JOURNAL OF COMPUTATIONAL PHYSICS **111**, 260–265 (1994)

On Destabilizing Implicit Factors in Discrete Advection–Diffusion Equations

JEAN-MARIE BECKERS*

GHER, Mécanique des fluides géophysiques, Sart Tilman B5, University of Liège, B-4000 Liège, Belgium

Received December 22, 1992; revised July 30, 1993

In the present paper, we find necessary and sufficient stability conditions for a simple one-time step finite difference discretization of an N-dimensional advection–diffusion equation. Furthermore, it is shown that when the implicit factors differ in each direction, a strange behavior occurs: By increasing one implicit factor in only one direction, a stable scheme can become unstable. It is thus suggested to use a single implicit direction (for efficient computing), or the same implicit factor in each direction. © 1994 Academic Press, Inc.

1. INTRODUCTION

In a recent paper [1], we found stability conditions for a discrete 3D advection–diffusion equation. We showed that a 3D scheme may need a fully implicit treatment in one direction if a 2D stability limit is attained in the remaining directions. In the present paper, we will generalize the study to an N-dimensional problem and find necessary and sometimes sufficient stability conditions.

Consider the advection–diffusion equation for a state-variable $y(t, x_i)$,

$$\frac{\partial y}{\partial t}+\sum_i u_i \frac{\partial y}{\partial x_i}=\sum_i \tilde{\kappa}_i \frac{\partial^2 y}{\partial x_i^2}, \tag{1}$$

where t is time, x_i are the space directions, u_i are the velocity components in the corresponding direction, and $\tilde{\kappa}_i$ are the diffusion coefficients.

If we apply a finite volume method with volumes of constant size on a system with constant velocities and diffusion coefficients, the discretization of Eq. (1) can be achieved by a centered scheme in space and an implicit Euler treatment in time. For this purpose, implicit factors α_i for advection and β_i for diffusion are introduced in such a way that the scheme reads[1]

$$\begin{aligned} y^{n+1} = y + \Delta t \sum_i \{ & -u_i[(1-\alpha_i)(y_{n_i+1}-y_{n_i-1})/(2\Delta x_i) \\ & +\alpha_i(y^{n+1}_{n_i+1}-y^{n+1}_{n_i-1})/(2\Delta x_i)] \\ & +\tilde{\kappa}_i[(1-\beta_i)(y_{n_i+1}+y_{n_i-1}-2y)/(\Delta x_i^2) \\ & +\beta_i(y^{n+1}_{n_i+1}+y^{n+1}_{n_i-1}-2y^{n+1})/(\Delta x_i^2)]\}, \end{aligned} \tag{2}$$

where n_i defines the n_ith point in direction x_i such that $x_i = n_i \Delta x_i$.

This scheme reduces to the classical FTCS (forward in time, centered in space) scheme for $\alpha_i = 0$ and $\beta_i = 0$, whereas for $\alpha_i = \frac{1}{2}$ and $\beta_i = \frac{1}{2}$ one retrieves the Crank–Nicolson method. The scheme is thus a generalization of well-established techniques. In the present paper, we focus on centered space derivatives, which gives a potentially very dispersive scheme, but the stability conditions that will be established remain valid for uncentered schemes that use upwind techniques and are thus less dispersive. Indeed, as it has been shown in [1], to generalize the stability conditions to these uncentered schemes it is sufficient to replace $\beta_i\tilde{\kappa}_i$ by $\beta_i\tilde{\kappa}_i+\alpha_i\kappa_i^{\text{num}}$ and $(1-\beta_i)\,\tilde{\kappa}_i$ by $(1-\beta_i)\,\tilde{\kappa}_i+(1-\alpha_i)\,\kappa_i^{\text{num}}$. The numerical diffusion κ_i^{num} depends upon the relative importance of the upwinding; for the classical upwind differencing, the numerical diffusion is for example given by $\kappa_i^{\text{num}} = u_i\,\Delta x_i/2$. For the purpose of a clear presentation we will thus concentrate on the centered version.

