MATHEMATICS OF COMPUTATION, VOLUME 28, NUMBER 126, APRIL 1974, PAGES 413–447

Stable Approximations for Hyperbolic Systems with Moving Internal Boundary Conditions*

By M. Goldberg and S. Abarbanel

Abstract. The work of Kreiss on the stability theory of difference schemes for the mixed initial boundary value problem for linear hyperbolic systems is extended to deal with the case of the pure initial value problem with an internal boundary. The case of an internal boundary X_B that moves with constant speed c is treated, i.e., $X_B = X_0 + ct$. In particular, the stability of "hybrid" schemes is studied by using the Lax-Wendroff scheme at points that are not on the internal boundary, while using a first order accurate scheme at the internal boundary points. Numerical evidence is given that the results of the linear stability analysis describes the qualitative behavior of such schemes for nonlinear cases, when the internal boundary is a shock.

0. Introduction. The treatment of boundary conditions associated with the finite difference approximations of initial value problems governed by systems of hyperbolic partial differential equations is of great importance in many computational applications. The analysis of the linear stability of such mixed initial boundary value problems in the one-dimensional quarter space case $x \geqq 0$, $t \geqq 0$ was given by Kreiss ([8], [9]), and it will be assumed that the reader is familiar with these works which are the basis of our paper.

There are many cases where the initial value problem may be thought of as having conditions specified at internal boundaries. Internal boundaries may model, for example, in hydrodynamics, shock waves, contact discontinuities and, in general, very narrow regions with strong gradients. In many such cases, there is no external boundary and the analysis cannot be confined to the quarter space but must consider the full half space $-\infty < x < \infty$, $t \geqq 0$. The stability theory of finite difference approximations in the half space, with conditions imposed on internal boundaries, is the main object of this paper.

We show first, in Chapter 1, that Kreiss' Main Theorem of [9] is valid for problems in the half space with slightly modified conditions applied at an internal boundary located at $x = 0$. In Chapter 2, we derive some general stability conditions for problems with internal boundaries moving at a constant speed. The motivation for seeking this type of result is clear, since, in general, the internal boundaries will not be stationary. In Chapters 3 and 4, we specialize and sharpen the results of Chapter 2 in the case of the Lax-Wendroff finite difference scheme with various types of boundary conditions. These boundary conditions are all in the form of first order accurate finite difference schemes. The reason why we pay this particular

Received April 4, 1973.

AMS (MOS) subject classifications (1970). Primary 65M10; Secondary 65N10.

* This research has been sponsored in part by the Air Force Office of Scientific Research (NAM) through the European office of Aerospace Research, AFSC, United Air Force, under Grant AFOSR-72-2370.

attention to the configurations of Chapters 3 and 4 is that, in order to avoid difficulties in regions of strong gradients when performing practical computations in fields like fluid mechanics, first order accurate schemes are used in those regions even when the basic overall finite difference approximation is of higher order accuracy. In Chapter 5, we present numerical evidence that the results of the linear stability analysis presented here are, as is usually expected, valid also for nonlinear problems.

It is hoped that the recent work of Gustafsson, Kreiss and Sundström [6] will allow the generalization of the present results to the case of nondissipative finite difference schemes, with variable coefficients.

The computations reported in this work were done on the CDC 6600 computer at the Tel Aviv University Computation Center.

1. Modification of Kreiss' Problem [9].

1.1 *Formulation of the Problem.* Consider the Cauchy problem for a first order hyperbolic system of partial differential equations with constant coefficients

$$(1.1) \qquad u_t = Au_x; \qquad -\infty < x < \infty, t \geqq 0, \qquad u(x, 0) = f(x).$$

Here, $u = (u^{(1)}(x, t), \cdots, u^{(n)}(x, t))'$ is a n-dimensional unknown vector and A is a $n \times n$ coefficient matrix.** The assumption of hyperbolicity is that A can be diagonalized by a similarity transformation and that its eigenvalues are real. Hence, we may assume without restriction that A is diagonal and has real elements. It is also assumed that A is nonsingular (Assumption 1) and that the dependent variables $u^{(j)}(x, t)$, $1 \leqq j \leqq n$, are arranged so that the diagonal terms of A, a_j, $1 \leqq j \leqq n$, form a monotone increasing sequence

$$(1.2) \qquad a_1 \leqq a_2 \leqq \cdots \leqq a_l < 0 < a_{l+1} \leqq \cdots \leqq a_n.$$

We take into account the possibility that $l = 0$ or $l = n$, i.e., that A can also be either positive or negative.

Define a mesh-size $h \equiv \Delta x > 0$, $\Delta t > 0$ such that $\lambda \equiv \Delta t/\Delta x =$ constant. Using the notation $x_\nu \equiv \nu\Delta x$ and $v_\nu(t) \equiv v(x_\nu, t)$, consider a consistent difference approximation to (1.1) of the form

$$(1.3a) \qquad v_\nu(t + \Delta t) = Qv_\nu(t); \qquad v_\nu(0) = f_\nu, \qquad \nu = 0, \pm 1, \pm 2, \cdots,$$

where

$$(1.3b) \qquad Q = \sum_{j=-r}^{p} A_j E^j, \qquad Ev_\nu = v_{\nu+1}.$$

The A_j's are constant matrices of order n and it is assumed (Assumption 2) that, if $r > 0$, then A_{-r} is nonsingular; similarly, if $p > 0$, then A_p is nonsingular. We call the scheme (1.3) the *basic scheme.*

We now let ν_0 and $q \geqq 0$ be two fixed integers and suppose that we wish to apply, at every time step, the basic scheme at all $\nu \neq \nu_0, \cdots, \nu_0 + q$, while at $\nu_0 \leqq \nu \leqq \nu_0 + q$ we define $v_\nu(t + \Delta t)$ by a constant coefficient scheme, usually different from (1.3). The transformation

** If y is a vector, then y' is its transpose, y^* is its adjoint and $|y| = (\sum_{(j)} |y_j|^2)^{1/2}$ its Euclidean norm. Similar notations hold for matrices; for example, $|A| = \sup |Ay|$, $|y| = 1$.

$$t \to t, \qquad x \to \zeta + \nu_0 h \tag{1.4}$$

shows that we may assume $\nu_0 = 0$, i.e., that the region of applicability of the basic scheme is all $\nu \neq 0, \cdots, q$, whereas the values of $v_\nu(t + \Delta t)$ in the range $0 \leqq \nu \leqq q$ are defined by $q + 1$ relations of the form

$$v_\nu(t + \Delta t) = \sum_{j=-k}^{s} C_{j,\nu} v_{\nu+j}(t), \qquad \nu = 0, \cdots, q. \tag{1.5}$$

Here $C_{j,\nu}$ are constant matrices of order n. This process defines a sort of *internal-boundary-conditions* which constitute a perturbation of the basic scheme. We designate these conditions as the *perturbation*, and the combined process of applying the basic scheme (1.3) at all $\nu \neq 0, \cdots, q$, together with applying the perturbation (1.5), shall be called the *perturbed scheme*. The integers r, p, k, s and q in (1.3) and (1.5) are called the indices of the perturbed scheme.

By analogy with [9], we denote by H the space of all the vector grid functions w_ν defined for $-\infty < \nu < \infty$, for which $\sum_{-\infty}^{\infty} |w_\nu|^2 < \infty$. H is a Hilbert space, if an inner product and a norm are defined by

$$(w, v)_h = \sum_{\nu=-\infty}^{\infty} w_\nu^* v_\nu h, \qquad ||w||_h^2 = (w, w)_h. \tag{1.6}$$

We say that the perturbed scheme is *stable* if there exists a constant K, independent of Δt, such that

$$||v(t)||_h \leqq K \, ||v(0)||_h \quad \text{for all } t \equiv m\Delta t \text{ and all } v(0) \in H. \tag{1.7}$$

In particular, if $K \leqq 1$, we say that the scheme is *strongly stable*. If we write the perturbed scheme in operator form, we obtain

$$v(t + \Delta t) = Gv(t), \qquad v(t), v(t + \Delta t) \in H, \tag{1.8}$$

where G is a bounded linear operator on H, defined by (1.3) for $v \neq 0, \cdots, q$, together with (1.5). Hence the perturbed scheme is stable iff there is a constant K such that

$$||G^m||_h \leqq K, \qquad m = 0, 1, 2, \cdots. \tag{1.9}$$

Hereafter, we shall assume that the basic scheme is strongly stable; i.e., we require that the norm of the operator, which represents Q of (1.3) in H, will be bounded by 1. This requirement shall take the following form. Evidently, H is isomorphic to the space l^2, and, by the Riesz-Fisher Theorem, the spaces l^2 and $L^2(-\pi, \pi)$ with the usual norms are isometric. Hence, it is clear that H is isometrically isomorphic to the space $\hat{H}$ of all functions $\hat{w}(\zeta)$ in $L^2(-\pi, \pi)$, with inner product and norm defined by

$$(\hat{w}, \hat{v}) = \frac{h}{2\pi} \int_{-\pi}^{\pi} \hat{w}^* \hat{v} \, d\zeta, \qquad ||w||^2 = (\hat{w}, \hat{w}). \tag{1.10}$$

The representation of the basic scheme in $\hat{H}$ is given by the *amplification matrix*

$$\hat{Q}(\zeta) = \sum_{j=-r}^{p} A_j e^{ij\zeta}, \qquad |\zeta| \leqq \pi, \tag{1.11}$$

and we have $||\hat{Q}(\zeta)|| = \max |\hat{Q}(\zeta)|$, $|\zeta| \leqq \pi$. Therefore, the assumption that the basic scheme is strongly stable (Assumption 3) is that

$$(1.12) \qquad |\hat{Q}(\zeta)| \leqq 1, \qquad |\zeta| \leqq \pi.$$

In addition, we assume (Assumption 4) that the basic scheme is dissipative in the sense of Kreiss, i.e., that there exists a constant $\tau > 0$ and a positive integer ω such that, for each ζ, $|\zeta| \leqq \pi$, the eigenvalues $\mu(\zeta)$ of $\hat{Q}(\zeta)$ satisfy

$$(1.13) \qquad |\mu(\zeta)| \leqq 1 - \tau\, |\zeta|^{2\omega}.$$

All of the above assumptions are exactly those made in [9].

Our aim in this chapter is to make use of the results in Sections 1–3 of [9] and Section 3 of [8], in order to obtain sufficiency conditions for the stability of the perturbed difference scheme; i.e., conditions sufficient for the uniform boundedness of the natural powers of the operator G.

The problem considered here is a modification of Kreiss' problem in [9]. He considered the system (1.1) in the quarter plane $0 \leqq x < \infty$, $t \geqq 0$, with boundary conditions given at a single time level, whereas we consider (1.1) in the half plane $-\infty < x < \infty$, $t \geqq 0$, with boundary conditions (1.5) given at two time levels.

The stability results for the modified problem will have exactly the same form as those in the Main Theorem of [9], and will be presented in the next section.

1.2. *Kreiss' Main Theorem* [9] *for the Modified Problem.* Let

$$(1.14) \qquad \det\left[\sum_{j=-r}^{p} A_j\kappa^j - zI\right] = 0$$

be the *characteristic equation* corresponding to the basic scheme (1.3). If p and r are nonnegative, then Lemma 2 of [9] assures us that, for all z with $z \neq 1$, $|z| \geqq 1$, this characteristic equation has exactly $(r + p)n$ roots κ_j; nr of them with $|\kappa_j| < 1$ and np with $|\kappa_j| > 1$. It is clear that, if $r \leqq 0$ ($p \leqq 0$), then (1.14) has only np (nr) roots, all with $|\kappa_j| > 1$ ($|\kappa_j| < 1$). Note that, by continuity, the roots of (1.14) satisfy milder inequalities for $|z| \geqq 1$; i.e., in each of the above cases, $|\kappa_j| < 1$ ($|\kappa_j| > 1$) becomes $\kappa_j \leqq 1$ ($|\kappa_j| \geqq 1$). We shall henceforth refer to the results contained in this paragraph as Lemma 2 of [9].

We remark that Lemma 2—as well as Lemma 7 of [9] which we shall quote later in this section—depends on the properties of the basic scheme only, and not on the boundary conditions. Therefore, they are clearly valid in our case.

We continue with the analysis in analogy with Section 1 of [9]. Let z with $z \neq 1$, $|z| \geqq 1$, be given. In order that G have an eigensolution $g \in H$, with eigenvalue z, g must fulfill the requirements

$$(1.15) \qquad (Q - z)g_\nu = 0, \qquad \nu \neq 0, \cdots, q;$$

$$(1.16) \qquad zg_\nu = \sum_{j=-k}^{s} C_{j,\nu} g_{\nu+j}, \qquad \nu = 0, \cdots, q.$$

Since Eq. (1.14) has roots κ_j with $|\kappa_j| < 1$ ($|\kappa_j| > 1$) only when $r > 0$ ($p > 0$), it follows that the most general solution of the ordinary difference equation (1.15), belonging to H, is

$$(1.17a) \qquad g_\nu = g_\nu(z) = \sum_{|\kappa_j|<1} P_j\kappa_j^\nu, \qquad \nu \geqq q + 1 - r,$$

if $r > 0$; or

$$g_\nu = 0, \qquad \nu \geqq q + 1, \tag{1.17b}$$

if $r \leqq 0$; and

$$g_\nu = g_\nu(z) = \sum_{|\kappa_j|>1} P_j \kappa_j^\nu, \qquad \nu \leqq p - 1, \tag{1.18a}$$

if $p > 0$; or

$$g_\nu = 0, \qquad \nu \leqq -1, \tag{1.18b}$$

if $p \leqq 0$. Here, $P_j = P_j(\nu)$ are polynomials in ν with vector coefficients, where the degree of $P_j(\nu)$ is one less than the multiplicity of the corresponding κ_j. The part of the solution given by (1.17a), (1.18a), involves nr (np) independent solutions, and, therefore, it depends on nr (np) parameters σ_j. Define

$$\begin{aligned} \tilde{r} &\equiv r, \quad r > 0, & \tilde{p} &\equiv p, \quad p > 0, \\ &\equiv 0, \quad r \leqq 0, & &\equiv 0, \quad p \leqq 0, \end{aligned} \tag{1.19}$$

then we see that the general solution in H of (1.15), given by (1.17), (1.18), depends on $(\tilde{r} + \tilde{p})n$ parameters σ_j. It is important to note that this solution is neither always defined nor always single valued. It might turn out to be either undefined or, on the contrary, defined twice, for certain values of ν, depending on r, p and q.

If $q = \tilde{r} + \tilde{p} - 1$, then g_ν is uniquely defined via (1.17) and (1.18) for all ν. In addition, as we saw, g must also satisfy (1.16) which constitutes a system of $(q + 1)n$ homogeneous equations in the $(\tilde{p} + \tilde{r})n$ unknowns, σ_j. Thus, for the present case of $q = \tilde{r} + \tilde{p} - 1$, the number of equations is the same as the number of unknown parameters to be defined.

If, on the other hand, $q > \tilde{p} + \tilde{q} - 1$, then g_ν remains undefined for $q - \tilde{r} - \tilde{p} + 1$ values of ν, $\tilde{p} \leqq \nu \leqq q - \tilde{r}$. Therefore, in order to complete the solution, we must define these g_ν's. The definition of each of these n-dimensional vectors involves n additional parameters σ_j, i.e.,

$$g_{\tilde{p}} = (\sigma_{(\tilde{r}+\tilde{p})n+1}, \cdots, \sigma_{(\tilde{r}+\tilde{p}+1)n})',$$

$$g_{\tilde{p}+1} = (\sigma_{(\tilde{r}+\tilde{p}+1)n+1}, \cdots, \sigma_{(\tilde{r}+\tilde{p}+2)n})', \quad \text{etc.}$$

This procedure adds $(q - \tilde{r} - \tilde{p} + 1)n$ parameters σ_j, and thus again g depends, all told, on $(q + 1)n$ parameters. The substitution of g in (1.16), again yields a homogeneous system of $(q + 1)n$ equations with $(q + 1)n$ unknown parameters σ_j.

