Improving the Convergence Rate to Steady State of Parabolic ADI Methods

SAUL S. ABARBANEL*

Tel Aviv University, Tel-Aviv, Israel and Massachusetts Institute of Technology, Cambridge, Massachusetts 02132

DOUGLAS L. DWOYER

NASA Langley Research Center, Hampton, Virginia 23665

AND

DAVID GOTTLIEB†

Tel Aviv University, Tel Aviv, Israel and Institute for Computer Applications in Science and Engineering

Received March 2, 1983; revised April 29, 1985

The rate of convergence to steady state of parabolic ADI solvers is analyzed in terms of the L_2-norms of the residuals. The analysis allows one to predict the number of iterations necessary for convergence as function of the Courant number. λ. A simple modification of existing ADI codes is devised. It improves the convergence rate substantially and is insensitive to the Courant number in a large range of λ.

INTRODUCTION

ADI methods for elliptic partial differential equations were proposed already in the 1950s. Particularly well known were the algorithms due to Peaceman and Rachford [11], and Douglas and Gunn [6]. In recent years, ADI methods were advocated to solve parabolic and hyperbolic p.d.e.'s, see Beam and Warming [2], and Briley and McDonald [4]. The motivation for ADI schemes was to combine the convenience of one-dimensional easily invertible operators and the unconditional stability of implicit methods.

* Research was supported in part by the U.S. Air Force Office of Scientific Research under Grant AFOSR-80-0249 and in part by the NASA Cooperative Agreement NCC1-45.

† Research was supported in part by the U.S. Army Research and Standardization Group (Europe) under Contract DAJA 38-80-C-0032 and in part by NASA Contract NAS 1-15810 while the author was in residence at ICASE, NASA Langley Research Center, Hampton, Virginia 23665.

0021-9991/86 $3.00

Experience with such ADI solvers, both for the Navier–Stokes, and Euler equations [9, 13] has not shown the large expected gains over direct explicit finite-difference algorithms. This is so because in approaching steady state the rate of convergence is very sensitive to the Courant number (appropriately defined for each case) and it decreases rapidly, i.e., the iteration count grows rapidly, when the calculation is carried out away from an optimal time-step. At the optimal Courant number, the convergence rate is comparable to those of explicit methods.

A procedure of using an optional sequence of time steps has been used by previous researchers [6, 10] to greatly accelerate the convergence to steady state of the heat equation. It is not at all clear that this procedure can be applied efficaciously to more complicated problems such as nonlinear systems of equation. We sought to make use of a preconditioning procedure whose rationale is immediately applicable, not only to the scalar parabolic problem, but to systems as well. The new preconditioned algorithm can be further accelerated to steady state by using the above-mentioned technique of sequencial times-step even if they are not the optimal ones for the unpreconditioned problem.

The present research concentrates on the parabolic case. A subsequent paper will deal with hyperbolic partial differential equations. In this work we set out to:

(a) Analyze the behaviour of the standard ADI algorithms in converging to steady state. The number of time-steps (or iterations) necessary to converge to steady state turns out to be most easily analyzed in terms of the square of the ratio of the L_2-norm of the residual to the L_2-norm of the initial residual. The underlying hypothesis of this analysis is that the residual attenuation is relatively independent of the initial and boundary conditions, provided they are of the (common) types that introduce many frequencies into the problem. The results of this research substantiate this hypothesis. All this is covered in Section 1.

(b) Devise a new algorithm which will "correct" the standard ADI schemes in the sense that it will speed up convergence to steady state without changing the accuracy of the solution and having only a weak dependence on the Courant number λ, for λ larger than its optimum value for the standard scheme. Also the new algorithm should leave the implicit part of the scheme unchanged. This is done in Section 2.

(c) Test the new corrected algorithm for a variety of situations: Dirichlet problems for uniform grids of many mesh sizes; mixed Dirichlet–Neumann problems; problems defined on stretched grids and/or problems with variable coefficients. The results of the numerical experimentations are described in Section 3.