Using the Von Neumann method for stability analysis with the definitions

$$c_i \equiv \frac{u_i\,\Delta t}{\Delta x_i}, \qquad d_i \equiv 2\,\frac{\tilde{\kappa}_i\,\Delta t}{\Delta x_i^2}, \tag{3}$$

$$\begin{aligned} y &\equiv \rho \exp\left(I \sum_i n_i k_i\,\Delta x_i\right) \\ &= \rho \exp\left(I \sum_i n_i \theta_i\right); \qquad I^2=-1, \end{aligned} \tag{4}$$

* National Fund for Scientific Research (Belgium).

[1] Superscripts and subscripts are only written if they differ from ^{n} and $_{n_i}$.

0021-9991/94 $6.00

we can easily compute the complex amplification factor ρ,

$$\rho = \frac{X + IY}{U + IV},$$

$$X = 1 - \sum_i (1 - \beta_i)\, d_i (1 - \cos\theta_i)$$

$$Y = -\left[\sum_i (1 - \alpha_i)\, c_i \sin\theta_i\right] \tag{5}$$

$$U = 1 + \sum_i \beta_i d_i (1 - \cos\theta_i)$$

$$V = \sum_i \alpha_i c_i \sin\theta_i.$$

At this stage, by looking at a numerical mode in one direction, it is easy to show that an absolute necessary stability condition is $d_i \geqslant 0$. This reflects only the fact the forward integration of a problem with negative diffusion or backward integration of a system with positive diffusion is not a well-posed mathematical problem. We thus suppose $d_i \geqslant 0$, leading immediately to $U^2 + V^2 > 1$.

The definition of

$$\varepsilon(\theta_i) \equiv U^2 - X^2 + V^2 - Y^2 \tag{6}$$

allows us then to write the strong Von Neumann stability condition as follows (because $U^2 + V^2 \neq 0$):

$$|\rho| \leqslant 1 \Leftrightarrow \varepsilon \geqslant 0. \tag{7}$$

This condition will now be translated into conditions on the discretization constants of (3).

Unfortunately, we are not able to find a general necessary and sufficient stability condition. Indeed, in Section 2 we will find two necessary stability conditions in the general case. Then, in Section 3 we will show that these stability conditions are sufficient if $\alpha_i = \alpha$ and $\beta_i = \beta$. In the case where the implicit factors vary coordinate-wise, we are able to derive sufficient stability conditions that differ from the necessary stability conditions found in Section 2. It will then be shown that the necessary and sufficient stability conditions lead to the strange possibility of destabilizing an implicit scheme by increasing one implicit factor in one direction.

2. NECESSARY LOCAL STABILITY CONDITIONS

A simple method to find *necessary* stability conditions is to verify the behavior of the scheme for long and short waves. Indeed, if we look at the behavior of short waves (the numerical mode) in direction i, we have $\theta_i = \pi$. By using all combinations of very long waves and short waves, it is easy to establish that there exists a necessary stability condition only due to the diffusion part of the equation. This condition can be proved to be

$$\sum_i \max(0, (1 - 2\beta_i)\, d_i) \leqslant 1; \qquad d_i \geqslant 0. \tag{8}$$

For long waves in every direction $\theta_i = 0$, $\forall i$, we are at the stability limit because

$$\varepsilon = 0. \tag{9}$$

Furthermore, for these long waves[2]

$$\frac{\partial \varepsilon}{\partial \theta_i} = 0, \tag{10}$$

$$\frac{\partial^2 \varepsilon}{\partial \theta_i\, \partial \theta_j} = 2[d_i \delta_{ij} - (1 - \alpha_i - \alpha_j)\, c_i c_j]. \tag{11}$$

In order to guarantee that $\varepsilon \geqslant 0$, it is thus necessary that the matrix representing the tensor $\mathbf{H}$,

$$\mathbf{H} = \sum_i \sum_j \mathbf{e}_i \mathbf{e}_j [d_i \delta_{ij} - (1 - \alpha_i - \alpha_j)\, c_i c_j], \tag{12}$$

is positive definite.[3] This condition is thus another necessary stability condition.

