y(1) + b_{11} y'(1) + O(\Delta^2),$$

but they violate property (ii) of Theorem 5.1. Here, the solution error has degraded to first order, although it still remains of higher order than the truncation error.

Motivated by the elimination of auxiliary unknowns in first-order systems, we have developed mathematical tools that, together with stability arguments, yield better estimates of accuracy. Our work has shown the limitations of the standard stability-consistency arguments when applied to nonuniform meshes. On the other hand, these results can also be obtained by combining stability with summation by parts; see, e.g., Kreiss et al. [16]. In addition, one should note the alternative approach of Grigorieff [9] utilizing stability and consistency with respect to a Spijker norm.

The reduction process described in Section 2 provides an efficient computational algorithm (see Manteuffel and White [19]). In addition, the concept of pairwise commuting matrices can be used to extend this analysis to higher-order differential equations. These results will appear in a subsequent paper.

Finally, it is easy to see how to extend these results to separable difference schemes on separable meshes. For example, consider the solution of 'the heat equation

$$(6.3) \qquad\qquad u_t = u_{xx}$$

on the grid (x_i, t_k). A Crank-Nicolson scheme for approximating (6.3) is

$$(L_h V)_i^k \equiv \frac{v_i^k - v_i^{k-1}}{\Delta t_k} - \frac{1}{2}\left(D_2 V^k + D_2 V^{k-1} \right)_i \equiv 0,$$

where D_2 is the second divided difference (1.3a), and

$$V^k = \left(v_0^k \; v_1^k \; \cdots \; v_N^k \right)^T.$$

In order to show that the error is second order, we follow the prescription given in Section 3.

Let

$$(6.4) \qquad\qquad \left(E_1^k \right)_i = \frac{1}{3} \sum_{j=1}^{i} \left(\Delta_j \right)^3 u_{xxx}\left(x_{j-1/2}, t^k \right),$$

and $E^k = E_1^k + E_2^k$. The resulting equation for E_2 is

$$\left(L_h E_2 \right)_i^k = \left(T_2 \right)_i^k - \frac{\left(E_1^k \right)_i - \left(E_1^{k-1} \right)_i}{\Delta t_k},$$

where $T_2 = O(\Delta_{max}^2 + \Delta t_{max}^2)$. Clearly, if the time-difference operator applied to E_1 is $O(\Delta^2)$, then stability will yield the desired result. Substituting from (6.4) yields

$$\frac{\left(E_1^k \right)_i - \left(E_1^{k-1} \right)_i}{\Delta t_k} = \frac{1}{3} \sum_{j=1}^{i} \left(\Delta_j \right)^3 \left[\frac{u_{xxx}\left(x_{j-1/2}, t_k \right) - u_{xxx}\left(x_{j-1/2}, t_{k-1} \right)}{\Delta t_k} \right]$$

$$\approx \frac{1}{3} \sum_{j=1}^{i} \left(\Delta_j \right)^3 u_{xxxt}\left(x_{j-1/2}, t_k \right).$$

Thus, if the exact solution is sufficiently smooth, then

$$\left| \frac{\left(E_1^k \right)_i - \left(E_1^{k-1} \right)_i}{\Delta t_k} \right| \leqslant C\Delta_{max}^2.$$

This same argument will work for difference schemes approximating elliptic, hyperbolic, or parabolic equations provided (1) that the solution is smooth, and (2) that the difference operators in one direction are not affected by the nonuniform mesh in another direction.

Computing and Communications Division
Los Alamos National Laboratory
Los Alamos, New Mexico 87545

1. C. M. ABLOW & S. SCHECHTER, "Campylotropic coordinates," *J. Comput. Phys.*, v. 27, 1978, pp. 351–362.

2. T. H. CHONG, "A variable mesh finite difference method for solving a class of parabolic differential equations in one space variable," *SIAM J. Numer. Anal.*, v. 15, 1978, pp. 835–857.

