JOURNAL OF COMPUTATIONAL PHYSICS **47**, 489–496 (1982)

Note

Stability of the Explicit Finite Differenced Transport Equation*

Recently Leonard [4] and, independently, Clancy [1] have presented necessary and sufficient conditions for the von Neumann stability of the explicit forward-time central-space (FTCS) one-dimensional transport equation in an infinite domain. Their results correct the work of Fromm [2] and Roache [5] which was only sufficient. It is the purpose of this note to show that the stability result obtained by Leonard and Clancy is also sufficient but not necessary in a finite computational domain. In doing so, we shall present the correct necessary and sufficient condition for the numerical solution of the transport equation. The result is given in a general form which includes different treatments of the convective term. The new result does not affect the stability region for grid Peclét numbers (Pe_Δ) of approximately two or smaller, but enlarges the stability region for large Pe_Δ. Although the enlargement is slight for the FTCS scheme, it is substantial for other schemes considered here.

The one-dimensional transport equation under consideration is

$$\frac{\partial T}{\partial t} + u\frac{\partial T}{\partial x} = \alpha\frac{\partial^2 T}{\partial x^2}, \tag{1}$$

and for definiteness we assume that $u \geqslant 0$; the case of $u < 0$ follows in an obvious fashion. The explicit finite difference equation which approximates (1) on a uniform grid is given by

$$T_j^{n+1} = T_j^n - c(T_{r_j}^n - T_{l_j}^n) + \Gamma(T_{j+1}^n - 2T_j^n + T_{j-1}^n), \tag{2}$$

where $T_j^n = T(x_j, t_n)$, $x_j = j\Delta x$, $t_n = n\Delta t$, and

$$c = u\Delta t/\Delta x, \qquad \Gamma = \alpha\Delta t/\Delta x^2, \tag{3}$$

$$\begin{aligned} T_{l_j} &= T_{j-1} - \tfrac{1}{2}I(T_{j-1} - T_j) - \tfrac{1}{6}J(T_{j-2} - 2T_{j-1} + T_j), \\ T_{r_j} &= T_j - \tfrac{1}{2}I(T_j - T_{j+1}) - \tfrac{1}{6}J(T_{j-1} - 2T_j + T_{j+1}). \end{aligned} \tag{4}$$

The values of I and J for different treatments of the convective term are given in Table I. As correctly pointed out by Leonard [3], T_{l_j} and T_{r_j} can be represented by a zeroth, linear, or quadratic upstream interpolation of T on the left and right faces of a control volume centered at j. The above interpolations give rise to the upwind,

* Work supported by the U.S. Department of Energy under Contract DE–AC04–76DP00789.

0021-9991/82/090489–08\$02.00/0

581/47/3-11

TABLE I

Necessary and Sufficient Stability Results for Different Approximations of Convective Term

Convective scheme	I	J	Stability condition	
Upwind	0	0	$\Gamma \leqslant \frac{1}{2}(1/(1+\frac{1}{2}\mathrm{Pe}_\Delta))$,	$0 \leqslant \mathrm{Pe}_\Delta < \infty$
Central	1	0	$\Gamma \leqslant \frac{1}{2}$,	$0 \leqslant \mathrm{Pe}_\Delta \leqslant 2$,
			$\leqslant C(\theta^*; 0, \mathrm{Pe}_\Delta)$,	$2 < \mathrm{Pe}_\Delta < \infty$
QUICK	1	$\frac{3}{4}$	$\Gamma \leqslant \frac{1}{2}(1/(1+\frac{1}{4}\mathrm{Pe}_\Delta))$,	$0 \leqslant \mathrm{Pe}_\Delta \leqslant \frac{1}{2}(1+\sqrt{17})$,
			$\leqslant C(\theta^*; \frac{3}{4}, \mathrm{Pe}_\Delta)$,	$\frac{1}{2}(1+\sqrt{17}) < \mathrm{Pe}_\Delta < \infty$
HOS	1	1	$\Gamma \leqslant \frac{1}{2}(1/(1+\frac{1}{3}\mathrm{Pe}_\Delta))$,	$0 \leqslant \mathrm{Pe}_\Delta \leqslant \frac{2}{3}(1+\sqrt{10})$,
			$\leqslant C(\theta^*; 1, \mathrm{Pe}_\Delta)$,	$\frac{2}{3}(1+\sqrt{10}) < \mathrm{Pe}_\Delta < \infty$

centered, or quadratic upstream interpolation for convective kinematics (QUICK) [3] treatment of the convective term. The higher order scheme (HOS) treatment is obtained by removing the central difference convective truncation error term $\frac{1}{6}u\Delta x^2(\partial^3 T/\partial x^3)$ in going from (1) to (2).