3. NECESSARY AND SUFFICIENT CONDITIONS

Suppose for now that the same implicit factor is used in each direction: $\alpha_i = \alpha$, $\beta_i = \beta$. The tensor $\mathbf{H}$ can then be written as[4]

$$\mathbf{H} = \mathbf{B} \cdot \mathbf{H}' \cdot \mathbf{B}, \qquad \mathbf{H}' = \mathbf{I} - (1 - 2\alpha)\, \mathbf{s}\,\mathbf{s}, \tag{13}$$

$$\mathbf{s} \cdot \mathbf{e}_i = \frac{c_i}{\sqrt{d_i}}, \qquad \mathbf{e}_i \cdot \mathbf{B} \cdot \mathbf{e}_j = \delta_{ij} \sqrt{d_i}, \tag{14}$$

and we suppose that $d_i > 0$; otherwise, only the necessary condition $\alpha \geqslant \frac{1}{2}$ remains. In the case $d_2 > 0$ the quadratic form associated with the tensor $\mathbf{H}$ and a vector $\mathbf{z}'$ is

$$\mathbf{z}' \cdot \mathbf{H} \cdot \mathbf{z}' = \mathbf{z} \cdot \mathbf{H}' \cdot \mathbf{z} = 2\mathbf{z} \cdot \mathbf{z} - (1 - 2\alpha)(\mathbf{z} \cdot \mathbf{s})^2, \qquad \mathbf{z} = \mathbf{B} \cdot \mathbf{z}' \tag{15}$$

and it is positive $\forall \mathbf{z}$ if and only if

$$1 - (1 - 2\alpha)\, \mathbf{s} \cdot \mathbf{s} > 0. \tag{16}$$

[2] δ_{ij} is the classical Kronecker symbol: $\delta_{ij} = 0$, $i \neq j$; $\delta_{ij} = 1$, $i = j$.

[3] $\mathbf{e}_i$ is the unit vector in direction i, and $\mathbf{e}_i \mathbf{e}_j$ is the dyadic vector product of vectors $\mathbf{e}_i$ and $\mathbf{e}_j$.

[4] $\mathbf{I}$ is the identity tensor.

We can now rewrite necessary stability conditions (8) and (16):

$$(1-2\alpha)\sum_i \frac{c_i^2}{d_i} \leqslant 1, \tag{17}$$

$$(1-2\beta)\sum_i d_i \leqslant 1, \qquad d_i \geqslant 0. \tag{18}$$

Luckily enough, conditions (17) and (18) are also sufficient conditions! To demonstrate this conjecture, we use Eq. (6):

$$\begin{aligned}\varepsilon &= (U+X)(U-X)+(V+Y)(V-Y)\\ &= \left[2-(1-2\beta)\sum_i d_i(1-\cos\theta_i)\right]\left[\sum_i d_i(1-\cos\theta_i)\right]\\ &\quad -(1-2\alpha)\left(\sum_i c_i \sin\theta_i\right)^2. \end{aligned}\tag{19}$$

$\geqslant$ For $\alpha \geqslant \frac{1}{2}$, we have thus

$$\varepsilon \geqslant Z[2-(1-2\beta)Z]; \qquad Z \equiv \sum_i d_i(1-\cos\theta_i). \tag{20}$$

The definition of Z and the necessary stability condition (18) allow us to affirm that

$$Z \geqslant 0; \qquad (1-2\beta)Z \leqslant 2, \tag{21}$$

and thus $\varepsilon \geqslant 0$, which assures stability, as we may expect for a more than semi-implicit scheme like the Crank–Nicolson scheme.