1

Consider the standard parabolic diffusion equation in two space dimensions

$$\frac{\partial u}{\partial t}=\nu\left(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}\right), \qquad 0\leqslant x\leqslant a, 0\leqslant y\leqslant a, 0\leqslant t. \tag{1.1}$$

Using the usual scaling procedures we get the following problem on a square,

$$\frac{\partial u}{\partial t}=\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}, \qquad 0\leqslant x,\ y\leqslant \pi, \qquad 0\leqslant t,$$
$$u(x,\ y,\ 0)=u_0(x,\ y), \tag{1.2}$$

with u (independent of t) prescribed on the boundaries. We are interested in the steady-state solution to (1.2).

A widely used numerical method for reaching the steady-state solution, $u^*(x,\ y)$, of (1.2) is to employ the unconditionally stable backward Euler algorithm in delta form with approximate factorization (AF) (see, e.g., Beam and Warming [2]). This finite-difference scheme may be written as follows

$$(1-\lambda\delta_x^2)(1-\lambda\delta_y^2)(v_{j,k}^{(n+1)}-v_{j,k}^{(n)})=\alpha\lambda(\delta_x^2+\delta_y^2)\,v_{j,k}^{(n)} \tag{1.3}$$

with $v_{0j,k}^{(0)}=U_0(j\,\Delta x,\ k\,\Delta y)$, $0\leqslant\alpha\leqslant 2$, and where $v_{j,k}^{(n)}=v(j\,\Delta x,\ k\,\Delta y,\ n\,\Delta t)$ is the finite-difference approximation to $u(x,\ y,\ t)$ and

$$\lambda=\Delta t/h^2, \qquad h=\Delta x=\Delta y=\pi/N.$$

The difference operators δ_x^2 and δ_y^2 are given by

$$\delta_x^2 v_{j,k}^{(n)}=v_{j+1,k}^{(n)}-2v_{j,k}^{(n)}+v_{j-1,k}^{(n)}$$

and

$$\delta_y^2 v_{j,k}^{(n)}=v_{j,k+1}^{(n)}-2v_{j,k}^{(n)}+v_{j,k-1}^{(n)}.$$

The case $\alpha=1$ is known as the Douglas-Gunn scheme; $\alpha=2$ is the Peaceman–Rachford algorithm.

We now define the deviation $\varepsilon_{j,k}^{(n)}$ of the evolving solution $v_{j,k}^{(n)}$ from the desired finite-difference approximation to the steady state, $v_{j,k}^*$. The function v^* satisfies the equation

$$\frac{1}{h^2}(\delta_x^2+\delta_y^2)\,v_{j,k}^*=0 \tag{1.4}$$

with v^* having boundary values which correspond appropriately to the boundary values of u^*. This deviation is given, therefore, at any time $t=n\,\Delta t$ by

$$\varepsilon_{j,k}^{(n)}=v_{j,k}^{(n)}-v_{j,k}^*. \tag{1.5}$$

At this point, we introduce the assumptions that $v_{j,k}^{(n)}=v_{j,k}^*$ on the boundaries and that the initial values of u, and hence also of v, are such that on the boundaries they coincide with the prescribed boundary conditions. Note that these are reasonable assumptions that agree with standard procedures for problems such as (1.2). A consequence of these assumptions is that ε^n for every $n\geqslant 0$ vanishes on the boundaries.

Using Eqs. (1.3), (1.4), and (1.5) we find that ε^n satisfies

$$(1-\lambda\delta_x^2)(1-\lambda\delta_y^2)(\varepsilon_{j,k}^{(n+1)}-\varepsilon_{j,k}^{(n)})=\alpha\lambda(\delta_x^2+\delta_y^2)\,\varepsilon_{j,k}^{(n)} \tag{1.6}$$

with homogenous boundary conditions. Since ε^n is zero on the boundaries: $\{x=0,$ $0\leqslant y\leqslant \pi;\ x=\pi,\ 0\leqslant y\leqslant\pi;\ 0\leqslant x\leqslant\pi,\ y=0;\ 0\leqslant x\leqslant \pi,\ y=\pi\}$, it may be represented as an interpolation polynomial in the form of a finite double-sine series. The explicit representation is

$$\varepsilon_{j,k}^{(n)}=\varepsilon^{(n)}(x_j,y_k)=\sum_{q=1}^{N-1}\sum_{p=1}^{N-1}A_{p,q}^{(n)}\sin(px_j)\sin(qy_k),\qquad 1\leqslant j,k\leqslant N-1, \tag{1.7}$$