It can be shown following the method of Warming and Hyett [6] that the solution of (2) represents the solution of the following modified transport equation

$$\frac{\partial T}{\partial t} + u\frac{\partial T}{\partial x} = \alpha' \frac{\partial^2 T}{\partial x^2} + \text{T.E.}, \tag{5}$$

where the truncation error is

$$\begin{aligned}\text{T.E.} = &-\frac{1}{2}\Delta t u^2 \frac{\partial^2 T}{\partial x^2} + \alpha'\Delta x\left[c - \frac{1-J}{6}\mathrm{Pe}'_\Delta\right]\frac{\partial^3 T}{\partial x^3} \\ &-\frac{1}{2}\alpha'\Delta x^2\left[\Gamma' - \frac{1}{6}(1 - J\mathrm{Pe}'_\Delta)\right]\frac{\partial^4 T}{\partial x^4} + O(\Delta t^2, \Delta x^4),\end{aligned} \tag{6}$$

where $\alpha' = \alpha(1 + \frac{1}{2}(1-I)\,\mathrm{Pe}_\Delta)$, $\Gamma' = \alpha'\Delta t/\Delta x^2$, $\mathrm{Pe}'_\Delta = u\Delta x/\alpha'$, and $\mathrm{Pe}_\Delta = u\Delta x/\alpha$ is the grid Peclét number. Each treatment of the convective term gives rise to different truncation errors. The upwind scheme has the poorest accuracy, being only first-order accurate in space. All other schemes result in second-order accurate approximations to (5), even though the HOS approximation for the convective term is third-order accurate. Each of the above schemes also has its corresponding stability region. As noted by Leonard [3], the essence of a stable approximation for the convective term must be such that its sensitivity to T_j is negative, i.e., $\partial C_j/\partial T_j < 0$, where C_j is the finite difference approximation to $-u(\partial T/\partial x)|_j$. The sensitivity of the four schemes are $-u/\Delta x$ for upwind (stable), zero for central (neutrally stable), $-\frac{3}{8}u/\Delta x$ for QUICK (stable, but less than upwind), and $-\frac{1}{2}u/\Delta x$ for HOS (stable, but less than upwind and more than QUICK). This same order in the stability of the four schemes will be shown to remain valid in our final stability result.

A standard von Neumann stability analysis studies the development of a Fourier component

$$T_j^n = A^n e^{ikx_j} = A^n e^{i\theta j}, \tag{7}$$

where $\theta = k\Delta x$. The complex amplification factor $G = A^{n+1}/A^n$ is obtained by substituting (7) into (2) to obtain

$$\begin{aligned} G = 1 &+ 2\Gamma'(\cos\theta - 1) + (\tfrac{1}{6}J)\, c(4\cos\theta - \cos 2\theta - 3) \\ &- ic[\sin\theta + (\tfrac{1}{6}J)(2\sin\theta - \sin 2\theta)]. \end{aligned} \tag{8}$$

Alternately, we could have written

$$T_j^n = Be^{\sigma t_n} e^{ikx_j} = Be^{\sigma n\Delta t} e^{i\theta j}, \tag{9}$$

where $\sigma = a + ib$ is the growth rate. It can now be seen immediately that we can write $G = |G| e^{i\phi}$, where $|G| = e^{a\Delta t}$, and $\phi = b\Delta t$. Thus we get

$$\begin{aligned} a = (1/2\Delta t) \ln\{&[1 + 2\Gamma'(\cos\theta - 1) + \tfrac{1}{6}Jc(4\cos\theta - \cos 2\theta - 3)]^2 \\ &+ c^2[\sin\theta + \tfrac{1}{6}J(2\sin\theta - \sin 2\theta)]^2\}, \end{aligned} \tag{10}$$

$$b = -\frac{1}{\Delta t}\tan^{-1}\left\{\frac{c[\sin\theta + \frac{1}{6}J(2\sin\theta - \sin 2\theta)]}{[1 + 2\Gamma'(\cos\theta - 1) + \frac{1}{6}Jc(4\cos\theta - \cos 2\theta - 3)]}\right\}. \tag{11}$$