$\geqslant$ If $\alpha < \frac{1}{2}$, Schwartz inequaliy and necessary condition (17) imply:

$$\begin{aligned}(1-2\alpha)\left(\sum_i c_i\sin\theta_i\right)^2 &\leqslant (1-2\alpha)\left(\sum_i \frac{c_i}{\sqrt{d_i}}\sqrt{d_i}\sin\theta_i\right)^2\\ &\leqslant (1-2\alpha)\left(\sum_i \frac{c_i^2}{d_i}\right)\left(\sum_i d_i\sin^2\theta_i\right)\\ &\leqslant \sum_i d_i\sin^2\theta_i\\ &\leqslant 4\sum_i d_i\sin^2\frac{\theta_i}{2} - 4\sum_i d_i\sin^4\frac{\theta_i}{2}. \end{aligned}\tag{22}$$

Using this last inequality we have

$$\varepsilon \geqslant 4\sum_i d_i\sin^4\frac{\theta_i}{2} - 4(1-2\beta)\left(\sum_i d_i\sin^2\frac{\theta_i}{2}\right)^2. \tag{23}$$

For $\beta \geqslant \frac{1}{2}$, we see immediately that $\varepsilon \geqslant 0$ and that the scheme is thus stable.

If $\beta < \frac{1}{2}$, Schwartz inequality and necessary condition (18) lead to

$$\begin{aligned}(1-2\beta)\left(\sum_i d_i\sin^2\frac{\theta_i}{2}\right)^2 &\leqslant (1-2\beta)\left(\sum_i \sqrt{d_i}\sqrt{d_i}\sin^2\frac{\theta_i}{2}\right)^2\\ &\leqslant (1-2\beta)\left(\sum_i d_i\right)\left(\sum_i d_i\sin^4\frac{\theta_i}{2}\right)\\ &\leqslant \left(\sum_i d_i\sin^4\frac{\theta_i}{2}\right). \end{aligned}\tag{24}$$

This demonstrates that $\varepsilon \geqslant 0$ if conditions (17) and (18) are satisfied.

Stability conditions (17) and (18) are thus *necessary and sufficient* stability conditions for the scheme (2) when $\alpha_i = \alpha$ and $\beta_i = \beta$. Our stability conditions are then a generalization of the well-known 1D conditions and the recently discovered conditions in 2D [2].

We found thus a necessary and sufficient stability condition in a case where Van Leer [3] thought it impossible. Furthermore, we correct a well-known error of Roache [4] and Fromm [5]. They postulate the existence of a stability condition based on the cell Reynolds number:

$$\frac{u\,\Delta x}{\tilde{\kappa}} \leqslant 2, \tag{25}$$

which in fact is neither a sufficient nor a necessary condition. Several authors (e.g., Leonard [6], Thompson *et al.* [7]) mention this error that is shown here to persist in N dimensions.

It is worth noting that the stability conditions have been obtained easily by a local analysis. Only afterwards did we prove that they are sufficient conditions as well.

4. DESTABILIZING IMPLICITNESS

Let us finally come back to the general discretization of Eq. (2). We showed already that, not very surprisingly, a more than semi-implicit scheme is unconditionally stable if the same implicit factor is used in each direction. In practice, once one decides to pay the price (in terms of computational requirements) for an implicit scheme in more than one direction, one would, of course, intuitively tend to this solution, but there could be reasons to use different implicit factors in each direction. If, for reasons of accuracy, for example, one would limit the implicitness differently according to the direction,[5] scheme (2) could be applied. It is then tempting to suppose that the conditions $\alpha_i \geqslant \frac{1}{2}$, $\beta_i \geqslant \frac{1}{2}$ assure stability. We will show that they will *not* guarantee

[5] In this case, the implicit factors would probably be chosen by some functional dependencies on discretization constants like $\alpha_i = \alpha_i(d_k, c_k)$, $\beta_i = \beta_i(d_k, c_k)$.

stability; even worse, there could be situations where an increase of implicitness could destabilize the scheme.