For the remaining possibility of $q < \tilde{p} + \tilde{r} - 1$, g_ν is indeed defined for every ν; but for $\tilde{p} + \tilde{r} - q - 1$ values of ν, $q - \tilde{r} + 1 \leqq \nu \leqq \tilde{p} - 1$, it is defined twice and thus it is possibly double valued over these points. Requiring uniqueness, we are led to a system of $(\tilde{p} + \tilde{r} - q - 1)n$ homogeneous equations in the $(\tilde{p} + \tilde{r})n$ parameters σ_j which define g. We also still have (1.16), which gives us $(q + 1)n$ additional equations.

To summarize, we get in every case a homogeneous system of

$$\gamma \equiv \max\{(q + 1)n, (\tilde{r} + \tilde{p})n\}$$

equations with an equal number of the unknowns σ_j, which define g completely and uniquely. This system may be written in the form

$$E(z)\sigma = 0, \qquad \sigma = (\sigma_1, \cdots, \sigma_\gamma)', \tag{1.20a}$$

where $E(z)$ is a $\gamma \times \gamma$ matrix. Since g is an eigensolution of G iff $E(z)\sigma = 0$ has a nontrivial solution, we are led to Lemma 3 of [9]: *z with $|z| \geqq 1$, $z \neq 1$, is an eigenvalue of G, iff* det $E(z) = 0$.

Now, Lemma 7 of [9] states that, as $z \to 1$ (for $|z| \geqq 1$, $z \neq 1$), precisely n roots of (1.14), κ_j, tend to 1; l of them from inside the unit disk and $n - l$ from its exterior. Therefore, $g(1)$, as defined above in (1.17), (1.18) and the subsequent discussion, does not in general belong to H; however, it does depend on the proper number of parameters σ_j. Substitution of $g(1)$ into (1.16) leads to the homogeneous system

$$E(1)\sigma = 0. \tag{1.20b}$$

Kreiss [9] defines $z = 1$ as a *generalized eigenvalue* of G iff (1.20b) has a nontrivial solution, i.e., if det $E(1) = 0$.

We now assert that Kreiss' Main Theorem of [9] is also valid in our case, i.e., for the modified problem. We rephrase it as follows:

The perturbed scheme is stable if

(1) *the Assumptions* 1–4 *are fulfilled,*

(2) *$z = 1$ is not a generalized eigenvalue of G,*

(3) *G has no eigenvalue z with $|z| \geqq 1$, $z \neq 1$.*

In order to prove this assertion, it is necessary to consider several steps needed in the various stages of the proof of the Main Theorem in Sections 2 and 3 of [9] and Section 3 of [8].

Following Kreiss in [9], we wish to estimate the solution $v(t)$ of the finite difference approximation in the form

$$v(t) = -\frac{1}{2\pi i}\oint_\Gamma z^m(G - zI)^{-1}\, dz\cdot v(0), \qquad t = m\Delta t, \tag{1.21}$$

where Γ is any contour in the complex plane which includes the spectrum of G in its interior. To estimate (1.21), we now study the resolvent $(G - zI)^{-1}$ of G.

Consider first the case of a *one-sided* basic scheme, i.e., $r \leqq 0$ or $p \leqq 0$. Say $r \leqq 0$ and let $v \in H$ be given. Let us compute explicitly $f = (G - zI)^{-1}v$; this is equivalent to finding the solution of $(G - zI)f = v$, i.e.,

$$[(G - zI)f]_\nu = v_\nu, \qquad -\infty < \nu < \infty. \tag{1.22}$$

Therefore, we first consider

$$(Q - zI)f_\nu = v_\nu, \qquad \nu \neq 0, \cdots, q, \tag{1.23a}$$

or, alternatively,

$$f_{\nu+p} = -A_p^{-1}\left(\sum_{j=-r}^{p-1} A_j f_{\nu+j} - zf_\nu - v_\nu\right), \qquad \nu \neq 0, \cdots, q. \tag{1.23b}$$

Define the vector

$$y_\nu = (f'_{\nu+p-1}, f'_{\nu+p-2}, \cdots, f'_{\nu-r}, \cdots, f'_\nu)', \tag{1.24}$$

which, together with (1.23b), leads to the one step formula

$$y_{\nu+1} = My_\nu + e_\nu, \qquad \nu \neq 0, \cdots, q, \tag{1.25a}$$

here M is an $np \times np$ matrix and e_ν is a vector, defined by

$$(1.25\text{b}) \qquad M = -\begin{pmatrix} A_p^{-1}A_{p-1} & \cdots & & A_p^{-1}A_{-r} & 0\cdots 0 & -zA_p^{-1} \\ -I & 0 & \cdot & & \cdots 0 & 0 \\ 0 & \cdot & & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & & \cdot & \cdot & \cdot & \cdot \\ \cdot & & \cdot & & 0 & \\ 0 & \cdot\ \cdot & \cdot & 0 & -I & 0 \end{pmatrix}, \qquad e_\nu = A_p^{-1}\begin{pmatrix} v_\nu \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{pmatrix}.$$

The most general solution of (1.25) is

$$(1.26\text{a}) \qquad y_\nu = \sum_{i=q+1}^{\nu-1} M^{\nu-i-1}e_i + M^{\nu-q-1}y_{q+1}, \quad \text{for } \nu \geqq q+2,$$

and

$$(1.26\text{b}) \qquad y_\nu = -\sum_{i=\nu}^{-1} M^{\nu-i-1}e_i + M^\nu y_0, \quad \text{for } \nu \leqq -1,$$

where y_{q+1} and y_0 are arbitrarily chosen. Now note that the eigenvalues of M are those of the characteristic equation (1.14), and since $r \leqq 0$, Lemma 2 of [9] shows that they are all outside the unit circle, provided $|z| \geqq 1$, $z \neq 1$. In this case, we have $|M^m| \to \infty$ as $m \to \infty$, and therefore, the part of the solution of (1.25), given by (1.26a), satisfies $\sum_{\nu=q+2}^{\infty}|y_\nu|^2 = \infty$, except when y_{q+1} takes on a particular value. Since $\sum_{\nu=-\infty}^{\infty}|y_\nu|^2 < \infty$ if and only if $||f||_h^2 < \infty$, we see that the general solution (1.26) does not usually define, through (1.24), a vector f which belongs to H. Thus, in order to obtain the most general solution of (1.25) which implies a vector f that belongs to H, we keep (1.26b), but replace (1.26a) by the unique special solution of (1.25) for which $\sum_{\nu=q+1}^{\infty}|y_\nu|^2 < \infty$. This special solution is

$$(1.26\text{c}) \qquad y_\nu = -\sum_{i=\nu}^{\infty} M^{\nu-i-1}e_i, \qquad \nu \geqq q+1.$$

Equation (1.26c) uniquely defines the f_ν's for all $\nu > q+1$, while (1.26b) depends on $y_0 = (f'_{p-1}, f'_{p-2}, \cdots, f'_0)'$. Hence, f is defined uniquely iff $f_0, \cdots, f_q$ are defined. The values of $f_0, \cdots, f_q$ must satisfy (1.22b) for $\nu = 0, \cdots, q$, and, with the aid of the boundary conditions (1.5), we obtain

$$(1.27) \qquad \sum_{i=-k}^{s} C_{\nu,i}f_{\nu+i} - zf_\nu = v_\nu, \qquad \nu = 0, \cdots, q.$$

If we set $y^0 \equiv (f'_q, f'_{q-1}, \cdots, f'_0)'$, (1.27) takes the form

$$(1.28) \qquad C(z)y^0 = e^0,$$

where $C(z)$ is a square matrix depending on the $C_{\nu,i}$'s as well as on z, and e^0 is a vector defined by the matrices A_i and the given $v \in H$ (note that $\dim y^0 \neq \dim y_0$ unless $p = q+1$).

Now Lemma 5 of [9] is established: *if z with $|z| \geqq 1$, $z \neq 1$, is not an eigenvalue of G, then $(G - zI)^{-1}$ exists in H and is bounded.* The proof is similar to Kreiss': If $C(z)^{-1}$ exists, then f is uniquely defined and the resolvent exists in H and is bounded.

If, on the other hand, $C(z)^{-1}$ does not exist, then the homogeneous system $C(z)\bar{y}^0 = 0$ has a nontrivial solution $\bar{y}^0 \equiv (\bar{f}'_q, \bar{f}'_{q-1}, \cdots, \bar{f}'_0)'$. Now define

$$\begin{aligned} \bar{y}_0 &= (\bar{f}'_{p-1}, \bar{f}'_{p-2}, \cdots, \bar{f}'_0)', && p \leqq q+1, \\ &= (0, \cdots, 0, \bar{f}'_q, \cdots, \bar{f}'_0)', && p > q+1. \end{aligned} \tag{1.29a}$$

Thus,

$$\begin{aligned} y_\nu &\equiv 0, && \nu \geqq q+1, \\ &\equiv \bar{y}_0, && \nu = 0, \\ &\equiv M^\nu \bar{y}_0, && \nu < 0, \end{aligned} \tag{1.29b}$$

is a solution of the homogeneous equation (1.25a) with $e_\nu = 0$. Equations (1.29) define a nontrivial solution of $(G - zI)f = 0$ and, hence, z is an eigenvalue of G;

From the above discussion, and with the help of the transformation $T(z)$ of Lemma 4 of [9], it should now be clear that Lemma 5 of [9] is valid also when the basic scheme is two-sided (i.e., $p > 0$ and $r > 0$). We remark that Lemma 6, which is the conclusion of Section 2 of [9], follows immediately, as in [9], from Lemma 5.

To complete the argument about the validity of the Main Theorem in our case, we note that Section 3 of [9] goes over with slight modifications; the same holds for Theorem 3 of [8], which is required at the end of the proof of the Main Theorem in [9].

At the end of this chapter, we should make the following remark: It would seem that the application of Kreiss' results to our problem could have been more easily effected by first transforming the problem from the half space $-\infty < x < \infty$, $t \geqq 0$, to the quarter space, by defining a new unknown $2n$-dimensional vector $W(x, t) = (u(x, t), u(-x, t))'$, $x \geqq 0$. This procedure would have yielded a new finite difference scheme

$$W_\nu(t + \Delta t) = \sum_{i=-S}^{S} B_i W_{\nu+i}(t), \tag{1.30a}$$

where

$$S = \max\{r, p\}. \tag{1.30b}$$

However, the matrix coefficients B_{-S} and B_S will be singular unless in the original basic scheme $r = p$. Therefore, Assumption 2 of the Main Theorem is not fulfilled and Kreiss' results cannot be taken over to the present problem. In most of the applications presented in the following chapters, we encounter a situation which is equivalent to having an original basic scheme with $p \neq r$, and hence the approach adopted here.

2. General Results. Before obtaining some general results concerning the stability of perturbed schemes, let us consider the basic scheme alone. Assume that, in (1.3), $r \leqq 0$ ($p \leqq 0$). Because of consistency and dissipativity, Lemma 2 of [9] guarantees, as we previously saw, that, for all z with $|z| \geqq 1$, $z \neq 1$, the characteristic equation (1.14) has exactly np (nr) roots κ_i, all of them satisfying $|\kappa_i| > 1$ ($|\kappa_i| < 1$). Now, suppose that the number of negative (positive) eigenvalues of the coefficient matrix A of (1.1) is $l > 0$ ($n - l > 0$), then, due to consistency, we have the result of Lemma 7 of [9] that l (or respectively $n - l$) of these κ_i's tend to 1 from inside (outside) the unit disc, as $z \rightarrow 1$. This is a contradiction and we have

COROLLARY 1. *If the coefficient matrix of the P.D.E. system* (1.1) *has any negative* (*positive*) *eigenvalues, then there cannot exist a finite difference scheme of the form* (1.3) *with* $r \leqq 0$ ($p \leqq 0$) *which is both consistent and dissipative.****

This result means that, in the general case, when A in (1.1) has both negative and positive eigenvalues, it is impossible to construct one-sided finite difference schemes which are consistent as well as dissipative. However, as we shall see in the next chapter, where A is positive (negative), it is possible to construct such consistent and dissipative right (left) one-sided schemes.

From now on, we consider the perturbed scheme, together with Assumptions 1–4 of the previous chapter. We start with

THEOREM 1. *If, in the basic scheme* (1.3) *and in the perturbation* (1.5), $r \leqq 0$ *and* $-k \geqq q$ (*or alternatively*, $p \leqq 0$ *and* $-s \geqq q$), *then the perturbed scheme is stable unless* $k = q = 0$ ($s = q = 0$). *In the latter case, the scheme is stable if the spectral radius of* $C_{0,0}$ *satisfies* $\rho(C_{0,0}) < 1$ *and is unstable if* $\rho(C_{0,0}) > 1$.

Proof. We consider the first set of conditions $r \leqq 0$ and $-k \geqq q$, and, in accordance with the Main Theorem of Kreiss, we try to determine whether the operator G has eigenvalues z with $|z| \geqq 1$. Since $r \leqq 0$, the most general solution of (1.15), for $\nu > q$, that belongs to H, is given, as we, saw, by (1.17b); i.e., we have $g_\nu = 0$ for all $\nu \geqq q + 1$.

Let us assume, to start with, that $-k \geqq q + 1$, rather than $-k \geqq q$. Then, using (1.17b), we get, from (1.16),

$$zg_\nu = \sum_{j=-k}^{s} C_{j,\nu} g_{\nu+j} = 0, \qquad \nu = 0, \cdots, q, \tag{2.1}$$

and, together with (1.17b), we have

$$g_\nu = 0, \qquad \nu \geqq 0. \tag{2.2}$$

For $\nu \leqq -1$, Eq. (1.15) can be written, for $r \leqq -1$ and $r = 0$, as

$$g_\nu = z^{-1}(A_p g_{\nu+p} + \cdots + A_{-r} q_{\nu-r}), \qquad \nu \leqq -1, \tag{2.3a}$$

or

$$g_\nu = (zI - A_0)^{-1}(A_p g_{\nu+p} + \cdots + A_1 g_{\nu+1}), \qquad \nu \leqq -1, \tag{2.3b}$$

respectively. Note that $zI - A_0$ is nonsingular for $|z| \geqq 1$; otherwise, $\kappa = 0$ is a root of (1.14), contradicting Lemma 2 of [9] which guarantees that all the roots of (1.14) satisfy $|\kappa_j| \geqq 1$ when $r \leqq 0$. Using (2.2), we get from (2.3) that $g_\nu = 0$ also for $\nu \leqq -1$ and this, together with (2.2), gives $g = 0$. Thus we have failed to construct a nontrivial eigensolution of $(G - zI)g = 0$ that belongs to H, and hence $|z| \geqq 1$ is not an eigenvalue (generalized eigenvalue if $z = 1$) of G and, by Kreiss' Main Theorem, we are assured of stability.

To finish the proof, we set $-k = q$. Again, using (1.17b), we get, from (1.16),

*** This result might have implications concerning the treatment of boundary conditions by one-sided finite difference approximations.

$$(2.4a)\qquad zg_\nu = \sum_{j=-k}^{s} C_{j,\nu}g_{\nu+j} = 0, \qquad \nu = 1, \cdots, q;$$

and, for $\nu = 0$,

$$(2.4b)\qquad zg_0 = \sum_{j=-k}^{s} C_{j,0}g_j = C_{-k,0}g_{-k}.$$

From (2.4), we see that, if $q > 0$, then $g_\nu = 0$ for $0 \leqq \nu \leqq q$, i.e., we are back at (2.1) and the theorem follows. If, on the other hand, $k = q = 0$, then (2.4b) gives

$$(2.5)\qquad (zI - C_{0,0})g_0 = 0.$$

Thus, if $\rho(C_{0,0}) < 1$, then $(zI - C_{0,0})^{-1}$ exists for all $|z| \geqq 1$ and, therefore, $g_0 = 0$, and again we are back at (2.1). If $\rho(C_{0,0}) > 1$, then there exists z_0 with $|z_0| > 1$ such that $(z_0I - C_{0,0})$ is singular and we can find a vector $g_0 \neq 0$ which satisfies (2.5). Therefore, for $z = z_0$, we have the ordinary finite difference equation (2.3a) (or (2.3b) if $r = 0$) which is of degree p, with the following p initial conditions:

$$(2.6)\qquad g_0 \neq 0, \qquad g_1 = \cdots = g_{p-1} = 0.$$

These initial conditions assume a nontrivial solution of (1.15) for $\nu < 0$. By Lemma 2 of [9], if $|z_0| \geqq 1$, $z_0 \neq 1$, then *all np* eigenvalues of (1.14) satisfy, in our case, ($r \leqq 0$) the inequality $|\kappa_j| > 1$; hence, it follows from (1.18a) that *any* solution of (1.15) for $\nu < 0$ belongs to H. Therefore, the nontrivial solution of (1.15) defined by the initial conditions (2.6) belongs to H and, together with (1.17b), it dictates an eigensolution of G, $g \neq 0$, with eigenvalue z_0. We now invoke Lemma 1 of [9] (which is independent of any other result of [9]). In this lemma, Kreiss shows that if G has an eigenvalue z_0 with $|z_0| > 1$, then G is unstable. This finishes our argument. The proof for the second set of conditions ($p \leqq 0$, $s \leqq -q$) is almost identical.