3. MELVYN CIMENT, STEPHEN W. LEVENTHAL & BERNARD C. WEINBERG, "The operator compact implicit method for parabolic equations," *J. Comput. Phys.*, v. 28, 1978, pp. 135–166.

4. STEPHEN F. DAVIS & JOSEPH E. FLAHERTY, "An adaptive finite element method for initial-boundary value problems for partial differential equations," *SIAM J. Sci. Statist. Comput.*, v. 3, 1982, pp. 6–27.

5. V. E. DENNY & R. B. LANDIS, "A new method for solving two-point boundary value problems using optimal node distribution," *J. Comput. Phys.*, v. 9, 1972, pp. 120–137.

6. EUGENIA KAUNAX DE RIVAS, "On the use of nonuniform grids in finite difference equations," *J. Comput. Phys.*, v. 10, 1972, pp. 202–210.

7. EUSEBIUS J. DOEDEL, "The construction of finite difference approximations to ordinary differential equations," *SIAM J. Numer. Anal.*, v. 15, 1978, pp. 450–465.

8. EUSEBIUS J. DOEDEL, "Finite difference methods for nonlinear two-point boundary value problems," *SIAM J. Numer. Anal.*, v. 16, 1979, pp. 173–185.

9. ROLF D. GRIGORIEFF, "Some stability inequalities for compact finite difference schemes," typescript.

10. J. D. HOFFMAN, "Relationship between the truncation errors of centered finite-difference approximations on uniform and nonuniform meshes," *J. Comput. Phys.*, v. 46, 1982, pp. 469–474.

11. EUGENE ISAACSON & H. B. KELLER, *Analysis of Numerical Methods*, Wiley, New York, 1966.

12. HERBERT B. KELLER, "Accurate difference methods for linear ordinary differential systems subject to linear constraints," *SIAM J. Numer. Anal.*, v. 6, 1969, pp. 8–30.

13. H. B. KELLER & A. B. WHITE, JR., "Difference methods for boundary value problems in ordinary differential equations," *SIAM J. Numer. Anal.*, v. 12, 1975, pp. 791–802.

14. HERBERT B. KELLER, *Numerical Solution of Two Point Boundary Value Problems*, SIAM, Regional Conference Series in Applied Mathematics, 24.

15. H.-O. KREISS, "Difference approximations for boundary and eigenvalue problems for ordinary differential equations," *Math. Comp.*, v. 26, 1972, pp. 605–624.

16. H.-O. KREISS, T. A. MANTEUFFEL, B. SWARTZ, B. WENDROFF & A. B. WHITE, JR., "Supra-convergent schemes on irregular grids," *Math. Comp.*, v. 47, 1986, pp. 537–554.

17. DAVID LEVERMORE, private communication.

18. R. E. LYNCH & J. R. RICE, *The HODIE method*, Report C5D-TR 188, Dept. of Computer Science, Purdue University, W. Lafayette, Indiana.

19. T. A. MANTEUFFEL & A. B. WHITE, JR., "On the efficient solution of systems of second-order boundary value problems," *SIAM J. Numer. Anal.* (To appear.)

20. M. R. OSBORNE, "Minimizing truncation error in finite difference approximations to ordinary differential equations," *Math. Comp.*, v. 21, 1967, pp. 133–145.

21. CARL E. PEARSON, "On a differential equation of boundary layer type," *J. Math. Phys.*, v. 47, 1968, pp. 134–154.

22. BLAIR K. SWARTZ, "The construction and comparison of finite difference analogs of some finite element schemes," in *Mathematical Aspects of Finite Elements Partial Differential Equation* (1974).

23. A. N. TIKHONOV & A. A. SAMARSKIĬ, "Homogeneous difference schemes on nonuniform nets," *Zh. Vychisl. Mat. i Mat. Fiz.*, v. 1, 1962, English translation (Russian).

24. A. B. WHITE, JR., "On selection of equidistributing meshes for two-point boundary value problems," *SIAM J. Numer. Anal.*, v. 16, 1979, pp. 472–502.