where $N=\pi/h$ is the number of mesh intervals in either direction and hence $x_j=j\pi/N$, $y_k=k\pi/N$. The coefficients $A_{p,q}^{(n)}$ are given by

$$A_{p,q}^{(n)}=\frac{4}{N^2}\sum_{j=1}^{N-1}\sum_{k=1}^{N-1}\varepsilon_{j,k}^{(n)}\sin(j\pi p/N)\sin(k\pi q/N),\qquad 1\leqslant p,q\leqslant N-1. \tag{1.8}$$

The identity (1.8) is easily established by substituting (1.7) into (1.8) and using the formula

$$\sum_{p=1}^{N-1}\sum_{q=1}^{N-1}\sin\left(\frac{j\pi p}{N}\right)\sin\left(\frac{k\pi q}{N}\right)\sin\left(\frac{p\pi l}{N}\right)\sin\left(\frac{q\pi m}{M}\right)=\frac{N^2}{4}\delta_{lj}\,\delta_{mk}, \tag{1.9}$$

where

$$\delta_{\mu\nu}=\begin{cases}1, & \mu=\nu\\ 0, & \mu\neq\nu.\end{cases} \tag{1.10}$$

Next we define the L_2-norm of $\varepsilon_{j,k}^n$ in the following way:

$$\|\varepsilon_{j,k}^{(n)}\|^2=\frac{1}{N^2}\sum_{j=1}^{N-1}\sum_{k=1}^{N-1}|\varepsilon_{j,k}^{(n)}|^2. \tag{1.11}$$

Substituting (1.7) into (1.11) and using (1.9) gives

$$\|\varepsilon_{j,k}^{(n)}\|^2=\frac{1}{N^2}\sum_{j=1}^{N-1}\sum_{k=1}^{N-1}|\varepsilon_{j,k}^{(n)}|^2=\frac{1}{4}\sum_{p=1}^{N-1}\sum_{q=1}^{N-1}[A_{p,q}^{(n)}]^2. \tag{1.12}$$

The formula (1.12) is basically Parseval's relation for the present case.

At this point, we would like to motivate the manner in which to investigate the convergence to steady state of the algorithm (1.3). Upon evolution (n increasing) the steady-state term

$$R_n=(1/h^2)(\delta_x^2+\delta_y^2)\,v_{j,k}^{(n)}=(N^2/\pi^2)(\delta_x^2+\delta_y^2)\,v_{j,k}^{(n)}$$

should decrease towards zero and its norms should behave similarly. In running practical codes, practitioners monitor some kind of a norm of what is known as the residual; i.e., the steady-state equation. Often the maximum norm is utilized. We propose to use the L_2-norm for the following reasons:

(a) it is more efficient on parallel computers while comparable in implementation on scalar machines,

(b) stability analysis usually uses the L_2-norm,

(c) the convergence analysis to be presented is conveniently carried out for the L_2-norm, and

(d) we do not know how to carry out the analysis for the maximum norm. Note that we are dealing with a different kind of convergence than the classical one. In the classical case, one considers the norm of $v_{j,k}^{(n)} - u(j\,\Delta x, k\,\Delta y, n\,\Delta t)$, as $\Delta x, \Delta y \to 0$ (i.e., $N \to \infty$) $\forall n\,\Delta t \leqslant T$. In the present case, however, we base the analysis on the deviation $\varepsilon_{j,k}^{(n)}$ as defined in (1.5), for N fixed as n increases. In these circumstances, the L_2 and the maximum norms are mathematically equivalent.