A stable numerical solution of (2) will generally contain errors in amplitude (dissipation) and phase (dispersion). Dissipation and dispersion errors can be analyzed by assuming the solution of (1) to be of the form $e^{\sigma t}e^{ikx}$. Then it can be shown that the amplitude ratio and phase shift per time step are $e^{-\Gamma\theta^2}$ and $-c\theta$, respectively. Consequently, the relative dissipation and dispersion per time increment are given by $-a\Delta t/\Gamma\theta^2$ and $-b\Delta t/c\theta$, respectively, i.e., the amplitude decays faster than the exact solution, and the wave speed is larger than that of the exact solution if the ratios are larger than unity; the opposite statement is true if they are less than unity. Since we are concerned in this paper about the stability of the schemes, we shall not elaborate further on the above errors.

For stability, it is required that $a \leqslant 0$, or, equivalently, from (10) we get

$$\Gamma' \leqslant \min_{\theta^* \leqslant \theta \leqslant \pi} C(\theta; J, \mathrm{Pe}'_\Delta), \tag{12}$$

where $\theta^* = (2\Delta x/l)\,\pi$, l being the largest dimension in the computational domain, and

$$\begin{aligned} C(\theta; J, \mathrm{Pe}'_\Delta) = {} & [(1-\cos\theta) - \tfrac{1}{12} J\,\mathrm{Pe}'_\Delta(4\cos\theta - \cos 2\theta - 3)] \\ & \div \{[(1-\cos\theta) - \tfrac{1}{12} J\,\mathrm{Pe}'_\Delta(4\cos\theta - \cos 2\theta - 3)]^2 \\ & + \tfrac{1}{4}\mathrm{Pe}'^2_\Delta[\sin\theta + \tfrac{1}{6}J(2\sin\theta - \sin 2\theta)]^2\}. \end{aligned} \tag{13}$$

Condition (12) is necessary and sufficient for (2) to be stable. Note that the smallest wavelength possible in a computation domain is $2\Delta x$, thus $\theta = k\Delta x = (2\pi/2\Delta x)\,\Delta x = \pi$, and the largest wavelength possible equals l, thus $\theta = (2\pi/l)\,\Delta x = (2\Delta x/l)\,\pi = \theta^*$. Hence θ can range from θ^* to π. We should comment here that the von Neumann stability analysis is *only* correct if the boundary conditions of (2) are periodic, and that is the reason why l is the largest wavelength that can be obtained in a computational domain.

As shown in Fig. 1, the minimum of $C(\theta;\, J,\, \mathrm{Pe}'_\Delta)$ for all Pe_Δ occurs either at $\theta = \theta^*$ or at $\theta = \pi$. Before discussing the results, we give some useful limits for expressions (12) and (13).

(i) $\mathrm{Pe}'_\Delta = 0$: $C(\theta; J, 0) = 1/(1-\cos\theta)$, so $\Gamma' \leqslant \frac{1}{2}$.

(ii) $\mathrm{Pe}'_\Delta \gg 1$: In this case we have that

$$\begin{aligned} \Gamma' & \leqslant 4/\mathrm{Pe}'^2_\Delta(1+\cos\theta^*), && \text{for} \quad J = 0, \\ & \leqslant -12J(4\cos\theta^* - \cos 2\theta^* - 3)/ \\ & \qquad \mathrm{Pe}'_\Delta\{[J(4\cos\theta^* - \cos 2\theta^* - 3)]^2 \\ & \quad + 36[\sin\theta^* + \tfrac{1}{6}J(2\sin\theta^* - \sin 2\theta^*)]^2\}, && \text{for} \quad J > 0. \end{aligned} \tag{14}$$

(iii) $\theta = \theta^*$: $\Gamma' \leqslant C(\theta^*; J, \mathrm{Pe}'_\Delta)$.

(iv) $\theta = \pi$: $\Gamma' \leqslant \frac{1}{2}(1/(1 + \frac{1}{3}J\,\mathrm{Pe}'_\Delta))$.

The necessary and sufficient conditions for the four schemes, using the above limits, can now be clearly tabulated as in Table I.