This can be proved by the following analysis. Let again write

$$\begin{aligned}\varepsilon &= (U+X)(U-X)+(V+Y)(V-Y)\\ &= \Big\{2-\sum_i (1-2\beta_i)\, d_i(1-\cos\theta_i)\Big\}\Big\{\sum_j d_j(1-\cos\theta_j)\Big\}\\ &\quad+\Big\{\sum_i c_i\sin\theta_i\Big\}\Big\{\sum_j (2\alpha_j-1)\, c_j\sin\theta_j\Big\}\\ &= \sum_i\sum_j \{2\delta_{ij}\, d_j(1-\cos\theta_j)\\ &\quad-(1-2\beta_i)\, d_i d_j(1-\cos\theta_i)(1-\cos\theta_j)\\ &\quad+c_i c_j(2\alpha_j-1)\sin\theta_i\sin\theta_j\}. \end{aligned} \tag{26}$$

Using the following decomposition,[6]

$$2d_j\delta_{ij}(1-\cos\theta_j)=\sqrt{d_i}\sqrt{d_j}\,\delta_{ij}\{\sin\theta_i\sin\theta_j + (1-\cos\theta_i)(1-\cos\theta_j)\}, \tag{27}$$

we can write (26) as a sum of two quadratic forms:

$$\varepsilon=\varepsilon_1+\varepsilon_2, \tag{28}$$

$$\varepsilon_1=\mathbf{p}\cdot\mathbf{H}'\cdot\mathbf{p}, \tag{29}$$

$$\varepsilon_2=\mathbf{q}\cdot\mathbf{D}\cdot\mathbf{q}, \tag{30}$$

where $\mathbf{H}'$ is the tensor used in Eq. (13), while the other vectors and tensors are defined by

$$\mathbf{p}\cdot\mathbf{e}_i=\sqrt{d_i}\sin\theta_i, \tag{31}$$

$$\mathbf{q}\cdot\mathbf{e}_i=\sqrt{d_i}(1-\cos\theta_i), \tag{32}$$

$$\mathbf{e}_i\cdot\mathbf{D}\cdot\mathbf{e}_j=\delta_{ij}-\sqrt{d_i}\sqrt{d_j}(1-\beta_i-\beta_j). \tag{33}$$

A *sufficient* stability condition is thus that the matrixes representing the tensors $\mathbf{H}'$ and $\mathbf{D}$ are positive definite. These are only sufficient conditions because it is not necessary that $\varepsilon_1 \geqslant 0$ *and* $\varepsilon_2 \geqslant 0$. Furthermore, for ε_2 to be positive, the matrix representing the tensor $\mathbf{D}$ does not need to be positive definite, because the vectors $\mathbf{q}$ are not arbitrary, as they have only positive components.

The conditions on $\mathbf{H}'$ and $\mathbf{D}$ are thus only sufficient conditions. Interestingly enough, the condition on tensor $\mathbf{H}'$ is a part of the necessary conditions. The difference between the necessary conditions and the sufficient conditions is thus linked to the diffusion part through the tensor $\mathbf{D}$. To illustrate the stability conditions, let us examine a 2D problem.

[6] We definetely suppose that $d_i > 0$.

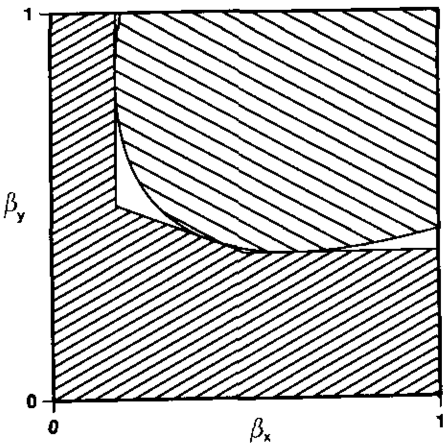

FIG. 1. The difference between necessary condition (35) and sufficient condition (38) shown as a function of implicit factors β_x, β_y for $d_x=\frac{3}{2}$, $d_y=4$. The narrowly hatched region corresponds to condition (35) and is thus an instability region. The other hatched region corresponds to condition (38) is thus a stability region provided that (34) is satisfied. Between the two, nothing can be stated about stability.