So far, we have dealt with the perturbation of the basic scheme on a set of fixed points of the spatial mesh. However, our aim is to apply the stability analysis to nonlinear problems containing in their solutions discontinuities or large gradients. These systems are most conveniently thought of as initial value problems with moving internal boundaries. In general, these internal boundaries (such as shock waves, etc.) do not stay in a fixed position on the x-axis but move with time along some trajectory in the (x, t)-plane. When, due to the large gradients, the basic scheme does not provide smooth enough results in the neighborhood of these moving internal boundaries, it might be effective to apply at each time step a local perturbation in the (small) region of discontinuities or large gradients. The rest of this chapter is devoted to this problem, with the stability analysis applied to the linearized versions.

Let $x = \chi(t)$ be a function that, at the various time levels $t_m = m\Delta t$, $m \geqq 0$, assumes grid-point values $\nu_m h$; i.e., $\nu_m h = \chi(t_m)$ where the ν_m's are integers. Suppose that, for every given $m \geqq 0$, we apply the basic scheme at every spatial grid point $\nu \neq \nu_m, \cdots, \nu_m + q$, while, in the range $\nu_m \leqq \nu \leqq \nu_m + q$, $v_\nu(t_{m+1})$ is defined, by analogy with (1.5), by $q + 1$ relations of the form

$$(2.7)\qquad v_\nu(t_{m+1}) = \sum_{j=-k}^{s} C_{j,\nu-\nu_m}v_{\nu+j}(t_m), \qquad \nu = \nu_m, \cdots, \nu_m + q.$$

Here also, $C_{j,\nu}$ are constant $n \times n$ matrices. This procedure defines, as before, internal boundary conditions, except that now the boundary is moving.

The stability problem of the perturbed scheme can be posed as follows. If G is the linear operator defined by (1.8) (i.e., the operator appropriate for the case $\nu_m = 0$, $m \geqq 0$), and T is the translation operator in H, namely $(Tv)_\nu = v_{\nu+1}$, then the basic finite difference scheme perturbed by (2.7) may be written in the form

$$v(t_{m+1}) = T^{\nu_m}GT^{-\nu_m}v(t_m). \tag{2.8}$$

Therefore, if we start with initial conditions $v(0) \in H$, we get

$$\begin{aligned} v(t_m) &= \left[\prod_{i=0}^{m-1} (T^{\nu_i}GT^{-\nu_i})\right]v(0) \\ &= T^{\nu_{m-1}}GT^{\nu_{m-2}-\nu_{m-1}}G \cdots GT^{\nu_0-\nu_1}GT^{-\nu_0}v(0). \end{aligned} \tag{2.9}$$

The difficulty we face here is that G and T do not commute (G is not a normal operator on H) and, therefore, even when we assume that the powers of G are uniformly bounded, and even though T is an isometry, we still cannot conclude anything concerning the boundedness of an operator of the form given in the square brackets of (2.9). Thus, we cannot deal with the most general perturbation; however, there are many important cases where we can solve the stability problem for the perturbed scheme.

Let us assume that

$$\chi(t) = (t \cdot \delta/\lambda) + \nu_0 h, \tag{2.10}$$

where δ is a given integer. In terms of the discrete variables, we have, taking without restrictions $\nu_0 = 0$,

$$\nu_m = (t_m \cdot \delta)/(\lambda h) + \nu_0 = m \cdot \delta. \tag{2.11}$$

The above choice for $\chi(t)$ means that we perturb the basic scheme along a straight line in the (x, t)-plane, and that the perturbed region moves at a constant speed δ/λ starting at $t = 0$ from the point $\nu_0 \Delta x$. It is clear that, when $\delta = \nu_0 = 0$, we go back to the original perturbation (1.5). Clearly, the line $x = \chi(t)$ can be thought of as representing the trajectory of a "shock-wave" or any other physical discontinuity.

Consider now the transformation

$$t \to t, \qquad x \to \xi + \chi(t). \tag{2.12}$$

Keep the time step Δt and choose $\Delta\xi = h$ as before. By this transformation, we obtain a new mesh

$$\{(\xi_\nu, t_m) \mid -\infty < \nu < \infty, m \geqq 0\}, \tag{2.13}$$

with the point (ξ_ν, t_m) being the image of $(x_{\nu+m\delta}, t_m)$. Designating the solution of the finite difference approximation on the new net by $w_\nu(t) \equiv w(\xi_\nu, t)$, we have $w_\nu(t_m) = v_{\nu+m\delta}(t_m)$. The basic scheme (1.3), transformed to the new coordinates, takes the form

$$w_\nu(t + \Delta t) = Rw_\nu(t), \qquad \nu \neq 0, \cdots, q, \tag{2.14a}$$

where

$$\begin{aligned} R \equiv QE^\delta &= \sum_{i=-r}^{p} A_i E^{i+\delta} \equiv \sum_{i=-r'}^{p'} B_i E^i, \qquad Ew_\nu = w_{\nu+1}, \\ r' &\equiv r - \delta, \qquad p' \equiv p + \delta, \qquad B_i \equiv A_{i-\delta}. \end{aligned} \tag{2.14b}$$

Scheme (2.14), which we call the *alternative basic scheme*, has all the properties of the basic scheme (1.3) except one. In particular, its amplification matrix is

$$\hat{R}(\zeta) = \hat{Q}(\zeta)e^{i\zeta\delta}, \tag{2.15}$$

and therefore, by (1.12) and (1.13), we see that scheme (2.14) is also strongly stable and dissipative in the sense of Kreiss. The only difference between the properties of (2.14) and (1.3) is that the alternative basic scheme is obviously not consistent with (1.1) but with the *alternative differential system*

$$\begin{aligned} u_t &= Bu_\xi, \qquad -\infty < \xi < \infty, t \geqq 0, \\ u(\xi, 0) &= f(\xi + \chi(0)), \\ B &= A + \left[\frac{d}{dt}\chi(t)\right]I = A + (\delta/\lambda)I, \end{aligned} \tag{2.16}$$

which is the image of the original system (1.1) under the transformation (2.12). B, like A, is a constant diagonal matrix.

Use of the transformation (2.12) means that, on the new net, the alternative basic scheme is applied (at *all* time steps) at $\nu \neq 0, \cdots, q$, and will be perturbed at $\nu = 0, \cdots, q$. The *alternative perturbation*, which is the image of (2.7) under the transformation (2.12), is

$$w_\nu(t + \Delta t) = \sum_{j=-k}^{s} C_{j,\nu}w_{\nu+j+\delta}(t) \equiv \sum_{j=-k'}^{s'} D_{j,\nu}w_{\nu+j}(t), \qquad \nu = 0, \cdots, q, \tag{2.17a}$$

where

$$D_{j,\nu} \equiv C_{j-\delta,\nu}, \qquad k' \equiv k - \delta, \qquad s' \equiv s + \delta. \tag{2.17b}$$

Now, the advantage of employing the transformation (2.2) has become clear: The results of the last section are immediately applicable to the *alternative perturbed scheme* (2.14) and (2.17).

Note that if $|\delta|$ is sufficiently large so that $p' = p + \delta \leqq 0$ or that $r' = r - \delta \leqq 0$, then the alternative basic scheme (2.14) becomes one-sided. In that case, even though the coefficient matrix A of (1.1) may have both positive and negative eigenvalues, Corollary 1 guarantees that the coefficient matrix $B = A + (\delta/\lambda)I$ of (2.16) satisfies $B \geqq 0$ or $B \leqq 0$, as δ is positive or negative. We are not able, at this point, to establish that B is always nonsingular, because A might have $-\delta/\lambda$ for an eigenvalue. Nevertheless, the results of Kreiss' Main Theorem are still valid for this case. This is so because the only place in the proof of Kreiss' theorem where use is made of the regularity of the coefficient matrix of the differential system is in Lemma 7 of [9]. There it is shown that, as $z \to 1$ ($|z| \geqq 1$), exactly n of the characteristic roots κ_j approach 1: l of them from inside the unit disc and $n - l$ from its exterior, l being the number of the negative eigenvalues in the coefficient matrix A. By using these particular n roots and their properties, the partition of the matrix M (defined by Eq. (2.4) of [9]) is accomplished. In the present case where the coefficient matrix B might be singular, it is seen from Eq. (3.3) of [9] that the number of roots κ_j tending to 1 as $z \to 1$ may exceed n. However, this is no longer of interest, due to the fact that if the finite difference scheme (2.14) is one-sided when $p' \leqq 0$ or $r' \leqq 0$, the roots of the characteristic equation belonging to this scheme are all either inside or outside the unit circle. Therefore, the partition of M, required in Lemma 7, is trivial; in fact, it is given by M itself.

At this point, we submit the conjecture that when $r' \leqq 0$ or $p' \leqq 0$, B is indeed regular. While we cannot prove this in the general case, we shall, in the following chapter, show it for a particular scheme, namely the Lax-Wendroff scheme.

From the transformation (2.12), it is clear that $v(t)$ and $w(t)$—the solutions of the original and alternative perturbed schemes—satisfy

$$(2.18) \qquad ||w(t)||_h = ||v(t)||_h$$

for all $t \geqq 0$. Therefore, we conclude that the original perturbed scheme is stable iff the alternative perturbed scheme is stable. Theorem 1 provides stability conditions concerning the indices of the alternative scheme r', p', k', s' and q. Thus, we are ready to state

THEOREM 2. *Let ν_0, δ and $q \geqq 0$ be given integers. Let the basic scheme* (1.3) *be perturbed at each time t_m by* (2.7) *with $\nu_m = m \cdot \delta + \nu_0$. If $r' \equiv r - \delta \leqq 0$ and $-k' \equiv \delta - k \geqq q$ ($p' \equiv p + \delta \leqq 0$ and $-s' \equiv -s - \delta \geqq q$), then the perturbed scheme is stable unless $-k' \equiv \delta - k = q = 0$ ($s' \equiv s + \delta = q = 0$). In the latter case, the scheme is stable if $\rho(C_{-k,0}) < 1$ ($\rho(C_{s,0}) < 1$) and is unstable if this spectral radius is greater than* 1.

Proof. When $r' = r - \delta \leqq 0$ and $-k \equiv \delta - k \geqq q$ (alternatively $p' = p + \delta \leqq 0$, $-s' \equiv s - \delta \geqq q$), all cases except $-k' \equiv q = 0$ are clear from Theorem 1. When $-k' = q = 0$, we conform to Theorem 1 by requiring $\rho(D_{0,0}) < 1$ for stability, and $\rho(D_{0,0}) > 1$ for instability. By (2.17), $D_{0,0} = C_{\delta,0} = C_{k,0}$ and Theorem 2 follows.

Since r, p, s, k and q are given, Theorem 2 insures stability if $|\delta|$ is sufficiently large. When $|\delta|$ is not large enough to satisfy one of the two alternative sets of conditions of the theorem, we have not been able to find stability conditions. However, in the particular case of various perturbations of the Lax-Wendroff scheme, the problem is completely solved as is shown in the next chapter.

Here we have dealt, in general terms, with perturbations moving at the speed $d\chi/dt = \delta\Delta x/\Delta t$ where we are free to fix Δt within the range permitted by the stability condition of the basic scheme, and δ is any integer. It might have been thought that we could have avoided carrying out the analysis for an infinite number of cases ($\delta = 0, \pm 1, \cdots$), by immediately transforming the original P.D.E. system (1.1), via (2.12), into (2.16), and then consider (2.16) as the given problem and construct a finite difference approximation to it, without further reference to the original problem. This would have meant that, from the beginning, we would have to consider only the case of a perturbation fixed at $x = 0$. All of this is quite true were we to restrict our attention only to linear systems with constant coefficients. However, the motivation for this work is to provide the (customary) linear stability analysis for nonlinear systems, in which case we have no a priori knowledge of the trajectories of internal boundaries such as shock-waves. The present approach which considers a spectrum of speeds of the internal boundaries allows us, as we shall see in Chapter 3, to find explicit stability criteria for the perturbed problem uniformly valid in δ.

3. Stability of Certain Perturbed Lax-Wendroff Schemes. In this chapter, the basic scheme which approximates the system (1.1) is the Lax-Wendroff (L-W) one [11], i.e.,

$$
\begin{aligned}
v_\nu^{m+1} &= Qv_\nu^m = \sum_{j=-1}^{1} A_j v_{\nu+j}^m, \qquad A_{\pm 1} = \tfrac{1}{2}(D^2 \pm D), \\
A_0 &= I - D^2, \qquad D \equiv \lambda A, \qquad \lambda = \Delta t/\Delta x,
\end{aligned} \tag{3.1}
$$

where $v_\nu{}^m \equiv v_\nu(t_m)$, and A is the coefficient matrix in (1.1). It is known (see for example [12, Chapter 12]) that the condition which we shall assume henceforth,

$$0 < \lambda\rho(A) < 1, \tag{3.2}$$

with $\rho(A)$ the spectral radius of A, assures the strong stability and dissipativity of (3.1). We now proceed to consider certain perturbations of (3.1).

3.1. *General Consistent and Diagonal Three-Point Perturbations Along the Trajectory of a Single Grid Point.* Let ν_0 and δ be given integers, as in the previous chapter. Consider a perturbation of (3.1), at each time level t_m, at the single grid point

$$\nu_m = m\delta + \nu_0, \tag{3.3}$$

given by a *three-point formula* of the form

$$v_{\nu_m}^{m+1} = \sum_{j=-1}^{1} C_{j,0} v_{\nu_m+j}^m. \tag{3.4}$$

This corresponds to (2.7) with $q = 0$. The perturbation considered as a scheme by itself is not necessarily consistent with (1.1); if it is, we call (3.4) a *consistent perturbation.* In the case of a consistent perturbation, it still does not necessarily follow that the $C_{j,0}$'s are diagonal, even though A is. When the matrices $C_{j,0}$ are diagonal, we say that we have a *diagonal perturbation.* Note[†] that diagonality does not necessarily imply that the $C_{j,0}$'s are polynomials in A. In practice, it turns out that if A is diagonal, then the perturbation (3.4) is usually chosen to be diagonal and consistent. We now prove

THEOREM 3. (a) *The L-W scheme* (3.1) *perturbed by the general three-point perturbation* (3.4) *is stable for all* $|\delta| \geqq 2$.

(b) $\rho(C_{\pm 1,0}) < 1$ *is a sufficient condition for stability when* $\delta = \pm 1$. *If the inequality is reversed, the perturbed scheme is unstable.*

(c) *If the perturbation* (3.4) *is consistent and diagonal, then the inequalities*

$$-\tfrac{1}{3}(1 + (4 - 3\lambda^2\,|a_j|^2)^{1/2}) < (c_{0,0})_j < 1 + \lambda\,|a_j|, \qquad 1 \leqq j \leqq n, \tag{3.5}$$

constitute a sufficient condition for stability when $\delta = 0$. *Here* a_j *and* $(c_{0,0})_j$ *are the eigenvalues of* A *and* $C_{0,0}$, *respectively. If, for any* $1 \leqq j \leqq n$, *one of the inequalities in* (3.5) *is reversed, the perturbed scheme is unstable.*

Proof. Since here $q = 0$ and $r = p = k = s = 1$, (a) and (b) follow immediately from Theorem 2. In this connection, we remark that for $|\delta| > 0$ the alternative scheme for (3.1) is one-sided. The dissipativity condition (3.2) implies that $|a_j| < 1/\lambda$ for all $1 \leqq j \leqq n$. Therefore, the eigenvalues of the coefficient matrix B of the alternative differential system (2.16), $b_j = a_j + \delta/\lambda$, are all positive or all negative ac-

[†] For example, approximate $u_t = Au_x$, $A \equiv \operatorname{diag}\{a, b\}$, by a finite difference operator $Q = \sum_{-1}^{1} C_j E^j$. If $C_{\pm 1} = \frac{1}{2}(\begin{smallmatrix}1\pm\lambda a & \lambda \\ \lambda & 1\pm\lambda b\end{smallmatrix} 3)$ and $C_0 = (\begin{smallmatrix}0 & -\lambda \\ -\lambda & 0\end{smallmatrix})$, then Q is consistent but not diagonal. If $C_{\pm 1} = \operatorname{diag}\{(1 \pm \lambda a - b)/2, (1 \pm \lambda b - a)/2\}$ and $C_0 = \operatorname{diag}\{b, a\}$, then Q is diagonal and consistent but the C_j's are not polynomials in A.

cording to the sign of δ, showing the conjecture on page 425 to be valid in this case.