How do results (12) and (13) differ from that obtained by Leonard [3, 4] for $I = 1$, $J = 0, \frac{3}{4}$, and by Clancy [1] for $I = 1$, $J = 0$? In general, the results differ when the stability is being controlled by $\theta = \theta^*$. To see this, G can be plotted in the complex plane (for $0 \leqslant \theta \leqslant \pi$), showing that the locus is a generalized semi-ellipse. In essence, the above authors have required that the curvature of G at $(1, 0)$ be greater than that of the unit circle, i.e., $\partial^2|G|/\partial\theta^2|_{\theta=0} \leqslant 0$ (see [3, 4]). This condition is equivalent to evaluating (13) at $\theta = 0$ (by using l'Hospital's rule twice), which is also equivalent to requiring that (2) be stable to infinite wavelength disturbances. In an actual computation, this is not necessary since it is not possible to have $l = \infty$ or $\Delta x = 0$, and hence the resulting restriction (for all cases except $I = J = 0$) of $\Gamma' \leqslant C(0;\, J,\, \mathrm{Pe}'_\Delta) = 2/\mathrm{Pe}^2_\Delta$ is verly conservative for large Pe_Δ. To illustrate this, in Fig. 2 we show

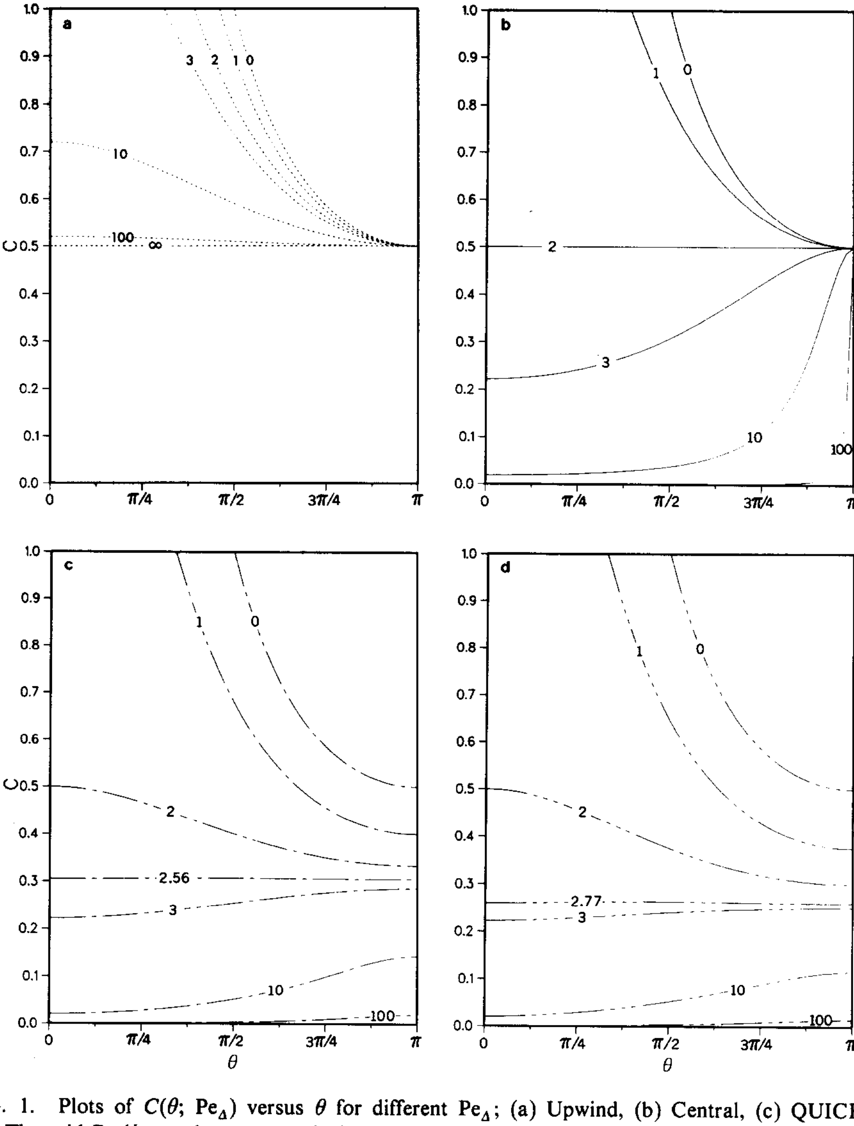

FIG. 1. Plots of $C(\theta;\ \mathrm{Pe}_\Delta)$ versus θ for different Pe_Δ; (a) Upwind, (b) Central, (c) QUICK, (d) HOS. The grid Peclét numbers are marked on the curves.