Because of Eq. (1.5), we have $(\delta_x^2 + \delta_y^2)\, v_{j,k}^{(n)} = (\delta_x^2 + \delta_y^2)\, \varepsilon_{j,k}^{(n)}$. We now evaluate the L_2-norm of the residual using this relation, Eq. (1.7) and the definitions of the operator δ_x^2, δ_y^2,

$$\begin{aligned}(\delta_x^2 + \delta_y^2)\, v_{j,k}^{(n)} &= (\delta_x^2 + \delta_y^2)\, \varepsilon_{j,k}^{(n)} \\ &= -4 \sum_{p=1}^{N-1} \sum_{q=1}^{N-1} A_{p,q}^{(n)} \left[\sin^2\left(\frac{\pi p}{2N}\right) + \sin^2\left(\frac{\pi q}{2N}\right)\right]\left[\sin\frac{\pi p j}{N} \sin\frac{\pi q k}{N}\right].\end{aligned}$$

Using (1.11) and (1.12) this becomes

$$\|(\delta_x^2 + \delta_y^2)\, \varepsilon_{j,k}^{(n)}\|^2 = 4 \sum_{p=1}^{N-1} \sum_{q=1}^{N-1} \left\{ A_{p,q}^{(n)} \left[\sin^2\left(\frac{\pi p}{2N}\right) + \sin^2\left(\frac{\pi q}{2N}\right)\right]\right\}^2. \tag{1.13}$$

At this point, it is necessary to establish the connection between $A_{p,q}^{(n)}$ and $A_{p,q}^{(0)}$. We do this by substituting (1.17) into (1.6) and operating with δ_x^2 and δ_y^2 to obtain

$$\sum_{j=1}^{N-1} \sum_{k=1}^{N-1} \{(1 + 4\lambda\xi_p^2)(1 + 4\lambda\eta_q^2)[A_{p,q}^{(n+1)} - G_{p,q} A_{p,q}^{(n)}]\} \sin\frac{\pi p j}{N} \sin\frac{\pi k q}{N} = 0, \tag{1.14}$$

where

$$\xi_p^2 = \sin^2(\pi p/2N), \qquad \eta_q^2 = \sin(\pi q/2N)$$

and

$$G_{p,q} = \frac{1 + \sigma^2 \xi_p^2 \eta_q^2 + (1-\alpha)\,\sigma(\xi_p^2 + \eta_q^2)}{(1 + \sigma\xi_p^2)(1 + \sigma\eta_q^2)} \tag{1.15}$$

with $\sigma = 4\lambda$. For (1.14) to be satisfied we must have

$$A_{p,q}^{(n+1)} = G_{p,q} A_{p,q}^{(n)}, \qquad \forall 1 \leqslant p, q \leqslant N-1 \tag{1.16}$$

Consequently,

$$A_{p,q}^{(n)} = G_{p,q}^{n} A_{p,q}^{(0)}. \tag{1.17}$$

This is the relation we were seeking and upon substituting into (1.13) we have

$$\|R_n\|^2 = \left\| \left(\frac{\delta_x^2 + \delta_y^2}{h^2} \right) \varepsilon_{j,k}^{(n)} \right\|^2 = \frac{4N^4}{\pi^4} \sum_{p=1}^{N-1} \sum_{q=1}^{N-1} G_{p,q}^{2n} \left[A_{p,q}^{(0)} \left(\sin^2 \frac{\pi p}{2N} + \sin^2 \frac{\pi q}{2N} \right) \right]^2. \tag{1.18}$$

Note, from (1.15), that if we let ξ_p^2 and η_q^2 vary continuously (i.e., $N \to \infty$) then $G_{p,q} = G$, the amplification factor for the pure initial value problem. For the sake of completness, and also because

$$\varepsilon^{(n)} \to 0 \qquad \text{as} \quad (\delta_x^2 + \delta_y^2)\, \varepsilon^{(n)} \to 0$$

we rewrite the norm of $\varepsilon^{(n)}$ equation (1.12) in term of the $A_{p,q}^{(0)}$'s; i.e.,

$$\|\varepsilon_{j,k}^{(n)}\|^2 = 4 \sum_{p=1}^{N-1} \sum_{q=1}^{N-1} G_{p,q}^{2n} [A_{p,q}^{(0)}]^2. \tag{1.19}$$