In the 2D case, the necessary stability condition associated with the tensor of Eq. (12) and condition (8) lead, indeed, to the *necessary* conditions[7]

$$(1-2\alpha_x)r_x+(1-2\alpha_y)r_y+(\alpha_x-\alpha_y)^2 r_x r_y \leqslant 1, \tag{34}$$

$$\max(0,(1-2\beta_x)\,d_x)+\max(0,(1-2\beta_y)\,d_y)\leqslant 1, \tag{35}$$

where

$$r_x\equiv\frac{c_x^2}{d_x},\qquad r_y\equiv\frac{c_y^2}{d_y}. \tag{36}$$

On the other hand, the *sufficient* conditions read

$$(1-2\alpha_x)r_x+(1-2\alpha_y)r_y+(\alpha_x-\alpha_y)^2 r_x r_y \leqslant 1, \tag{37}$$

$$(1-2\beta_x)d_x+(1-2\beta_y)d_y+(\beta_x-\beta_y)^2 d_x d_y \leqslant 1. \tag{38}$$

It is readily verified that if (38) is satisfied then (35) is also satisfied. The difference between the two conditions is depicted on Fig. 1 for $d_x=\frac{3}{2}$, $d_y=4$. Provided that (34) is satisfied, satisfying (38) assures stability, whereas violating (35) leads to instability. Between the two conditions, nothing can be said in general, but in practice, a fine-scale parametric scanning of the amplification factor allows us to compute the real stability limit. On Fig. 2, we show the computation of these limits for the same numerical values used in the other examples with $\alpha_x=1$, $\alpha_y=0.8$ (which assures us

[7] The condition on tensor $\mathbf{H}$ of Eq. (12) provides also the necessary conditions $(1-2\alpha_x)r_x\leqslant 1$ and $(1-2\alpha_y)r_y\leqslant 1$, but it is easily shown that they are less constraining than (34).

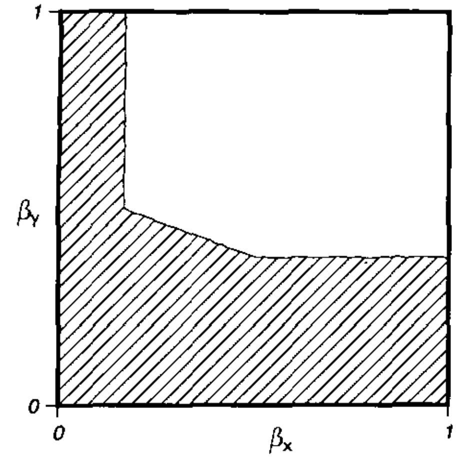

FIG 2. Numerical computation of the stability condition as a function of β_x, β_y for $d_x=\frac{3}{2}$, $d_y=4$.

that (34) is satisfied). It turns out that the curve found numerically is very close to the necessary stability condition.

But suppose now that we even satisfy sufficient condition (38). We are thus sure that condition (34) is the separation between a stable and an unstable scheme because when it is not satisfied, a necessary condition is not satisfied (instability), whereas if condition (34) is satisfied, all the sufficient conditions are satisfied (stability). We can then see a strange behavior: in the (α_x, α_y) plane, the stability limit (34) describes a parabola (shown on Fig. 3 for $r_x=\frac{15}{2}$, $r_y=\frac{9}{2}$). If we increase one implicit factor in one direction, but fix the second one, the scheme may become unstable although it was stable for smaller implicitness. It can even be unstable for $\alpha_i \geqslant \frac{1}{2}$!

Increasing the implicit factor in only one direction may thus change a stable scheme into an unstable one. This unexpected behavior stems from the special form of the amplification factor. Indeed, if the implicit factors are not identical, their influence on the numerator and denominator of (5) are not as straightforward as in the case where they are identical.