Next, we set $\delta = 0$ and prove (c). Since, in this case, the perturbation is taken to be diagonal, the various components of the perturbed scheme are decoupled and we may assume without restrictions that $v_\nu{}^m$, $C_{j,0}$, A, and A_j are all scalars. Hence, the condition (3.2) takes the form

$$0 < \lambda\ |A| < .1 \quad \text{or equivalently} \quad 0 < |D| < 1\,. \tag{3.6}$$

In order to be able to use Kreiss' Main Theorem, we will show that z, with $|z| \geqq 1$, is not an eigenvalue (generalized eigenvalue if $z = 1$) of the appropriate operator G given in (1.8).

We start by presenting the characteristic equation (see (1.14)) of the L-W scheme, which is

$$z = A_{-1}/\kappa + A_0 + A_1\kappa, \tag{3.7}$$

the A_j's being given in (3.1). By Lemma 2 of [9], both roots of (3.7) satisfy, for $|z| \geqq 1$, $|\kappa_1| \leqq 1$ and $|\kappa_2| \geqq 1$.

Now we let z satisfy $|z| \geqq 1$ and seek an eigensolution $g \in H$ of G with z being its eigenvalue. From (1.17a) and (1.18a), we find that the most general solution in H of the (scalar) equation (1.15) is

$$g_\nu = \sigma_1\kappa_1^\nu, \quad \nu \geqq 0; \qquad g_\nu = \sigma_2\kappa_2^\nu\,, \quad \nu \leqq 0. \tag{3.8}$$

We see that g_0 has been defined twice, and the requirement of uniqueness implies

$$\sigma_1 = \sigma_2. \tag{3.9a}$$

In addition, g must also be an eigenfunction of the perturbation (3.4) (with $\nu_m = \nu_0 = 0$ for simplicity); therefore, (1.16) becomes

$$z\sigma_1 = C_{-1,0}\kappa_2^{-1}\sigma_2 + C_{0,0}\sigma_1 + C_{1,0}\kappa_1\sigma_1. \tag{3.9b}$$

Equations (3.9) define the homogeneous linear system $E(z)\sigma = 0$, $\sigma \equiv (\sigma_1, \sigma_2)'$, and we obtain

$$\det E(z) = C_{-1,0}\kappa_2^{-1} + C_{0,0} + C_{1,0}\kappa_1 - z. \tag{3.10}$$

Since we do not want G to have eigenvalues outside or on the unit circle, then, by Lemma 3 of [9] and (1.20), we have to look for conditions for which $\det E(z) \neq 0$, for all $|z| \geqq 1$. From (3.7), we have $\kappa_1\kappa_2 = A_{-1}/A_1 = (D-1)/(D+1)$ and, therefore,

$$\kappa_2^{-1} = [(D+1)/(D-1)]\kappa_1. \tag{3.11}$$

The consistency of the perturbation implies $\sum_{j=-1}^{1} C_{j,0} = 1$ and $\sum_{j=-1}^{1} jC_{j,0} = \lambda A \equiv D$, from which we obtain

$$C_{\pm1,0} = \tfrac{1}{2}(1 \pm D - C_{0,0}). \tag{3.12}$$

Substituting (3.11), (3.12), and (3.7) with $\kappa = \kappa_1$, into (3.10), we find that $\det E(z) = 0$ iff κ_1 satisfies

$$D(2C_{0,0} - 1 + D^2)\kappa_1^2 + 2(1-D)(C_{0,0} + D^2 - 1)\kappa_1 + D(1-D)^2 = 0. \tag{3.13}$$

Hence, we look for the range of $C_{0,0}$ for which (3.13) is contradicted for all $|z| \geqq 1$.

If the coefficient of κ_1^2 in (3.13) vanishes, i.e., if

$$C_{0,0} = \tfrac{1}{2}(1 - D^2), \tag{3.14}$$

then the only solution of (3.13) is $\kappa_1 = D/(D + 1)$. Using (3.7), we obtain $z = \frac{1}{2}$. Hence, in this case, det $E(z) \neq 0$ when $|z| \geqq 1$; therefore, from now on, we take

$$C_{0,0} \neq \tfrac{1}{2}(1 - D^2), \tag{3.15}$$

in which case (3.13) has the two solutions

(3.16)

$$\kappa_1 = [(1 - D)/D][1 - C_{0,0} - D^2 \pm ((1 - C_{0,0})^2 - D^2)^{1/2}]/[2C_{0,0} - 1 + D^2].$$

Substituting (3.16) into (3.7), we find that

$$z = C_{0,0}[C_{0,0} \pm ((1 - C_{0,0})^2 - D^2)^{1/2}]/[2C_{0,0} - 1 + D^2]. \tag{3.17}$$

Let $\Delta \equiv (1 - C_{0,0})^2 - D^2$ and assume first that $\Delta < 0$, i.e.,

$$1 - |D| < C_{0,0} < 1 + |D|. \tag{3.18}$$

In that case, z is complex and, from (3.17), we get

$$|z| = C_{0,0}/(2C_{0,0} - 1 + D^2)^{1/2}. \tag{3.19}$$

Since here $\Delta < 0$, it is easily checked that the R.H.S. of (3.19) is smaller than 1. Thus, whenever $C_{0,0}$ satisfies (3.18), det $E(z) \neq 0$ for $|z| \geqq 1$ and G has no eigenvalues outside the unit disc.

Next, we check out the case $\Delta \equiv 0$, i.e.,

$$C_{0,0} = 1 - |D| \tag{3.20a}$$

or

$$C_{0,0} = 1 + |D|. \tag{3.20b}$$

We first treat the case (3.20a) (for which it is easy to verify from (3.17) that $z = 1$). Here, (3.16) has a single solution κ_1. The corresponding value of κ_2 is given by (3.11), and we have

$$\kappa_1 = [|D|(1 - D)]/[D(1 - |D|)], \quad \kappa_2 = [D(|D| - 1)]/[|D|(D + 1)]. \tag{3.21}$$

By (3.6), we see that $|\kappa_2| < 1$ or $|\kappa_1| > 1$, when $D > 0$ or $D < 0$, respectively. By Lemma 2 of [9], this is a contradiction to the fact that $|\kappa_1| \leqq 1$ and $|\kappa_2| \geqq 1$ for all $|z| \geqq 1$. Thus again, det $E(z) \neq 0$ for $|z| \geqq 1$. The case of (3.20b) will be treated later.

There remains the case of $\Delta > 0$, where

$$C_{0,0} < 1 - |D| \tag{3.22a}$$

or

$$C_{0,0} > 1 + |D|. \tag{3.22b}$$

As in the case $\Delta \equiv 0$, we first consider only (3.22a). The κ_1 in (3.16) now has two real values. If we take the larger of the two $|\kappa_1|$ and obtain the corresponding κ_2

from (3.11), we get

(3.23)

$$\max\{|\kappa_1|, |\kappa_2|^{-1}\}$$
$$= (1 + |D|)(1 - C_{0,0} - D^2 + ((1 - C_{0,0})^2 - D^2)^{1/2})/[|D|\cdot|2C_{0,0} + 1 - D^2|]$$
$$> [(1 + |D|)(1 - C_{0,0} - D^2)]/[|D|\cdot|2C_{0,0} + 1 - D^2|].$$

Using (3.22a), a short calculation shows that the R.H.S. of (3.23) is greater than or equal to 1, and we have a contradiction of the same sort as that following (3.21). This does not complete the discussion in the (3.22a) case, since (3.16) has an additional value.

So, consider the smaller of the two $|\kappa_1|$ in (3.16). The corresponding value of z is obtained from (3.17) with the minus sign, i.e.,

$$z = C_{0,0}[C_{0,0} - ((1 - C_{0,0})^2 - D^2)^{1/2}]/[2C_{0,0} - 1 + D^2]. \tag{3.24a}$$

Recalling that, by (3.22a), $C_{0,0} < 1 - |D|$, we look for the range of $C_{0,0}$ such that z in (3.24a) will satisfy $|z| < 1$. When $C_{0,0} \geqq 0$, it is easy to verify directly that z in (3.24a) satisfies $0 \leqq z < 1$. When $C_{0,0} < 0$, (3.24a) implies

$$|z| = |C_{0,0}|\cdot[|C_{0,0}| + ((1 + |C_{0,0}|)^2 - D^2)^{1/2}]/[2\ |C_{0,0}| + 1 - D^2], \tag{3.24b}$$

and the requirement $|z| < 1$ leads to the condition $-(1 + (4 - 3D^2)^{1/2})/3 < C_{0,0} < 0$. Thus, the required condition on $C_{0,0}$ yielding $|z| < 1$ in the present case is

$$-\tfrac{1}{3}(1 + (4 - 3D^2)^{1/2}) < C_{0,0} < 1 - |D|. \tag{3.25}$$

By combining (3.18), (3.20a), and (3.25), we find that for $C_{0,0}$ such that

$$-\tfrac{1}{3}(1 + (4 - 3D^2)^{1/2}) < C_{0,0} < 1 + |D| \qquad (D = \lambda A), \tag{3.26}$$

$\det E(z) \neq 0$ for all $|z| \geqq 1$, and, by Kreiss' Theorem, the perturbed scheme is stable for these values of $C_{0,0}$.

This agrees with (3.5) in the scalar case. It is clear that when we have an $n \times n$ diagonal system, (3.26) would have to be satisfied component-wise and (3.5) follows.

In order to complete the proof of part (c) of the theorem, it remains to investigate the following cases:

$$C_{0,0} \geqq 1 + |D|, \qquad C_{0,0} \leqq -\tfrac{1}{3}(1 + (4 - 3D^2)^{1/2}). \tag{3.27}$$

For any value of $C_{0,0}$ obeying one of the strict inequalities in (3.27), we can show that there exists a z_0 with $|z_0| > 1$, such that $\det E(z_0) = 0$; hence, z_0 is an eigenvalue of G outside the unit disc and, by Lemma 1 of [9], the perturbed scheme is unstable. This completes the proof. We note that when $C_{0,0} = 1 + |D|$, then $\det E(1) = 0$, i.e., $z = 1$ is a generalized eigenvalue; and when $C_{0,0} = -\tfrac{1}{3}(1 + (4 - 3D^2)^{1/2})$, it can be shown that $\det E(-1) = 0$. Therefore, in these two special cases, G has eigenvalues on the unit disc and we cannot determine whether or not the scheme is stable.

Next, we illustrate the application of Theorem 3. Consider the scheme

$$v_\nu^{m+1} = [\theta_\nu^m L + (1 - \theta_\nu^m)Q]v_\nu^m, \tag{3.28}$$

where Q represents the L-W scheme (3.1) and $Lv_\nu{}^m$ is some constant coefficient finite difference approximation, consistent with (1.1), using the points ν, $\nu \pm 1$. It is clear that (3.28) itself is consistent with (1.1), provided $\{\theta_\nu{}^m\}$ is bounded. We now take

$$\text{(3.29)}\qquad \begin{aligned} \theta_\nu^m &= \theta = \text{constant}, & \nu &= \nu_m, \\ &= 0, & \nu &\neq \nu_m, \end{aligned} \quad \nu_m = m\delta + \nu_0 .$$

Note that here ν_m is that of (3.3). With this definition of $\theta_\nu{}^m$, (3.28) becomes a L-W scheme perturbed, at each time level t_m, at the single point $\nu = \nu_m$, the perturbation being given by

$$\text{(3.30)}\qquad v_{\nu_m}^{m+1} = [\theta L + (1 - \theta)Q]v_{\nu_m}^m .$$

As a first example, we take L to be the staggered Lax-Friedrichs (L-F) scheme [10]:

$$\text{(3.31a)}\qquad Lv_\nu^m = K_{-1}v_{\nu-1}^m + K_{+1}v_{\nu+1}^m, \qquad K_{\pm 1} = \tfrac{1}{2}(I \pm D), \qquad D = \lambda A .$$

This L is a first order accurate operator, stable under the condition $\lambda\rho(A) \leqq 1$. The perturbation (3.30) is now put in the form of (3.4) with

$$\text{(3.31b)}\qquad C_{\pm 1,0} = \tfrac{1}{2}(I \pm D)(\theta \pm D \mp \theta D), \qquad C_{0,0} = (1 - \theta)(I - D^2),$$

and we prove

COROLLARY 2. (a) *The Lax-Wendroff scheme perturbed by* (3.30), *with L being the L-F operator, is stable for all* $|\delta| \geqq 2$, *for all* θ.

(b) *For* $\delta = 1$ *or* $\delta = -1$, *the perturbed scheme is stable for every θ satisfying, respectively,*

$$\text{(3.32a)}\qquad 1 + \max_j[(\lambda a_j - 3)/(1 - \lambda^2 a_j^2)] < \theta < 1 + \left(1 - \lambda \cdot \min_j a_j\right)^{-1},$$

$$\text{(3.32b)}\qquad 1 - \min_j[(\lambda a_j - 3)/(1 - \lambda^2 a_j^2)] < \theta < 1 + \left(1 + \lambda \cdot \max_j a_j\right)^{-1}.$$

(c) *For* $\delta = 0$ *the perturbed scheme is stable for every θ satisfying*

$$\text{(3.33)}\qquad \lambda \cdot \eta(A)/(\lambda \cdot \eta(A) - 1) < \theta < 1 + [1 + (4 - 3\lambda^2\eta^2(A))^{1/2}]/[3 - 3\lambda^2\eta^2(A)],$$

where $\eta(A) \equiv \min_{1 \leqq j \leqq n} |a_j|$. *Specifically, the most stringent conditions obtained from* (3.32) *and* (3.33) *are* $-\frac{1}{2}(1 + \sqrt{8}) < \theta < \frac{3}{2}$ *and* $0 \leqq \theta \leqq 2$, *respectively. Hence, stability is obtained, uniformly in δ, for all* $0 \leqq \theta \leqq \frac{3}{2}$.

Proof. (a) follows immediately from Theorem 3. To prove (b) again via Theorem 3, all that is necessary to do is to determine the range of θ for which the inequalities $\rho(C_{\pm 1,0}) < 1$ are valid, where the $C_{\pm 1,0}$ are given by (3.31b). We see that θ must satisfy

$$\text{(3.34)}\qquad |\tfrac{1}{2}(1 \mp \lambda a_j)(\theta \mp \lambda a_j \pm \theta\lambda a_j)| < 1, \qquad 1 \leqq j \leqq n,$$

where the upper and lower signs correspond to $\delta = 1$ and $\delta = -1$, respectively. This implies that θ must satisfy, for all $1 \leqq j \leqq n$, the inequalities (3.32), taken without the extrema notation. Hence, the L.H.S.'s of (3.32) follow directly. From (3.2), we have $|\lambda a_j| < 1$ and the L.H.S.'s, taken as functions of the argument λa_j in the interval $(-1, 1)$, attain their maxima at the points $\lambda a_j = \pm(3 - \sqrt{8})$, where both take the value $-\frac{1}{2}(1 + \sqrt{8})$. The R.H.S.'s of (3.32) follow because $(1 \mp \lambda a_j)^{-1}$ is a monotone function of λa_j over the interval $(-1, 1)$, increasing or decreasing according to the minus or plus sign. The R.H.S.'s are always greater than $\frac{3}{2}$, and, therefore, for $|\delta| = 1$, the perturbed scheme is stable for all $-\frac{1}{2}(1 + \sqrt{8}) < \theta \leqq \frac{3}{2}$.