It is clear from (1.18) and (1.15) that the norm of the residual depends on the number of grid points, N the number of time-steps, n, the Courant number λ and the initial deviations as reflected in $A_{p,q}^{(0)}$. Of course the $A_{p,q}^{(0)}$'s are not known a priori, and are in a sense arbitrary. The deviations vanish on the boundaries, however, and so it is reasonable to consider a category that generalizes to our two-dimensional case the triangular initial distribution used by Richtmyer and Morton [1], in their expository first chapter. This family is characterized by having Fourier coefficients $A_{p,q}^{(0)}$ $(1 \leqslant p, q \leqslant N-1)$ whose magnitude is given by

$$|A_{p,q}^{(0)}| = \begin{cases} 0 & \text{for some } p, q \\ \dfrac{K}{N^2} \cdot \dfrac{1}{\sin^2(p\pi/2N) + \sin^2(q\pi/2N)} & \text{for rest of } p, q, \end{cases} \tag{1.20}$$

where K is a constant. Two examples of members of this family are given by the regular square-based pyramid (see Fig. 1) of height C,

$$A_{p,q}^{(0)} = \begin{cases} 0, & p \neq q, \text{ or } p = q \text{ even} \\ \dfrac{2C}{N^2} \cdot \dfrac{1}{\sin^2(\pi p/2N)}, & p = q \text{ odd}. \end{cases} \tag{1.21}$$

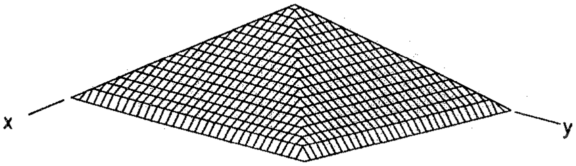

FIG. 1. Initial condition for test problem.

and by a square-skirted cone of height C

$$A_{p,q}^{(0)}=\begin{cases} 0, & p,\ q \text{ even} \\ (-)^{(p+q)/2}\dfrac{2C}{N^2}\cdot\dfrac{1}{\sin^2(\pi p/2N)+\sin^2(\pi q/2N)} & \end{cases} \tag{1.22}$$

substituting (1.21) and (1.22) into (1.18), we get for the norms of the residuals the following expressions, respectively,

$$\|R_n\|^2=\frac{16C^2}{\pi^2}\sum_{\substack{p=1\\ p\text{ odd}}}^{N-1} G_{p,p}^{2n} \tag{1.23}$$

and

$$\|R_n\|^2=\frac{16C^2}{\pi^2}\sum_{\substack{p=1\\ p\text{ odd}}}^{N-1}\sum_{\substack{q=1\\ q\text{ odd}}}^{N-1} G_{p,q}^{2n}. \tag{1.24}$$

We are most interested in the attenuation rate of the L_2-norm of the residual, given by

$$\rho_n^2=\frac{\|R_n\|^2}{\|R_0\|^2}. \tag{1.25}$$

For the initial deviations described by (1.21) and (1.22) we get, respectively,

$$\rho_n^2=\frac{2}{N}\sum_{\substack{p=1\\ \text{odd}}}^{N-1} G_{p,p}^{2n} \tag{1.26}$$

and

$$\rho_n^2=\frac{4}{N^2}\sum_{\substack{p=1\\ \text{odd}}}^{N-1}\sum_{\substack{q=1\\ \text{odd}}}^{N-1} G_{p,q}^{2n}. \tag{1.27}$$

Note, from Holders' inequality, that

$$\rho_n^2\geqslant\rho_1^{2n}. \tag{1.28}$$

This provides us with an upper bound on the rate of attenuation of the norm of the residual per time-step.

Let us first start by considering the properties of ρ_1^2 for the Douglas–Gunn case, $\alpha=1$. For the family of initial conditions given by (1.22), we write the expression for ρ_1^2, using (1.27) and (1.15),

$$\rho_1^2=\frac{4}{N^2}\sum_{\substack{p=1\\ \text{odd}}}^{N-1}\sum_{\substack{q=1\\ \text{odd}}}^{N-1}\left[\frac{1+\sigma^2\sin^2\theta_p\sin^2\psi_q}{(1+\sigma\sin^2\theta_p)(1+\sigma\sin^2\psi_q)}\right]^2, \tag{1.29}$$

where

$$\theta_p=\frac{\pi p}{2N}, \qquad \psi_q=\frac{\pi q}{2N}, \qquad \sigma=4\lambda.$$