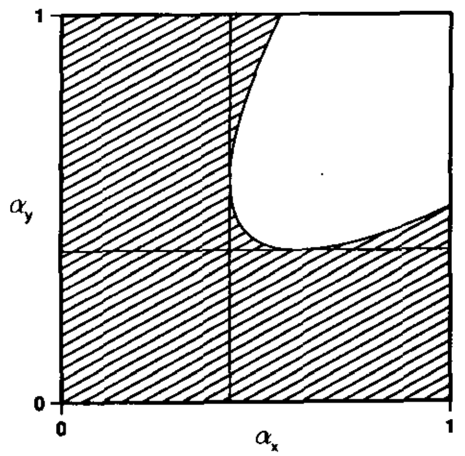

FIG. 3. The stability domain (white) given by (34) as a function of implicit factors α_x, α_y for $r_x=\frac{15}{2}$, $r_y=\frac{9}{2}$. The horizontal and vertical lines correspond to the necessary 1D stability conditions.

A similar necessary stability condition can be written for the 3D scheme and the condition on the tensor **H** is

$$\begin{aligned}(1-2\alpha_x)r_x+(1-2\alpha_y)r_y+(1-2\alpha_z)r_z\\ +(\alpha_x-\alpha_y)^2 r_x r_y+(\alpha_x-\alpha_z)^2 r_x r_z\\ +(\alpha_z-\alpha_y)^2 r_y r_z \leqslant 1. \qquad (38)\end{aligned}$$

This condition allows also for a destabilizing effect of the implicitness.

Even more, for $\beta_x=\beta_y$, conditions (35) and (38) are identical, and in N dimensions, if $\beta_i=\beta_j$, then condition (8) and the condition on tensor $\mathbf{H}'$ are necessary and sufficient conditions; our result on the destabilizing implicitness is thus also demonstrated in this case.

5. DISCUSSION

By analyzing the stability of a classical one-time-step discretization of an N-dimensional advection–diffusion equation, we were able to find necessary and sufficient stability conditions when the same implicit factors are used in each direction. When different implicit factors are used in each direction, the local analysis provides generally only necessary stability conditions, which show, however, that the scheme can become unstable when only one implicit factor is increased.

Whether or not our result is peculiar to the scheme analyzed here, is not clear. In our opinion, similar strange behaviors may be found in other schemes that favor the treatment in one direction compared to the others. Our study suggests, then, a more careful look at the behavior of such schemes.

The Von Neumann stability analysis does, of course, not take into account the boundary conditions and finite domains. From this point of view, the stability of the scheme can be modified by the type and peculiar implementation of the boundary conditions. This is, however, a general remark on stability conditions and is not characteristic of the scheme analyzed here. The observation of the strange behavior, when increasing one implicit factor, is thus not necessarily observable in practice, because instabilities due to boundary conditions could hide this effect or the boundary conditions could stabilize the scheme.

ACKNOWLEDGMENTS

The author gratefully acknowledges the financial support of the European Communities in the scope of the research programs MAS2-CT92-0041 (EUROMODEL). Supercomputer facilities and technical assistance were supplied by IBM and the National Fund for Scientific

Research (Belgium). Thanks are due to Kevin G. Ruddick who made a significant contribution by suggesting one of the demonstrations of sufficient conditions.

REFERENCES

1. J. M. Beckers, *SIAM J. Numer. Anal.* **29**, 701 (1992).
2. A. Rigal, *Int. J. Numer. Methods Fluids* **13**, 579 (1991).
3. B. Van Leer, *J. Comput. Phys.* **3**, 473 (1969).
4. P. J. Roache, Computational fluid dynamics (Hermosa Alburquerque, New Mexico, 1976).
5. J. E. Fromm, *Methods Comput. Phys.* **3**, 345 (1964).
6. B. P. Leonard, *Appl. Math. Modelling* **4**, 401 (1980).
7. H. D. Thompson, B. W. Webb, and J. D. Hoffmann, *Int. J. Numer. Methods Fluids* **5**, 305 (1985).