To prove (c), we have to check out the inequalities (3.5) of Theorem 3 where, by (3.31b), $(c_{0,0})_i = (1 - \theta)(1 - \lambda^2 |a_i|^2)$ in which, from (3.2), $\lambda|a_i|$ belongs to the interval (0, 1). This leads to a condition identical with (3.33), except that we have a_i rather than $\eta(A)$ for every $1 \leqq j \leqq n$. The right- and left-hand sides of this condition are, respectively, monotone increasing and monotone decreasing, as functions of $\lambda|a_i|$ over the interval (0, 1). Therefore, this condition is equivalent to (3.33). In particular, it is easily checked that the most stringent condition is $0 \leqq \theta \leqq 2$.

As a second example, we take for L of (3.30) the first order accurate, consistent, unconditionally *unstable* Euler scheme

$$(3.35) \qquad Lv_\nu^m = M_{-1}v_{\nu-1}^m + v_\nu^m + M_1 v_{\nu+1}^m, \qquad M_{\pm 1} = \pm D/2.$$

The resulting perturbation (3.35) is again put in the form of (3.4) with

$$(3.36) \qquad C_{\pm 1,0} = D(D \pm I - \theta D)/2, \qquad C_{0,0} = I + (\theta - 1)D^2.$$

Despite the fact that the Euler operator is unconditionally unstable, we have

COROLLARY 2′. (a) *The Lax-Wendroff scheme perturbed by* (3.30), *with L being the Euler operator, is stable for all* $|\delta| \geqq 2$, *for all* θ.

(b) *For* $|\delta| = 1$, *the perturbed scheme is stable for every* θ *satisfying*

$$(3.37) \qquad 1 \mp \min_i(\lambda a_i \pm 2)/(\lambda a_i)^2 < \theta < 1 + \min_i[(2 \mp \lambda a_i)/(\lambda a_i)^2],$$

where the upper and lower signs refer to $\delta = 1$ *and* $\delta = -1$, *respectively.*

(c) *For* $\delta = 0$, *the perturbed scheme is stable if*

$$(3.38) \qquad 1 - [4 + (4 - 3\lambda^2\rho^2(A))^{1/2}]/[3\lambda^2\rho^2(A)] < \theta < [1 + \lambda\rho(A)]/[\lambda\rho(A)].$$

Specifically, the most stringent conditions obtained from (3.37) *and* (3.38) *are, respectively,* $0 \leqq \theta \leqq 2$ *and* $-\frac{2}{3} \leqq \theta \leqq 2$. *Therefore, stability is assured, uniformly in* δ *for all* $0 \leqq \theta \leqq 2$.

The proof follows exactly the same lines of argument as that of Corollary 2. At this point, we note that the results of Ciment in [3], regarding perturbed schemes, constitute special cases of Corollary 2 with $\delta = 0$, $\theta = 1$, and of Corollary 2′ with $\delta = 0$ and $\theta = \frac{1}{2}, 1$.

3.2. *A Five-Point Perturbation Along the Trajectory of a Single Grid Point.* Consider the five-point approximation ([1], [2]) which is based on a single iteration of the L-W operator Q of (3.1),

$$(3.39) \qquad v_\nu^{m+1} = [I + (1 - \theta)(Q - I) + \theta(Q - I)Q]v_\nu^m, \qquad 0 \leqq \theta = \text{const} \leqq 1.$$

This scheme is conditionally stable [5] and is consistent with (1.1), but unlike the L-W scheme, is only first order accurate for $\theta > 0$; when $\theta = 0$, it coincides with L-W.

As before, let ν_0 and δ be given integers and let ν_m be defined by (3.3). Next, in (3.39) take $\theta \equiv \theta_\nu{}^m$, where $\theta_\nu{}^m$ is given by (3.29). This means that, for all $\nu \neq \nu_m$, the P.D.E. system (1.1) is approximated by the L-W scheme, while, at $\nu = \nu_m$, we have the perturbation (3.39) which, in accordance with (2.7), here takes the form

$$(3.40a) \qquad v_{\nu_m}^{m+1} = \sum_{i=-2}^{2} C_{i,0} v_{\nu_m+i}^m;$$

(3.40b) $$C_{\pm 2,0} = \theta D^2(1 \pm D)^2/4, \qquad C_{\pm 1,0} = D(D \pm 1)(\tfrac{1}{2} - \theta D^2),$$

$$C_{0,0} = I - D^2 + \theta D^2(3D^2 - 1)/2, \qquad D = \lambda A, \qquad 0 \leqq \theta \leqq 1.$$

We now prove

THEOREM 4. *The L-W scheme perturbed by the iterated approximation* (3.40) *is stable for all* δ.

Proof. In agreement with our notation of Chapter 1, we have that the indices of the perturbed scheme are $r = p = 1$, $k = s = 2$ and $q = 0$; so, by Theorem 2, we have stability for all δ with $|\delta| \geqq 3$. When $\delta = \mp 2$, then, according to Theorem 2, the scheme is stable, provided that $\rho(C_{\pm 2,0}) < 1$ where $C_{\pm 2,0}$ is given in (3.40b). Since $0 \leqq \theta \leqq 1$ and, by (3.2), $\rho(D) < 1$, this condition is clearly met.

When $\delta = 1$, the alternative basic scheme is

(3.41a) $$w_\nu^{m+1} = A_{-1} w_\nu^m + A_0 w_{\nu+1}^m + A_1 w_{\nu+2}^m,$$

where the A_j's are the L-W coefficients given in (3.1). The alternative perturbation to (3.40a) is (with $\nu_0 = 0$, for simplicity)

(3.41b) $$w_0^{m+1} = \sum_{j=-1}^{3} C_{j-1,0} w_j^m ,$$

the $C_{l,0}$'s given by (3.40b). Since we are dealing with a diagonal finite difference approximation, we may assume, without restrictions, that it is scalar. Hence, the characteristic equation corresponding to (3.41a) is

(3.42) $$\tfrac{1}{2}(D^2 + D)\kappa^2 + (1 - D^2)\kappa + \tfrac{1}{2}(D^2 - D) - z = 0 \qquad (D = \lambda A),$$

and, by Lemma 2 of [9], the two roots κ_1 and κ_2 are not inside the unit disc for all z with $|z| \geqq 1$. Here, the indices of the alternative scheme are $r' = q = 0$, $k' = 1$, $p' = 2$ and $s' = 3$. Therefore, if $g \in H$ is an eigensolution of the appropriate operator G (that represents the perturbed scheme), with an eigenvalue $|z| \geqq 1$, then, by (1.17b) and (1.18a), we have $g_\nu = 0$ for all $\nu \geqq 1$, and $g_\nu = \sum P_j \kappa_j{}^\nu$ for $\nu \leqq 1$.

Assume first that $\kappa_1 = \kappa_2$. From (3.42), this happens when

(3.43a) $$z = z_0 \equiv (D - 1)/(2D),$$

from which

(3.43b) $$\kappa_1 = \kappa_2 = (D - 1)/D.$$

The case $|z_0| < 1$ is of no interest for us and since, by (3.6), $|z_0| \neq 1$, we assume that $|z_0| > 1$. In this case, g takes the form

(3.44) $$g_\nu = 0, \quad \nu \geqq 1; \qquad g_\nu = \sigma''_1 \kappa_1^\nu + \sigma''_2 \nu \kappa_2^\nu, \quad \nu \leqq 1,$$

where σ_1 and σ_2 are two parameters yet to be defined. The uniqueness requirement at $\nu = 1$ gives

(3.45a) $$\kappa_1 \sigma_1 + \kappa_2 \sigma_2 = 0.$$

Now, g must be an eigensolution of the internal boundary conditions (3.41b) as well, namely

$$z_0 g_0 = \sum_{i=-1}^{3} C_{i-1,0} g_i. \tag{3.46}$$

By substituting the $g_{-1}, \cdots, g_3$ from (3.44), (3.46) takes the form

$$(\kappa_1^{-1} C_{-2,0} + C_{-1,0} - z_0)\sigma_1 - \kappa_1^{-1} C_{-2,0}\sigma_2 = 0. \tag{3.45b}$$

Equations (3.45) define the system $E(z_0)\sigma = 0$. Substituting z_0 and κ_i from (3.43) and $C_{-2,0}$, $C_{-1,0}$ from (3.40b), we obtain

$$\det E(z_0) = (1 - D)^2(1 - D^2 + \theta D^4)/(2D^2), \tag{3.47}$$

which does not vanish for $0 < |D| < 1$ and $0 \leqq \theta \leqq 1$. Thus, if z_0 in (3.43a) satisfies $|z_0| > 1$, then, by Lemma 3 of [9], it is not an eigenvalue of G.

Next, assume that the roots of (3.42) satisfy $\kappa_1 \neq \kappa_2$ and that $|z| \geqq 1$. Then

$$g_\nu = 0, \quad \nu \geqq 1; \qquad g_\nu = \sigma_1 \kappa_1^\nu + \sigma_2 \kappa_2^\nu, \quad \nu \leqq 1. \tag{3.48}$$

The uniqueness requirement at $\nu = 1$ again gives (3.45a), while (3.46) with $g_{-1}, \cdots, g_3$ from (3.48) leads to

$$(\kappa_1^{-1} C_{-2,0} + C_{-1,0} - z)\sigma_1 + (\kappa_2^{-1} C_{-2,0} + C_{-1,0} - z)\sigma_2 = 0. \tag{3.49}$$

Equations (3.45a) and (3.49) define $E(z)\sigma = 0$, where

$$\det E(z) = [(\kappa_1 + \kappa_2)C_{-2,0} + \kappa_1\kappa_2(C_{-1,0} - z)](\kappa_1 - \kappa_2)/(\kappa_1\kappa_2). \tag{3.50}$$

The determinant is zero iff the expression in the square brackets vanishes. Computing $\kappa_1 + \kappa_2$ and $\kappa_1\kappa_2$ from the characteristic equation (3.42) and using $C_{-2,0}$ and $C_{-1,0}$ from (3.40b), we find that $\det E(z) = 0$ iff

$$\tilde{E}(z) \equiv 4z^2 + 4D(1 - D)(1 - \theta D^2)z + D^2(1 - D)^2(1 - \theta - \theta D^2) = 0. \tag{3.51}$$

We shall show that z_1 and z_2, the roots of (3.51), are inside the unit disc; hence, G does not have eigenvalues z with $|z| \geqq 1$, thereby proving stability.

It is readily shown that

$$|z_{1,2}| \leqq \tfrac{1}{2}\, |D|(1 + |D|)[1 - \theta D^2 + (\theta - \theta D^2 + \theta^2 D^4)^{1/2}]. \tag{3.52a}$$

Thus, to show that $|z_{1,2}| < 1$, it is sufficient to verify that

$$\Lambda(\Lambda + 1)(\theta - \theta\Lambda^2 + \theta^2\Lambda^4)^{1/2} < 2 + \Lambda(\theta\Lambda + 1)(\theta\Lambda^2 - 1), \qquad \Lambda \equiv |D|, \tag{3.52b}$$

where, by (3.6), $0 < \Lambda < 1$. Squaring both sides leads to the requirement

$$\psi(\theta, \Lambda) \equiv (1 - \Lambda)[4 - \Lambda^2(3 + \Lambda)] - \theta\Lambda^2(1 - \Lambda^2)(1 - 2\Lambda - \Lambda^2) > 0, \quad 0 \leqq \theta \leqq 1. \tag{3.53}$$

The expression for $\psi(\theta, \Lambda)$ is linear in θ and we find that $\psi(0, \Lambda)$ and $\psi(1, \Lambda)$ are positive for all $0 < \Lambda < 1$. Hence, the requirement (3.53) is met.

For $\delta = -1$, the proof is almost identical.

It remains to consider the case $\delta = 0$. Again we take, without restrictions, the scalar case. The basic alternative scheme reverts back to the L-W one and, following the proof of part (c) of Theorem 3, we obtain the characteristic equation (3.7), the relation (3.11), and, for $|z| \geqq 1$, g is given by (3.8) with (3.9a). Again, requiring that

g be an eigensolution of the internal boundary conditions as well (Eq. (3.40a) with $\nu_m = \delta m + \nu_0 = 0$ for simplicity), we obtain

$$zg_\nu = \sum_{j=-2}^{2} C_{j,0} g_j, \tag{3.54a}$$

which, after using the values of g_j from (3.8), takes the form

$$(C_{0,0} - z + C_{1,0}\kappa_1 + C_{2,0}\kappa_1^2)\sigma_1 + (C_{-2,0}\kappa_2^{-2} + C_{-1,0}\kappa_2^{-1})\sigma_2 = 0. \tag{3.54b}$$

This together with (3.9a) defines

$$\det E(z) = C_{-2,0}\kappa_2^{-2} + C_{-1,0}\kappa_2^{-1} + C_{0,0} + C_{1,0}\kappa_1 + C_{2,0}\kappa_1^2 - z. \tag{3.55}$$

Since $|z| \geqq 1$, then a necessary condition for z to be an eigenvalue of G is that $\det E(z) = 0$. We shall show that this condition is not fulfilled and we first consider the case of $0 < D < 1$.

Assume that $\det E(z) = 0$, and in (3.55) substitute κ_1 from (3.11), z from (3.7) with $\kappa = \kappa_2$, and $C_{j,0}$ from (3.40b). We obtain the following equation for $\tau \equiv \kappa_2^{-1}$:

$$\begin{aligned} \phi(\tau) \equiv \theta D(1 - D)^2\tau^3 + (1 - D)(4\theta D^2 - 1)\tau^2 \\ + \theta D(3D^2 - 1)\tau - (D + 1) = 0. \end{aligned} \tag{3.56}$$

By Lemma 2 of [9], we have that $|\tau| \equiv |\kappa_2^{-1}| \leqq 1$ for $|z| \geqq 1$. However, we shall show that (3.56) has no roots τ with $|\tau| \leqq 1$ when $0 < D < 1$, hence, contradicting the assumption of $\det E(z) = 0$. The product of the roots of (3.56), say Ω, satisfies

$$\Omega = (D + 1)/[\theta D(1 - D)^2] > 1. \tag{3.57a}$$

We also have

$$\phi(1) = 2(\theta D^2 - 1) < 0, \qquad \phi(\tilde{\tau}) = 2(2 + 2\theta D^2 - \theta)/[\theta(1 - D)] > 0, \tag{3.57b}$$

where

$$1 < \tilde{\tau} \equiv 1/[\theta D(1 - D)] = \Omega(1 - D)/(1 + D) < \Omega. \tag{3.57c}$$

This shows that (3.56) has a root τ_1 with

$$1 < \tau_1 < \Omega. \tag{3.57d}$$

Equation (3.56) does not have zeros in the interval [0, 1], since, for all τ with $0 \leqq \tau \leqq 1$, we have

$$\begin{aligned} \phi(\tau) &\leqq \theta D(1 - D)^2\tau^3 + 4\theta D^2(1 - D)\tau^2 + \theta D(3D^2 - 1)\tau - (1 + D) \\ &\leqq \theta D(1 - D^2)\tau + 4\theta D^2(1 - D)\tau + \theta D(3D^2 - 1)\tau - 1 - D \\ &= 2\theta D^2\tau - D - 1 < 0. \end{aligned} \tag{3.58a}$$

Equation (3.56) does not have zeros in the interval $[-1, 0]$ either, because, for all τ with $-1 \leqq \tau \leqq 0$,

$$\begin{aligned} \phi(\tau) &\leqq 4\theta D^2(1 - D) + \theta D - D - 1 \leqq 4\theta D^2(1 - D) - 1 \\ &\leqq 4\theta D \cdot \max_{0 \leqq D \leqq 1} [D(1 - D)] - 1 = \theta D - 1 < 0. \end{aligned} \tag{3.58b}$$

It follows, therefore, that the two remaining roots of (3.56) are either real and outside the interval $[-1, 1]$, or complex conjugate and, from (3.57d), we get $\tau_2\tau_3 = \Omega/\tau_1 > 1$, as was to be shown.