It is shown in [1, Appendix 1] that the summation in (1.29) can be carried out explicitly and be expressed in terms of elementary functions as follows:

$$\rho_1^2=[(1+\sigma)^{-1/2}\tanh(N\zeta/2)-\Lambda_1]^2+[1-\Lambda_1-(1+\sigma)^{-1/2}\tanh(N\zeta/2)]^2+2\Lambda_1^2, \tag{1.30}$$

where

$$\Lambda_1=(1/2)\,\sigma(1+\sigma)^{-3/2}\tanh(N\zeta/2)+(1/2)\,N(L+\sigma)^{-1}[\cosh(N\zeta/2)]^{-2} \tag{1.31}$$

and

$$\zeta=2\tanh^{-1}[(1+\sigma)^{-1/2}]. \tag{1.32}$$

It may be verified that ρ_1^2 is a positive concave function taking the value of unity at $\sigma=0$ and $\sigma=\infty$. Numerical evaluations of (1.30) have shown that the location of the minimum hardly changes for $N\geqslant 4$. In fact, a very accurate prediction for this minimum may be found by letting $N\to\infty$ in (1.30). One then finds

$$\lim_{N\to\infty}\rho_1^2=1-3z+3z^2+z^3-2z^4+z^6 \qquad (z=(1+\sigma)^{-1/2}). \tag{1.33}$$

The minimum of ρ_1^2 is found at $z=0.5128685$ or $\sigma=2.801788$.

The value of ρ_1^2, which is the attenuation rate per time-step of the norm of the residual, at that point, is $\|R_1\|^2_{\min}/\|R_0\|^2=0.265223$.

The behaviour of ρ_1^2 shows that even though the AF-scheme is unconditionally stable it is not beneficial to use too large a time-step ($\sigma\gg 1$). This phenomenon is also known for hyperbolic p.d.e.'s [1].

A different analysis, for the diffusion equation case, was carried out by Wachspress (see [10] and references therein) who considered the behaviour of the solution (rather than the residual) in a single time-step. He investigated the quantity

$$\min_{\sigma}\max_{p,q}|G_{p,q}|$$

and showed for the Peaceman–Ratchford scheme, $\alpha = 2$, that for constant time-step the optimum λ is proportional to N. The difference in the results stems from the fact that we considered the behaviour of the norm of the residual for a particular family of initial deviations rather than the behaviour of the solution.

While the motivation for investigating ρ_1^2, see Eq. (1.28), is based on the idea of obtaining an upper bound on the rate of attenuation of the norm of the residual, there clearly are features of the behaviour of the residual that cannot be adequatley explained by ρ_1^2. We are thus constrained to consider also $(\alpha = 1)$

$$\rho_n^2 = \frac{4}{N^2} \sum_{\substack{p=1 \\ \text{odd}}}^{N-1} \sum_{\substack{q=1 \\ \text{odd}}}^{N-1} \left[\frac{1 + \sigma^2 \sin^2 \theta_p \sin^2 \psi_q}{(1 + \sigma \sin^2 \theta_p)(1 + \sigma \sin^2 \psi_q)} \right]^{2n}. \tag{1.34}$$

In [1], Appendix 1] it is shown that

$$\rho_n^2 = \sum_{l=0}^{2n} \binom{2n}{l} \Lambda_l^2, \tag{1.35}$$

where for $0 \leqslant l \leqslant 2n-1$,

$$\Lambda_l = -\frac{1}{(2n-1)!} \sigma^l \left\{ \frac{\partial^{2n-1}}{\partial \sigma^l \, \partial a^{2n-l-1}} \left[(a^2 + \sigma a)^{-1/2} \tanh Ne \right] \right\}_{a=1} \tag{1.36}$$

and for $l = 2n$,

$$\Lambda_{2n} = 1 - \sum_{j=0}^{2n-1} \binom{2n}{j} \Lambda_j \tag{1.37}$$

with

$$e = \tanh^{-1} \left[\left(1 + \frac{\sigma}{a}\right)^{-1/2} \right]. \tag{1.38}$$