$F = C(\theta^*; J, \mathrm{Pe}_\Delta)/C(0; J, \mathrm{Pe}_\Delta)$ as a function of θ^* for different Pe_Δ and J, for $I = 1$. Here F is the time step factor that can be obtained over that given by Leonard for the FTCS and QUICK schemes, and by Clancy for the FTCS scheme. For FTCS with a realistic number of gridpoints, F remains close to one. For the other schemes, however, it increases with increasing J, and can be very large for large Pe_Δ, thus also illustrating the advantage of using the QUICK or HOS formulation over central difference for the convective term in the equation. Note that for the upwind scheme

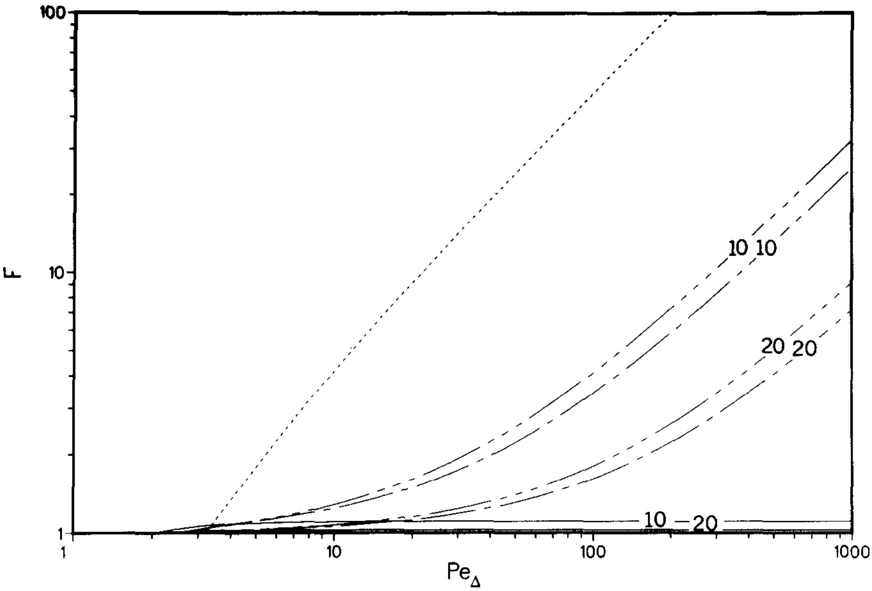

FIG. 2. Plot of the stability time step factor $F = C(\theta^*)/C(0)$ versus Pe_Δ for $\mathrm{Pe}_\Delta \geqslant 2$, and $\theta^* = (2\Delta x/l)\,\pi$, with $l/\Delta x = 10, 20$; Upwind (---), Central (—), QUICK (——), HOS (———).

$F = \frac{1}{4}\mathrm{Pe}_\Delta^2/(1 + \frac{1}{2}\mathrm{Pe}_\Delta)$ is independent of θ^*, and is included in Fig. 2 for completeness. Of course, F is largest for the upwind scheme, though this scheme may be less accurate since it introduces large artificial diffusion. Notice that for the FTCS scheme, $F \to 2/(1 + \cos\theta^*)$ as $\mathrm{Pe}_\Delta \to \infty$, while for the other schemes $F \sim K(\theta^*; J) \times \mathrm{Pe}_\Delta$ for large Pe_Δ, where K is a constant (for upwind, $K = \frac{1}{2}$). A different way of looking at the results is to plot the stability region represented by Γ versus Pe_Δ. This is illustrated for one value of θ^* in Fig. 3. Here the actual gain in time step over that given by the FTCS scheme can be clearly observed.

One important limit of the transport equation is that of $\mathrm{Pe}_\Delta \to \infty$ (inviscid limit). In this case, the upwind scheme remains stable for Courant numbers c less than or equal to unity. For the FTCS scheme, we find that the limiting Courant number is zero, independent of θ^*. This limit is nonzero, however, for the QUICK and HOS formulations for θ^* greater than zero. The actual values of c can be obtained directly from (14ii) for $J > 0$. For example, for the QUICK scheme, $c \leqslant 0.048$, 0.012 for $\theta^* = 0.2\pi$ and 0.1π, respectively. For the HOS scheme, $c \leqslant 0.062$, 0.016 for $\theta^* = 0.2\pi$ and 0.1π, respectively.