Next, we consider the case $-1 < D < 0$ and again assume that det $E(z) = 0$. In (3.55) substituting κ_2^{-1} from (3.11), z from (3.7) with $\kappa = \kappa_1$, and $C_{j,0}$ from (3.40b), we find that κ_1 must satisfy

$$\begin{aligned} -\theta D(1 + D)^2\kappa_1^3 + (1 + D)(4\theta D^2 - 1)\kappa_1^2 \\ - \theta D(3D^2 - 1)\kappa_1 + D - 1 = 0. \end{aligned} \tag{3.59}$$

Note that (3.59) yields (3.56) when we replace D by $-D$. Hence, by the previous argnment, (3.59) does not have roots κ_1 with $|\kappa_1| \leqq 1$, when $|z| \geqq 1$. It follows that det $E(z) \neq 1$ for all $|z| \geqq 1$, and, by Kreiss' Main Theorem, stability is assured.

4. Certain Divergenceless Perturbed Lax-Wendroff Schemes.

In practice, we often consider schemes which are in *conservation-form* [12, Chapter 12]. Such schemes are also called *divergence-free* or *divergenceless*; by this we mean that, for any given range $N_1 \leqq \nu \leqq N_2$, the solution $v_\nu{}^m$ satisfies

$$\sum_{\nu=N_1}^{N_2} v_\nu^{m+1} - \sum_{\nu=N_1}^{N_2} v_\nu^m = B(N_1, N_2, m). \tag{4.1}$$

Here, $B(N_1, N_2, m)$ consists of those terms that are the contribution of the external boundaries $N_1\Delta x$ and $N_2\Delta x$ only.

The advantage of conservation law schemes lies in the fact that they approximate weak solutions more correctly than nondivergenceless algorithms. For example, they are known to predict the correct propagation-speed of discontinuities in the solution, such as shock-waves. This is of particular importance in the nonlinear problems; however, the stability analysis of the corresponding finite difference approximation is usually carried out for the linear version. In this chapter, we shall construct certain divergence-free perturbed L-W schemes and prove their stability.

It should be noted that when divergence-free basic schemes are perturbed at a single grid point in the manner discussed in the previous chapter, the resulting perturbed schemes are *never* in conservation form, even though the perturbation itself might be divergenceless. For example, if we take a perturbed scheme of the form given in (3.28) together with (3.29), we obtain

$$\sum_{\nu=N_1}^{N_2} v_\nu^{m+1} - \sum_{\nu=N_1}^{N_2} v_\nu^m = \theta(L - Q)v_{\nu_m}^m + B(N_1, N_2, m), \tag{4.2}$$

thus showing that the conservation-form is lost unless we have the trivial case of $\theta = 0$ or $L \equiv Q$.

We would like now to construct and analyze divergenceless versions of the perturbed schemes considered in Corollaries 2 and 2′, and in Theorem 4 of Chapter 3. We shall see that the analysis of the divergence-free approximations is closely related in each case to the analysis of the analogous case discussed in the previous chapter.

We start with the case considered in Corollary 2, namely with the perturbed combination, described in (3.28), of the L-W scheme (3.1) and the L-F approximation (3.31a). A divergence-free version of this perturbed scheme is

$$\begin{aligned} v_\nu^{m+1} &= \theta_{\nu-1/2}^m K_{-1} v_{\nu-1}^m + \theta_{\nu+1/2} K_{+1} v_{\nu+1}^m \\ &\quad + (1 - \theta_{\nu-1/2}^m)(A_{-1} v_{\nu-1}^m + \tfrac{1}{2} A_0 v_\nu^m) + (1 - \theta_{\nu+1/2}^m)(\tfrac{1}{2} A_0 v_\nu^m + A_1 v_{\nu+1}^m), \end{aligned} \tag{4.3}$$

where the A_j's and K_j's are defined in (3.1) and (3.31a), respectively. It is easy to see that this scheme is indeed in conservation form for any definition of $\theta_{\nu+1/2}{}^m$.

As before, let δ and ν_0 be given integers and take

$$\begin{aligned} \theta_{\nu+1/2}^m &= \theta = \text{constant}, & \nu &= \nu_m, \\ &= 0, & \nu &\neq \nu_m, \end{aligned} \qquad \nu_m = m\delta + \nu_0. \tag{4.4}$$

We see that by this definition the basic L-W scheme is perturbed at each time step, at only two grid points, ν_m and $\nu_m + 1$. Obviously, this is the minimal number of grid points at which a perturbation can be applied so that the achieved perturbed scheme will be divergence-free. We write the perturbation in the form of (2.7):

$$v_{\nu_m}^{m+1} = \sum_{j=-1}^{1} C_{j,0} v_{\nu_m+j}^m, \qquad v_{\nu_m+1}^{m+1} = \sum_{j=-1}^{1} C_{j,1} v_{\nu_m+1+j}^m, \tag{4.5a}$$

where

$$\begin{aligned} C_{-1,0} &= (D^2 - D)/2, \qquad C_{0,0} = C_{0,1} = (1 - \theta/2)(I - D^2), \\ C_{1,1} &= (D^2 + D)/2, \qquad C_{1,0} = (I + D)(\theta + D - \theta D)/2, \\ C_{-1,1} &= (I - D)(\theta - D + \theta D)/2 \qquad (D = \lambda A). \end{aligned} \tag{4.5b}$$

Note that each of the internal boundary conditions in (4.5) is *inconsistent* with (1.1). It seems that this is the price one pays for obtaining a divergenceless perturbed scheme. We now prove

THEOREM 5. *The L-W approximation perturbed by* (4.5) *is stable under the conditions of Corollary* 2.

Proof. The indices of the perturbed scheme satisfy $r = p = s = k = q = 1$, hence, by Theorem 2, the scheme is stable for $|\delta| \geqq 2$.

When $\delta = 1$, the alternative basic scheme is given by (3.46a) and the alternative perturbation is

$$w_0^{m+1} = \sum_{j=0}^{2} C_{j-1,0} w_j^m, \qquad w_1^{m+1} = \sum_{j=0}^{2} C_{j-1,1} w_{j+1}^m, \tag{4.6}$$

with the $C_{j,l}$'s given by (4.5b). The corresponding characteristic equation is (3.42) and, for all z with $|z| \geqq 1$, its roots κ_1 and κ_2 satisfy $|\kappa_1| \geqq 1$, $|\kappa_2| \geqq 1$. The indices of this alternative perturbed scheme are $r' = k' = 0$, $p' = s' = 2$ and $q = 1$. Therefore, by (1.17b) and (1.18a), g—an eigensolution of G with an eigenvalue $|z| \geqq 1$—must satisfy $g_\nu = 0$ for all $\nu \geqq 2$ and $g_\nu = \sum P_j \kappa_j{}^\nu$ for $\nu \leqq 1$. In addition, g must be an eigensolution of (4.6) and, since $g_2 = g_3 = 0$, we get

$$zg_0 = C_{-1,0} g_0 + C_{0,0} g_1, \qquad zg_1 = C_{-1,1} g_1. \tag{4.7}$$

Assume at first $\kappa_1 = \kappa_2$; then

$$g_\nu = 0, \quad \nu \geqq 2; \qquad g_\nu = \sigma_1 \kappa_1^\nu + \sigma_2 \nu \kappa_1^\nu, \quad \nu \leqq 1. \tag{4.8a}$$

Substituting these values of g_ν into (4.7), we obtain the system $E(z)\sigma = 0$, where

$$\det E(z) = \kappa_1(C_{-1,0} - z)(C_{-1,1} - z). \tag{4.9a}$$

Therefore, $\det E(z) \neq 0$ for all $|z| \geqq 1$ iff

$$|C_{-1,0}| < 1 \quad \text{and} \quad |C_{-1,1}| < 1. \tag{4.10}$$

Next, take the case $\kappa_1 \neq \kappa_2$. We have

$$g_\nu = 0, \quad \nu \geqq 2; \qquad g_\nu = \sigma_1\kappa_1^\nu + \sigma_2\kappa_2^\nu, \quad \nu \leqq 1; \tag{4.8b}$$

and the same procedure as above gives

$$\det E(z) = (\kappa_2 - \kappa_1)(C_{-1,0} - z)(C_{-1,1} - z). \tag{4.9b}$$

It follows that a sufficient condition for stability is again (4.10). Clearly, in the $n \times n$ case, this generalizes to

$$\rho(C_{-1,0}) < 1 \quad \text{and} \quad \rho(C_{-1,1}) < 1, \tag{4.11}$$

where we recall that $C_{-1,0}$ and $C_{-1,1}$ are given in (4.5b).

By (3.2), we do have $\rho(C_{-1,0}) < 1$. Now notice that $C_{-1,1}$ in (4.5b) is identical with $C_{-1,0}$ in (3.31b), therefore, using part (b) of the proof of Corollary 2, we also obtain $\rho(C_{-1,1}) < 1$, provided that θ satisfies (3.32a). Similarly, for $\delta = -1$, it follows that the scheme is stable if θ satisfies (3.32b).

Lastly, we consider the case $\delta = 0$. Then the alternative basic scheme reverts back to the L-W approximation, and the alternative perturbation is given by (4.5) with $\nu_m = 0$. The characteristic equation is (3.7) and, for $|z| \geqq 1$, its two roots satisfy $|\kappa_1| \leqq 1$ and $|\kappa_2| \geqq 1$. We also have the relation (3.11). Since $r = p = q = 1$, (1.17a) and (1.18a) yield (for $|z| \geqq 1$)

$$g_\nu = \sigma_1\kappa_1^\nu, \quad \nu \geqq 1; \qquad g_\nu = \sigma_2\kappa_2^\nu, \quad \nu \leqq 0. \tag{4.12}$$

Again, requiring that g also be an eigensolution of the internal boundary conditions, we find that

$$zg_0 = \sum_{j=-1}^{1} C_{j,0}g_j, \qquad zg_1 = \sum_{j=-1}^{1} C_{j,1}g_{j+1}. \tag{4.13}$$

Substituting the values of g_ν given by (4.12) into (4.13), we arrive at a pair of linear equations for which

$$\det E(z) = \kappa_1[C_{1,0}C_{-1,1} - (\kappa_2^{-1}C_{-1,0} + C_{0,0} - z)(C_{0,1} + \kappa_1 C_{1,1} - z)]. \tag{4.14a}$$

Once again, we recall that a necessary condition for z with $|z| \geqq 1$ to be an eigenvalue of G is that $\det E(z) = 0$. Substituting for z from (3.7) with $\kappa = \kappa_1$, for κ_2^{-1} from (3.11), and for $C_{j,l}$ from (4.5b), we conclude that $\det E(z) = 0$ iff

$$(D - 1)[D(1 + D)(2\theta - 1)\kappa_1^2 + 2\theta(1 - D^2)\kappa_1 - D(1 - D)] = 0. \tag{4.14b}$$

Using $C \equiv (1 - \theta)(1 - D^2)$ to eliminate θ from (4.14b), we find that κ_1 must satisfy (3.13), with $C_{0,0}$ replaced by C. Therefore, by the argument following (3.13), stability is assured if c_j, the eigenvalues of C, satisfy (see Eq. (3.5))

$$-\tfrac{1}{3}(1 + (4 - 3\lambda^2\,|a_j|^2)^{1/2}) < c_j \equiv (1 - \theta)(1 - \lambda^2 a_j^2) < 1 + \lambda\,|a_j|, \tag{4.15}$$
$$1 \leqq j \leqq n.$$

These inequalities were dealt with in proving part (c) of Corollary 2, and the resulting restrictions on θ, needed for stability, are given by (3.33). This completes the proof of Theorem 5.

As a second illustration of a divergence-free perturbed scheme, we take the case considered in Corollary 2′, namely, we perturb the L-W operator by the Euler scheme (3.35) in the manner described by (3.28). A divergence-free version of this perturbed scheme is

$$(4.16)\quad \begin{aligned} v_\nu^{m+1} = {} & \theta^m_{\nu-1/2}(M_{-1}v^m_{\nu-1} + \tfrac{1}{2}v^m_\nu) + \theta^m_{\nu+1/2}(\tfrac{1}{2}v^m_\nu + M_1 v^m_{\nu+1}) \\ & + (1 - \theta^m_{\nu-1/2})(A_{-1}v^m_{\nu-1} + \tfrac{1}{2}A_0 v^m_\nu) + (1 - \theta^m_{\nu+1/2})(\tfrac{1}{2}A_0 v^m_\nu + A_1 v^m_{\nu+1}), \end{aligned}$$

where the A_j's and the M_j's are defined in (3.1) and in (3.35), respectively. Define $\theta_{\nu+1/2}{}^m$ as in (4.7). Then again the L-W scheme is perturbed at each time step at the two grid points ν_m and $\nu_m + 1$. The perturbation is in the form of (4.5a), where

$$(4.17)\quad \begin{aligned} & C_{-1,0} = (D^2 - D)/2, \qquad C_{0,0} = C_{0,1} = I - D^2 + \theta D^2/2, \\ & C_{1,1} = (D^2 + D)/2, \qquad C_{1,0} = D(D + I - \theta D^2)/2, \\ & C_{-1,1} = D(D - I - \theta D)/2. \end{aligned}$$

We now prove

THEOREM 5′. *The L-W approximation perturbed by* (4.5a), *with its coefficients given by* (4.17), *is stable under the conditions of Corollary* 2′.

Proof. The proof for the case $|\delta| \geqq 2$ is the same as in Theorem 5. When $\delta = 1$, an argument identical to the one used in Theorem 5 leads us to the stability condition (4.11), in which $C_{-1,0}$ and $C_{-1,1}$ are now defined by (4.17). By (3.2), it is clear that $\rho(C_{-1,0}) < 1$. Since $C_{-1,1}$ in (4.17) is identical with $C_{-1,0}$ in (3.36), the proof of part (b) of Corollary 2′ shows that $\rho(C_{-1,1}) < 1$, provided θ satisfies (3.37) with the upper signs. Similarly, when $\delta = -1$, we find that the stability condition is given by (3.37) with the lower signs.

When $\delta = 0$, we follow the argument in the proof of Theorem 5 and arrive at (4.14a). As in Theorem 5, we substitute in (3.17a) for z from (3.7) with $\kappa = \kappa_1$, for $\kappa_2{}^{-1}$ from (3.11), but for $C_{j,l}$ from (4.17). We obtain that det $E(z) = 0$ iff

$$(4.18)\quad D(2\theta D^2 + 1 - D^2)\kappa_1^2 + 2\theta D^2(1 - D)\kappa_1 + D(1 - D)^2 = 0.$$

Using $C' \equiv 1 + (\theta - 1)D^2$ to eliminate θ from (4.18), we see that κ_1 must satisfy (3.13), with $C_{0,0}$ being replaced by C'. Again, by the argument following (3.13), stability is assured if (see Eq. (3.5)) the eigenvalues of C', say c_j', satisfy (4.15) with $c_j' \equiv 1 + (\theta - 1)\lambda^2 a_j{}^2$ replacing c_j. Following the proof of part (c) of Corollary 2′, the restrictions on θ, needed for stability, are given by (3.38) and the proof of the theorem is completed.

As a last example of a divergence-free perturbed scheme, consider the iterated L-W approximation (3.39) with $\theta \equiv \theta_\nu{}^m$ defined in (3.29). Let $Q = \sum A_j E^j$ be the L-W operator defined in (3.1), then, by the consistency condition $\sum A_j = I$, the operator $Q - I$ is found to be

$$(4.19)\quad (Q - I)v^m_\nu = (A_{-1}v^m_{\nu-1} - A_1 v^m_\nu) + (A_1 v^m_{\nu+1} - A_{-1}v^m_\nu).$$

With this we can build the following divergence-free form of (3.39):

$$\begin{aligned}(4.20)\quad v_\nu^{m+1} &= v_\nu^m + (1-\theta_{\nu-1/2}^m)(A_{-1}v_{\nu-1}^m - A_1v_\nu^m) + (1-\theta_{\nu+1/2}^m)(A_1v_{\nu+1}^m - A_{-1}v_\nu^m)\\ &\quad + \theta_{\nu-1/2}^m(A_{-1}Qv_{\nu-1}^m - A_1Qv_\nu^m) + \theta_{\nu+1/2}^m(A_1Qv_{\nu+1}^m - A_{-1}Qv_\nu^m).\end{aligned}$$

This scheme is obviously in conservation form for any choice of $\theta_{\nu+1/2}{}^m$. Define $\theta_{\nu+1/2}{}^m$ as in (4.4), with $0 \leqq \theta \leqq 1$ as in (3.39); then (4.20) describes a L-W scheme perturbed at the two grid points ν_m and $\nu_m + 1$. Notice that, unlike the two previous examples, we now have a five-point perturbation which may be written in accordance with (2.7) in the form

$$(4.21a)\qquad v_{\nu_m}^{m+1} = \sum_{j=-2}^{2} C_{j,0}v_{\nu_m+j}^m,\qquad v_{\nu_m+1}^{m+1} = \sum_{j=-2}^{2} C_{j,1}v_{\nu_m+1+j}^m.$$

Defining $J \equiv \theta D/2$, it can be verified that

$$\begin{aligned} &C_{-2,0} = C_{2,1} = 0, \qquad C_{-2,1} = (D-1)JA_{-1},\\ &C_{-1,0} = (I + J - DJ)A_{-1},\\ &C_{-1,1} = A_{-1} + JA_1 - DJ(2A_{-1} + A_1),\\ (4.21b)\qquad &C_{0,0} = A_0 - JA_1 + DJ(2A_{-1} + A_1),\\ &C_{0,1} = A_0 + JA_{-1} + DJ(2A_1 + A_{-1}),\\ &C_{1,0} = A_1 - JA_{-1} - DJ(2A_1 + A_{-1}),\\ &C_{1,1} = (I - J - DJ)A_1, \qquad C_{2,0} = (D+1)JA_1.\end{aligned}$$

Here, the A_j's are given in (3.1). We now prove

THEOREM 6. *The L-W scheme, perturbed by* (4.21), *is stable for all* δ.

Proof. The indices of the perturbed scheme are $r = p = q = 1$ and $k = s = 2$; hence, by Theorem 2, the scheme is stable for all $|\delta| \geqq 3$.

For $\delta = 2$, the alternative basic scheme is

$$(4.22)\qquad w_\nu^{m+1} = A_{-1}w_{\nu+1}^m + A_0w_{\nu+2}^m + A_1w_{\nu+3}^m,$$

where the A_j's are defined in (3.1). The perturbation alternative to (4.21a) is

$$(4.23)\qquad w_0^{m+1} = \sum_{j=0}^{4} C_{j-2,0}w_j^m,\qquad w_1^{m+1} = \sum_{j=0}^{4} C_{j-2,1}w_{j+1}^m,$$

with the $C_{j,l}$'s given by (4.21b),

Again, for simplicity, we consider the scalar case and find that the characteristic equation corresponding to (4.22) is

$$(4.24)\qquad A_1\kappa^3 + A_0\kappa^2 + A_{-1}\kappa - z = 0.$$

By Lemma 2 of [9], for all z with $|z| \geqq 1$, each of the three roots κ_j of (4.24) satisfies $|\kappa_j| \geqq 1$. In the alternative perturbed scheme, $r' = -1$, $p' = 3$, $k' = 0$, $s' = 4$, and $q = 1$. Therefore by (1.17b) and (1.18a), we see that for $g \in H$ to be an eigensolution of G, with eigenvalue z, $|z| \geqq 1$, it must satisfy

$$(4.25)\qquad g_\nu = 0,\quad \nu \geqq 2;\qquad g_\nu = \sum P_j\kappa_j^\nu,\quad \nu \leqq 2.$$

For $\nu \leqq 2$, there are three possible solutions:

$$
\begin{aligned}
g_\nu &= \sigma_1\kappa_1^\nu + \sigma_2\nu\kappa_1^\nu + \sigma_3\nu^2\kappa_1^\nu, && \text{if } \kappa_1 = \kappa_2 = \kappa_3,\\
&= \sigma_1\kappa_1^\nu + \sigma_2\nu\kappa_1^\nu + \sigma_3\kappa_3^\nu, && \text{if } \kappa_1 = \kappa_2 \neq \kappa_3,\\
&= \sigma_1\kappa_1^\nu + \sigma_2\kappa_2^\nu + \sigma_3\kappa_3^\nu, && \text{if } \kappa_i \neq \kappa_j \text{ for } i \neq j.
\end{aligned}
\tag{4.26}
$$

As we see, g_2 has been defined twice and for uniqueness we require $g_2 = 0$. In addition, g must be an eigensolution of the internal boundary conditions (4.23); since $g_\nu = 0$ for $\nu \geqq 2$ and $C_{-2,0} = 0$, we get

$$
zg_0 = C_{-1,0}g_1, \qquad zg_1 = C_{-2,1}g_1. \tag{4.27}
$$

Using g_1 and· g_2 of (4.26) in (4.27) and in the equation $g_2 = 0$, we obtain a homogeneous system $E(z)\sigma = 0$, $\sigma = (\sigma_1, \sigma_2, \sigma_3)'$, with

$$
\begin{aligned}
\det E(z) = z[z - \tfrac{1}{4}\theta D^2(1 - D)^2]\cdot 2\kappa_1^3, &\qquad \text{if } \kappa_1 = \kappa_2 = \kappa_3,\\
\cdot\kappa_1(\kappa_3 - \kappa_1)^2, &\qquad \text{if } \kappa_1 = \kappa_2 \neq \kappa_3,\\
\cdot(\kappa_2 - \kappa_1)(\kappa_3 - \kappa_1)(\kappa_3 - \kappa_2), &\qquad \text{if } \kappa_i \neq \kappa_j \text{ for } i \neq j.
\end{aligned}
\tag{4.28}
$$

Since $0 < |D| < 1$ and $0 \leqq \theta \leqq 1$, det $E(z)$ cannot vanish for $|z| \geqq 1$. Hence, z with $|z| \geqq 1$ is not an eigenvalue of the operator G and Kreiss' Main Theorem assures stability. The proof in the case $\delta = -2$ is almost identical.

When $\delta = 1$, the alternative basic scheme is given by (3.41a), and the alternative perturbation is

$$
w_0^{m+1} = \sum_{j=-1}^{3} C_{j-1,0}w_j^m; \qquad w_1^{m+1} = \sum_{j=-1}^{3} C_{j-1,1}w_{j+1}^m. \tag{4.29}
$$

The characteristic equation is (3.42) and, for all z with $|z| \geqq 1$, the roots of this equation, κ_1 and κ_2, are not inside the unit disc. The indices of the perturbed scheme are $r' = 0$, $p' = 2$, $s' = 3$ and $k' = q = 1$. Therefore, from (1.17b) and (1.18a), we have that g, an eigensolution of G with an eigenvalue $|z| \geqq 1$, must satisfy

$$
g_\nu = 0, \quad \nu \geqq 2; \qquad g_\nu = \sum P_j\kappa_j^\nu, \quad \nu \leqq 1. \tag{4.30}
$$

For $\nu \leqq 1$, there are two possible solutions

$$
\begin{aligned}
g_\nu &= \sigma_1\kappa_1^\nu + \sigma_2\nu\kappa_1^\nu, && \text{if } \kappa_1 = \kappa_2,\\
&= \sigma_1\kappa_1^\nu + \sigma_2\kappa_2^\nu, && \text{if } \kappa_1 \neq \kappa_2.
\end{aligned}
\tag{4.31}
$$

Since g must also be an eigensolution of (4.29) and $g_2 = g_3 = g_4 = C_{-2,0} = 0$, g must satisfy

$$
zg_0 = C_{-1,0}g_0 + C_{0,0}g_1, \qquad zg_1 = C_{-2,1}g_0 + C_{-1,1}g_1, \tag{4.32}
$$

which leads us to the system $E(z)\sigma = 0$ with

$$
\begin{aligned}
\det E(z) = \tfrac{1}{4}\tilde{E}(z)\cdot\kappa_1, &\qquad \kappa_1 = \kappa_2,\\
\cdot(\kappa_2 - \kappa_1), &\qquad \kappa_1 \neq \kappa_2.
\end{aligned}
\tag{4.33}
$$

Here, $\tilde{E}(z)$ is quadratic in z and is given by (3.51). In the proof of Theorem 4, we have shown that the two roots z_1 and z_2 of $\tilde{E}(z) = 0$ are inside the unit disc. Hence, for $|z| \geqq 1$, $\det E(z) \neq 0$, and z with $|z| \geqq 1$ is not an eigenvalue of G, thus assuring stability. For $\delta = -1$, the proof is similar.

Finally, we consider $\delta = 0$ and, as usual, assume the scalar case. The basic scheme remains the L-W scheme (3.1), and the corresponding characteristic equation is (3.7). The roots of (3.7), κ_1 and κ_2, satisfy (3.11), and for all z with $|z| \geqq 1$, we have $|\kappa_1| \leqq 1$, $|\kappa_2| \geqq 1$. Here, $p = r = q = 1$ and g, an eigensolution of G, is defined by (1.17a) and (1.18a) to be

$$(4.34) \qquad g_\nu = \sigma_1 \kappa_1^\nu, \quad \nu \geqq 1; \qquad g_\nu = \sigma_2 \kappa_2^\nu, \quad \nu \leqq 0.$$

The appropriate internal boundary conditions are the original ones; namely (4.21a), and g, being their eigensolution, must satisfy

$$(4.35) \qquad z g_0 = \sum_{j=-1}^{2} C_{j,0} g_j, \qquad z g_1 = \sum_{j=-2}^{1} C_{j,1} g_{j+1}.$$

In these two equations, we substitute for the g_j's from (4.34), for the $C_{j,l}$'s from (4.21b), and for z from (3.7) with $\kappa = \kappa_2$. In addition, eliminating κ_1 by (3.11), we obtain a homogeneous system for the vector $\sigma = (\sigma_1, \sigma_2)$, with

$$(4.36) \qquad \det E(z) = \tfrac{1}{4} D^2 (D - 1) \kappa_2 \phi(\tau),$$

where $\tau \equiv \kappa_2^{-1}$, and $\phi(\tau)$ is defined by (3.56). Now, if z with $|z| \geqq 1$ is an eigenvalue of G, we must have $\det E(z) = 0$, and since $0 < |D| < 1$, $\det E(z)$ vanishes iff $\phi(\tau) = 0$. We have shown in the last part of the proof of Theorem 4 that, for $0 < D < 1$, all the roots of $\phi(\tau) = 0$ are outside the unit disc, which contradicts the fact that $|\kappa_2| \geqq 1$ for $|z| \geqq 1$. If, on the other hand, $-1 < D < 0$, we use (3.11) to eliminate κ_2 from (4.36) and thereby find that $\det E(z) = 0$ iff Eq. (3.59) is satisfied. As we argued after (3.59), this does not occur for κ_1 with $|\kappa_1| \leqq 1$ and again a contradiction results. Therefore, $\det E(z) \neq 0$ for all z with $|z| \geqq 1$, and we have stability.

5. Numerical Results.

In this chapter, we test the practical applicability of some of the linear stability analysis results of the two previous chapters, to nonlinear systems of the form

$$(5.1) \qquad \begin{aligned} &u_t = F(u)_x; \qquad A(u) \equiv \partial F(u)/\partial u; \\ &\qquad\qquad -\infty < x < \infty,\ t \geqq 0, \qquad u(x, 0) = f(x). \end{aligned}$$

The scalar test problem we selected is [4]

$$(5.2a) \qquad u_t = (-u^2/2)_x; \qquad A(u) = -u; \qquad -\infty < x < \infty,\ t \geqq 0;$$

$$(5.2b) \qquad \begin{aligned} u(x, 0) &= \ \ 1, & x &\leqq -1, \\ &= -x, & -1 &\leqq x \leqq 0, \\ &= \ \ x, & 0 &\leqq x \leqq 1, \\ &= \ \ 1, & 1 &\leqq x. \end{aligned}$$

The solution of (5.2) remains continuous until $t = 1$, at which point the compression region, initially located on $-1 \leqq x \leqq 0$, becomes a discontinuity ("shock-wave"), while the rarefaction wave on its right has positive gradients which continue to decrease with time. The solution for $t < 1$ is given by

$$
\begin{aligned}
u(x, t) &= 1, & x &\leqq t - 1,\\
&= x/(t - 1), & t - 1 &\leqq x \leqq 0,\\
&= x/(t + 1), & 0 &\leqq x \leqq t + 1,\\
&= 1, & t + 1 &\leqq x;
\end{aligned}
\tag{5.3a}
$$

and, for $t \geqq 1$, by

$$
\begin{aligned}
u(x, t) &= 1, & x &\leqq t + 1 - (2(t + 1))^{1/2},\\
&= x/(t + 1), & t + 1 - (2(t + 1))^{1/2} &\leqq x \leqq t + 1,\\
&= 1, & t + 1 &\leqq x.
\end{aligned}
\tag{5.3b}
$$

In this problem, we regard the shock-wave as a moving internal boundary, its trajectory being given by

$$x = \chi(t) = t + 1 - (2(t + 1))^{1/2}. \tag{5.4}$$

The choice of (5.2) as a test problem was dictated by the following considerations. As we saw, at $t = 1$, the solution develops a discontinuity which is preceded by a *nonuniform* rarefaction region with time varying gradients. This allows us to compare the numerical and exact solutions with regard to shock-speed and order of accuracy of the finite difference approximation in the rarefaction region. Note that the piecewise smooth initial values (5.2b) describe a polygonal function. Therefore, the solution (5.3) is polygonal as well, hence easily calculated—for comparison purposes with the numerical solution—to any degree of accuracy.

5.1. *Three-Point Perturbations.* The Lax-Wendroff finite difference approximation to (5.1) is well known to be [11].

$$
\begin{aligned}
v_\nu^{m+1} = \tilde{Q} v_\nu^m \equiv v_\nu^m + \tfrac{1}{2}\lambda(F_{\nu+1}^m - F_{\nu-1}^m) + \tfrac{1}{2}\lambda^2&\\
\cdot[A_{\nu+1/2}^m(F_{\nu+1}^m - F_\nu^m) - A_{\nu-1/2}^m(F_\nu^m - F_{\nu-1}^m)]&,
\end{aligned}
\tag{5.5}
$$

where we use the abbreviations $F_\nu{}^m \equiv F(v_\nu{}^m)$ and $A_{\nu\pm1/2}{}^m \equiv A[\frac{1}{2}(v_\nu{}^m + v_{\nu\pm1}{}^m)]$. The Lax-Friedrichs approximation to (5.1) is (see [10])

$$v_\nu^{m+1} = \tilde{L} v_\nu^m \equiv \tfrac{1}{2}(v_{\nu+1}^m + v_{\nu-1}^m) + \tfrac{1}{2}\lambda(F_{\nu+1}^m - F_{\nu-1}^m). \tag{5.6}$$

Consider now the scheme

$$v_\nu^{m+1} = [\theta_\nu \tilde{L} + (1 - \theta_\nu^m)\tilde{Q}]v_\nu^m. \tag{5.7}$$

If $\theta_\nu{}^m$ is given by (3.29), then (5.7) is exactly the nondivergenceless nonlinear version of the perturbed scheme dealt with in Corollary 2. In applying (5.7) to nonlinear problems, it must be understood that the condition of constant perturbation speed, stipulated in (3.29), cannot be realized in practice. In practical computations therefore, $\theta_\nu{}^m$ is taken to be nonzero only on the *actual* trajectory of the internal boundary. For our particular test-case, it means that $\theta_\nu{}^m$ will differ from zero only along the locus of the numerically created shock-wave which should approximate (5.4).

While the basic L-W approximation, applied to problems like (5.2), is well known to preserve the correct shock-speed as well as second order accuracy in smooth regions of the solution, it also introduces strong overshoots directly behind the shock. On the other hand, it is known that the employment of first order accurate schemes, such as the L-F approximation, yields monotone solutions in regions of discontinuities and strong gradients. In addition, shock-profiles computed by L-W are usually much steeper than those computed by L-F. This gave us the motivation for using the L-W scheme, perturbed *locally* by a L-F type of perturbation. However, three questions remain to be answered by the numerical computations:

(a) Will local employment of first order approximations be as effective as their global employment in smoothing out post-shock oscillations and overshoots?

(b) Will the overall second order accuracy of the L-W scheme be adversely affected by the introduction of first order perturbations at the moving boundary?

(c) Will the shock-profiles computed by the perturbed scheme be steeper than those computed by first order accurate schemes?