In conclusion, the correct necessary and sufficient condition for the discretized explicit transport equation on a uniform grid is given by (12) and (13). In addition, the advantage of using either the QUICK or HOS formulation over central difference for the convective term has been illustrated, while still retaining second-order accuracy. As a footnote, we point out that the given formulations (except upwind)

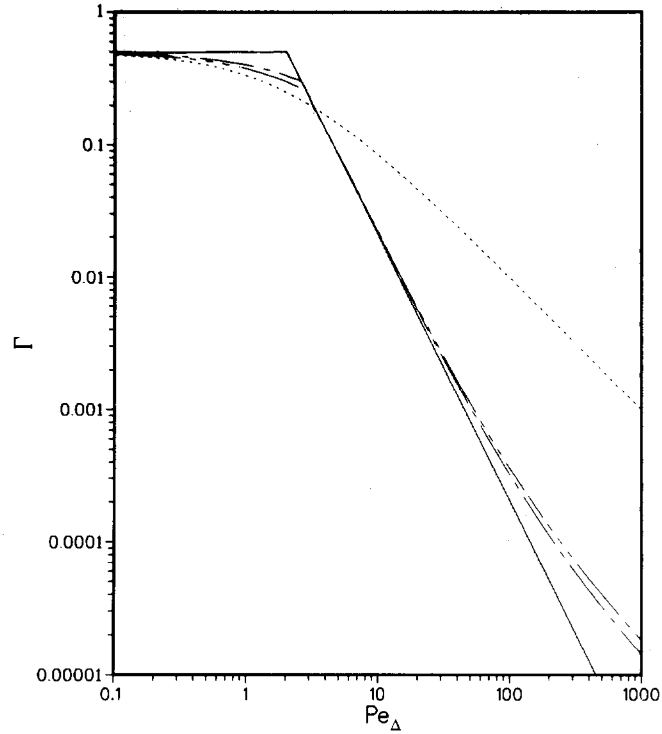

FIG. 3. Stability region for $\theta^* = (2\Delta x/l)\,\pi$, with $l/\Delta x = 20$, represented by Γ versus Pe_Δ; Upwind (---), Central (—), QUICK (——), HOS (——-—).

may lead to stable oscillatory results for $Pe_\Delta > 2$, 2.56, 2.77, for the FTCS, QUICK, and HOS schemes, respectively. This oscillatory behavior is well known to be related to local nonresolution of the physics somewhere in the computational domain. In *nonlinear* equations, these oscillations can result in a divergence of the solution. This *blow-up* is distinct, however, and not addressed by the von Neumann analysis (which only holds for the *linear* constant coefficients equation with periodic boundary conditions), for which our result is strictly correct for any $Pe_\Delta > 2$. Finally, we note that numerical stability is usually thought to be determined by the smallest resolvable wavelength in the problem; this is apparently due to limiting computations to $Pe'_\Delta \lesssim 2$. It was shown, however, that stability can be extended by considering the long wavelength cutoff in the numerical solution. The reason for this effect is that the minimum time step stability restriction is controlled by the highest frequency present in the numerical solution. It is typical in a physical problem that the smallest wavelengths are the ones responsible for the highest frequencies. In a numerical solution, however, when all relevant scales are not resolved, some of the energy which should have been attributed to the smallest scales reappears at the largest scales, resulting in a shift in the stability of the numerical problem. This shift can also be observed by looking at the truncation error terms which give rise to the highest frequency. These terms are due to diffusion for small Pe'_Δ, and to convection for large Pe'_Δ. When this occurs, the finite difference solution may somewhere violate monotonicity and/or boundedness conditions of the continuum problem. The standard way of eliminating this behavioral error of wiggles is to reduce the effective grid Peclét number, as done by the upwind scheme. It is not necessarily true, however, that the resulting oscillation-free solution is more accurate than the one containing the oscillations because the amplitude of the oscillations may be small and thus may not corrupt the solution.

References

1. R. M. Clancy, *Mon. Weather Rev.* **109** (1981), 1807.
2. J. E. Fromm, *Methods Comput. Phys.* **3** (1964), 345.
3. B. P. Leonard, *Comput. Methods Appl. Mech. Eng.* **19** (1979), 59.
4. B. P. Leonard, *Appl. Math. Modelling* **4** (1980), 401.
5. P. J. Roache, "Computational Fluid Dynamics," Hermosa, Albuquerque, N. M., 1976.
6. R. F. Warming and B. J. Hyett, *J. Comput. Phys.* **14** (1974), 159.

Received: December 29, 1981

Samuel Paolucci and Donald R. Chenoweth

Thermal Sciences Division,
Sandia National Laboratories,
Livermore, California 94550