Our test problem (5.2) was first solved numerically using the perturbed scheme (5.7) where θ_ν^m, by analogy to (3.29), is different from zero at a single point only, at each time level, and is defined by

$$\begin{aligned} \theta_\nu^m &= \theta = \text{constant}, & \nu &= \nu_{m,\max}, \\ &= 0, & \nu &\neq \nu_{m,\max}. \end{aligned} \tag{5.8a}$$

Here, $\nu_{m,\max}$ is the grid-index at which the numerical shock-wave is defined to be located, i.e., the index which satisfies

$$|v^m_{\nu_{m,\max}+1} - v^m_{\nu_{m,\max}-1}| = \max |v^m_{\nu+1} - v^m_{\nu-1}|. \tag{5.8b}$$

The numerical computations were carried out for $\Delta x = 0.1, 0.02, 0.01$; $\theta = 0, 0.2, 0.5, 1.0, 1.25, 1.5$; and $0 \leqq t \leqq 2$. Note that the range of θ is dictated by the mutual range of stability given in Corollary 2, and that $\theta = 0$ is the case for which the perturbed L-W scheme reduces to the L-W scheme itself.

The numerical results allow us to answer the question posed above as follows: (a) The local employment of our L-F-type first order accurate perturbation, at a single point only at each time level, is not quite as effective as its global employment in reducing the post-shock overshoot. However, as θ is increased, the overshoot becomes *significantly* smaller. (b) The overall accuracy of the solution, outside the shock region, is well preserved. (c) The shock-profiles are much steeper than those computed by the L-F scheme, though not as steep as the L-W results.

In order to increase the effectiveness of the perturbation in reducing the post-shock overshoot, we decided to use the perturbed scheme (5.7), with the perturbation still employed locally though not necessarily at a single point only. Therefore, instead of θ_ν^m of (5.8), we use, in (5.7),

$$\begin{aligned} \theta_\nu^m &= \theta = \text{const}, \quad \text{for all } \nu \text{ satisfying } |v^m_{\nu+1} - v^m_{\nu-1}| \geqq \alpha\Delta x \ (\alpha = \text{const}), \\ &= 0, \qquad\qquad \text{otherwise}. \end{aligned} \tag{5.9}$$

Hence, θ_ν^m is defined by a dynamic process, and the number of grid points at which θ_ν^m differs from zero—which depends on the predetermined positive value of α—varies from one time level to another. Obviously, as the *threshold gradient*, $\alpha/2$, is set at lower values, the number of points at which the perturbation appears increases.

The numerical computations were carried out for the same range of parameters as in the previous case with $\alpha\Delta x = 0.8, 0.4, 0.2, 0.1$. The expected reduction in the post-shock overshoot was realized in all cases except $\alpha\Delta x = 0.8$. For this value of $\alpha\Delta x$, we found that the perturbation was employed, on the average, at only one point approximately, at each time level. Hence, its effect was similar to the previous cases, where $\theta_\nu{}^m$ was defined by (5.8). It turns out that the mean number of perturbed grid points per time level, say N_p, depends only on $\alpha\Delta x$. Since the discontinuity in the solution first appears at $t = 1$, the computation of N_p is based on all $t_m \geqq 1$. For $\alpha\Delta x = 0.4, 0.2, 0.1$, $N_p \sim 2, 3, 4$, respectively. These values of N_p constitute a small fraction of the total number of grid points in the nonuniform region of the solution, provided Δx is small enough. Thus, we expected the second order accuracy to be preserved. This is indeed the case when $\Delta x = 0.02, 0.01$. When $\Delta x = 0.1$, then $N_p = 4$ means that about 25% of the grid points in the nonuniform region are perturbed. This explains why the overall results are only first order accurate in this case. On the basis of the above empirical results, we conclude that, for a given Δx, the perturbed scheme (5.7) with (5.9) is effective in reducing post-shock overshoots and maintaining second order accuracy, provided α is neither too large nor too small. In addition, it seemed to us that $\theta = 0.5$ gave the best results.

Another point which was checked had to do with the steepness of the computed shock-profile. The shock, described as a discontinuity in the exact solution, is smeared over several mesh intervals in the numerical case. Accordingly, we denote by N_s the number of mesh-intervals over which 99% of the *expected* jump across the shock is spread.

A sample of the numerical results of (5.7), relating to the two choices of $\theta_\nu{}^m$, defined in (5.8) and (5.9), are given in Table 1. For comparison purposes, we give the first order accurate results of the pure L-F scheme as well.

Type of scheme	θ	$\alpha\Delta x$	*No. of time steps*	N_s	N_p	**Mean over-shoot*	**$10^{-5}\times$ *error* (2, 2)	*Error in shock location*	*Remarks*
(5.6) Pure L-F	—	—	200	8	—	.0000	244.800	0	D, I
(5.5) pure L-W	—	—	213	2	—	.1228	1.414	0	D
(5.7) + (5.6) + (5.8)	1.5	—	204	2	1	.0340	1.468	$-18\Delta x$	
(5.7) + (5.6) + (5.9)	0.5	0.8	205	3	1.10	.0516	1.464	$-\ 5\Delta x$	
(5.7) + (5.6) + (5.9)	0.5	0.4	201	3	2.19	.0109	1.490	$-\ 3\Delta x$	
(5.7) + (5.6) + (5.9)	0.5	0.2	200	4	3.02	.0025	1.497	$-\ 2\Delta x$	
(5.7) + (5.6) + (5.9)	0.5	0.1	200	4	3.65	.0007	1.498	$-\ \Delta x$	

TABLE 1. Nondivergenceless L-F-type perturbations. All results are for $\Delta x = 0.01$, and $t \sim 2$. The letters I and D denote first order accuracy and divergenceless schemes, respectively.

* Behind the shock, the correct value of the solution is always 1; therefore, we define the overshoot to be $\max_\nu \{v_{\nu m} - 1\}$.

** We define error(2, 2) $\equiv |v(2, 2) - u(2, 2)|$ where $v(2, 2)$ and $u(2, 2)$ are the numerical and exact solutions at the point $(x, t) = (2, 2)$, respectively. When $t = 2$, then $x = 2$ is near the center of the rarefaction region.

We see from Table 1 that the L-W or L-F schemes predict the correct shock location. However, scheme (5.7), with $\theta_\nu{}^m$ given either by (5.8) or (5.9), does not agree with the theoretical location given by (5.4). This is expected, since (5.7) is not in conservation form.

In order to get the divergence-free version of (5.7), we have to construct a nonlinear analogue of (4.3), which takes the form, (see [7]),

$$\begin{aligned} v_\nu^{m+1} &= v_\nu^m + \tfrac{1}{2}[\theta^m_{\nu+1/2}(v^m_{\nu+1} - v^m_\nu) - \theta^m_{\nu-1/2}(v^m_\nu - v^m_{\nu-1})] + \tfrac{1}{2}\lambda(F^m_{\nu+1} - F^m_{\nu-1}) \\ &\quad + \tfrac{1}{2}\lambda^2[(1 - \theta^m_{\nu+1/2})A^m_{\nu+1/2}(F^m_{\nu+1} - F^m_\nu) \\ &\qquad - (1 - \theta^m_{\nu-1/2})A^m_{\nu-1/2}(F^m_\nu - F^m_{\nu-1})]. \end{aligned} \tag{5.10}$$

The fact that (4.3) is indeed the linearized version of (5.10) is verified by setting in the last scheme $A_{\nu\pm 1/2}{}^m = A$ and $F_\nu{}^m = Av_\nu{}^m$.

Now, if $\theta_{\nu+1/2}{}^m$ is given by (4.4), then (5.10) is exactly the nonlinear version of the perturbed scheme considered in Theorem 5. Obviously, the remarks following (5.7), concerning the nonlinearity of the shock-trajectory, are valid here as well. Therefore, by analogy with (5.8), we take

$$\begin{aligned} \theta^m_{\nu+1/2} &= \theta = \text{const}, \qquad & \nu &= \nu_{m,\max}, \\ &= 0, & \nu &\neq \nu_{m,\max}, \end{aligned} \tag{5.11a}$$

where $\nu_{m,\max}$ is the grid-index which now satisfies

$$|v^m_{\nu_{m,\max}+1} - v^m_{\nu_{m,\max}}| = \max_\nu |v^m_{\nu+1} - v^m_\nu|. \tag{5.11b}$$

The numerical computations using (5.10) and (5.11) yield, as far as the overall accuracy and post-shock overshoot reduction are concerned, almost identical results to those obtained from (5.7) with (5.8). However, this time, we obtain the *correct* shock location.

Again, in order to increase the effectiveness of the perturbation in reducing post-shock oscillation, we redefine $\theta_{\nu+1/2}{}^m$ by analogy with (5.10) to be

$$\begin{aligned} \theta^m_{\nu+1/2} &= \theta = \text{const}, \quad && \text{for all } \nu \text{ satisfying } |v^m_{\nu+1} - v^m_\nu| \geqq \beta\Delta x \ (\beta = \text{const}), \\ &= 0, && \text{otherwise}, \end{aligned} \tag{5.12}$$

where β is the present threshold gradient. The computations using (5.10) with (5.12) were carried out for the same range of parameters as before. Note that this means using $\beta\Delta x = 0.4$, 0.2, 0.1 and 0.05. The numerical calculations show, as we had hoped, that this scheme gives the best results in the sense that the overall second order accuracy is maintained, a correct shock location is obtained and the post-shock overshoots are drastically reduced. A sample of the numerical results of using (5.10) with (5.11) or (5.12) are given in Table 2.

5.2. *Five-Point Perturbations.* So far, we consider in this chapter the nonlinear versions of three-point perturbations. We now turn our attention to five-point perturbations such as the ones considered in Theorems 4 and 6.

We start with the nonlinear analogue of (3.44):

$$v_\nu^{m+1} = [I + (1 - \theta^m_\nu)(\tilde{Q} - I) + \theta^m_\nu(\tilde{Q} - I)\tilde{Q}]v^m_\nu, \qquad 0 \leqq \theta^m_\nu \leqq 1, \tag{5.13}$$

$\theta_{\nu+1/2}{}^m$	$\beta\Delta x$	*No. of time steps*	N_s	N_p	*Mean over-shoot*	$10^{-5}\times$ *error*(2, 2)
(5.11)	-	203	3	1	.0267	1.474
(5.12)	0.4	202	3	1.00	.0222	1.482
(5.12)	0.2	200	3	2.05	.0034	1.497
(5.12)	0.1	200	4	2.74	.0007	1.498
(5.12)	0.05	200	4	3.38	.0001	1.498

TABLE 2. Divergence-free L-F-type perturbations (Eq. (5.10)). $\Delta x = 0.01$, $\theta = 0.5$, $t \sim 2$. No error in shock location.

where $\tilde{Q}$, the nonlinear L-W difference operator, is defined in (5.5). When $\theta_\nu{}^m$ is given by (3.29), (5.13) becomes the nonlinear version of the perturbed scheme considered in Theorem 4. However, due to the nonlinearity of the shock trajectory $\theta_\nu{}^m$ must be redefined. If we take $\theta_\nu{}^m$ to be given by either (5.8) or (5.9) we have, respectively, the cases where the perturbation, at each time level, is applied either at a single point only, or, depending on the value of the preset threshold gradient, at a variable number of points. These two versions of (5.13) are not divergence-free and, using the same range of parameters as for the two corresponding versions of (5.7) (except that here θ is bounded by 1), we obtained very similar results.

The divergence-free version of (5.13) is

$$
\begin{aligned}
v_\nu^{m+1} &= v_\nu^m + \tfrac{1}{2}\lambda[(1-\theta_{\nu+1/2}^m)(F_{\nu,+1}^m + F_\nu^m) - (1-\theta_{\nu-1/2}^m)(F_\nu^m + F_{\nu-1}^m)] \\
&\quad + \tfrac{1}{2}\lambda^2[(1-\theta_{\nu+1/2})A_{\nu+1/2}^m(F_{\nu+1}^m - F_\nu^m) - (1-\theta_{\nu-1/2}^m)A_{\nu-1/2}^m(F_\nu^m + F_{\nu-1}^m)] \\
&\quad + \tfrac{1}{2}\lambda[\theta_{\nu+1/2}^m(\tilde{F}_{\nu+1}^m + \tilde{F}_\nu^m) - \theta_{\nu-1/2}(\tilde{F}_\nu^m + \tilde{F}_{\nu-1}^m)] \\
&\quad + \tfrac{1}{2}\lambda^2[\theta_{\nu+1/2}^m\tilde{A}_{\nu+1/2}^m(\tilde{F}_{\nu+1}^m - \tilde{F}_\nu^m) - \theta_{\nu-1/2}^m\tilde{A}_{\nu-1/2}^m(\tilde{F}_\nu^m - \tilde{F}_{\nu-1}^m)],
\end{aligned}
\tag{5.14a}
$$

where

$$\tilde{F}_\nu^m \equiv F(\tilde{Q}v_\nu^m), \qquad \tilde{A}_{\nu\pm1/2}^m \equiv [A(\tilde{Q}v_{\nu\pm1}^m) + A(\tilde{Q}v_\nu^m)]/2. \tag{5.14b}$$

If $\theta_{\nu+1/2}{}^m$ is given by (4.4), then (5.14) is the nonlinear version of the perturbed scheme considered in Theorem 6. In practice, however, $\theta_\nu{}^m$ is chosen, as before, to be given either by (5.11) or by (5.12). The numerical results of these two versions of (5.14) are again of the same quality as those found in the corresponding cases of the three-point perturbation (5.10).

We remark that in all of the numerical computations described in subsections 5.1 and 5.2 the time step at each time level, Δt_m, was chosen to be the maximal one allowed by the stability condition of the L-W scheme, i.e.,

$$t_m = \frac{\Delta x}{\max\limits_\nu\{v_\nu^m\}}. \tag{5.15}$$

By taking smaller time steps, the results obtained were not improved in any respect, and usually were worse.

Department of Mathematical Sciences
Tel Aviv University
Tel Aviv, Israel

1. S. Abarbanel & M. Goldberg, "Numerical solution of quasi-conservative hyperbolic systems—the cylindrical shock problem," *J. Comput. Phys.*, v. 10, 1972, pp. 1–21.

2. S. Abarbanel & G. Zwas, "An iterative finite difference method for hyperbolic systems," *Math. Comp.*, v. 23, 1969, pp. 549–565. MR **40** #1044.

3. M. Ciment, "Stable matching of difference schemes," *SIAM J. Numer. Anal.*, v. 9, 1972, pp. 695–701.

4. M. Goldberg & S. Abarbanel, *A Note on Discontinuities in a Nonlinear Hyperbolic Equation with Piecewise Smooth Data*, Dept. of Math. Sciences, Tel Aviv Univ. Report, 1972.

5. M. Goldberg, "A note on the stability of an iterative finite-difference method for hyperbolic systems," *Math. Comp.*, v. 27, 1973, pp. 41–44.

6. B. Gustafsson, H. O. Kreiss & A. Sundström, "Stability theory of difference approximations for mixed initial boundary value problems. II," *Math. Comp.*, v. 26, 1972, pp. 649–686.

7. A. Harten & G. Zwas, "Self-adjusting hybrid schemes for shock computations," *J. Comput. Phys.*, v. 9, 1972, pp. 568–583.

8. H. O. Kreiss, "Difference approximations for the initial-boundary value problem for hyperbolic differential equations," *Numerical Solutions of Nonlinear Differential Equations* (Proc. Adv. Sympos., Madison, Wis., 1966), Wiley, New York, 1966, pp. 141–166. MR **35** #5156.

9. H. O. Kreiss, "Stability theory for difference approximations of mixed initial boundary value problems. I," *Math. Comp.*, v. 22, 1968, pp. 703–714. MR **39** #2355.

10. P. D. Lax, "Weak solutions of nonlinear hyperbolic equations and their numerical computations," *Comm. Pure Appl. Math.*, v. 7, 1954, pp. 159–193. MR **16**, 524.

11. P. D. Lax & B. Wendroff, "Systems of conservation laws," *Comm. Pure Appl. Math.*, v. 13, 1960, pp. 217–237. MR **22** #11523.

12. R. D. Richtmyer & K. W. Morton, *Difference Methods for Initial Value Problems*, 2nd ed., Interscience Tracts in Pure and Appl. Math., vol. 4, Interscience, New York, 1967. MR **36** #3515.