We were not able to find a simpler exact analytic representation to ρ_n^2 than (1.35)–(1.38). However, a good approximation, uniformly valid in σ, N and n is derived in [1, Appendix 2]. Moreover, (1.35) is easily and speedily evaluated numerically for various ranges of the parameters σ, n, N. In particular, one may ask what is the number of iterations required to reduce the initial residual by a certain large factor 10^{ω}. Figure 2 shows the predictions of Eq. (1.35) (for the families of initial deviations) in comparison with values obtained using the AF-solver (1.3). The results obtained from the finite-difference equations (FDE) solver are insensitive to initial conditions. Details of the numerical experimentations are described in Section 3. One feature which the numerical work (both evaluating (1.35) and using the FDE solver) brought out was that for σ large enough, the number of iterations necessary to converge is independent of the mesh N. Another feature was that n, the number of steps necessary for convergence, increases linearly with σ for

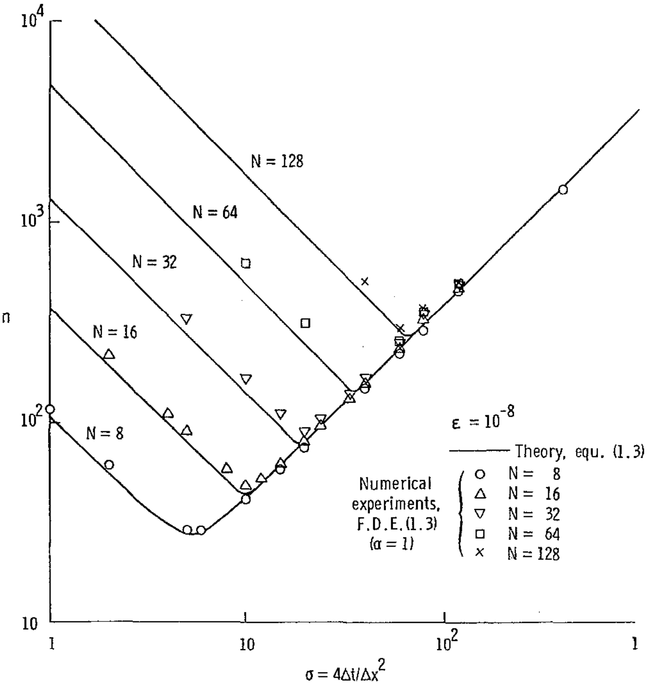

FIG. 2. Number of iterations to converge to steady state for approximate factorization implicit backward Euler scheme vs. Courant No. Comparison of prediction and numerical experiments.

large σ. This linear-dependence behaviour can be extracted directly from the derivation in [1, Appendix 2].

The behaviour and nature of the solution to the AF-solver is thus explained. It is clear that it is desirable to improve its performance, particularly for large σ. This topic is covered in the next section.

2

The problem of solving the Laplace operator is well known and ADI methods were proposed for it in the 1950s, see [10, 12]. Convergence to steady state was greatly improved by taking cyclical relaxation factors (or, in terms of the parabolic evolution operator, cyclical time-steps). Later multigrid methods [3], accelerated the convergence process. These improvements became more difficult to apply to various cases of practical importance such as variable and/or stretched mesh grids, mixed Neuman–Dirichlet boundary conditions, nonconstant coefficients, etc.

Because of the great effort already invested in existing AF-algorithms, we were looking for an algorithmic correction which would satisfy the following constraints:

(a) leave the implicit AF part of the algorithm unchanged,

(b) the correction to be introduced should be explicit and easy to program,

(c) the correction should be robust in the sense that it need not be fine-tuned for different mesh sizes, different grid stretchings, different mixes of the boundary conditions, etc.,

(d) generalization to the three-dimensional case should be obvious, and finally

(e) the number of iterations to convergence should be roughly the same for a large range of σ and be competitive with other methods of accelerating convergence.

It is clear from the analysis of the last section that the nonbenign behaviour of ρ_n^2 for large σ stems from the fact that as the Courant number $\lambda = \sigma/4$ increases, every $G_{p,q}$ for any fixed mesh approaches unity uniformly for all the frequencies p and q. In these circumstances ρ_n^2 will also approach unity as σ increases. This behaviour is due to the quadratic term, $\sigma^2 \sin^2 \theta_p \sin^2 \psi_q$, introduced by the AF. This fact has been noted before, see for example [5, 7].

The main idea was not to increase the efficiency of the AF-algorithm by devices such as cyclic time-step and multi-griding but rather by modifying the scheme itself. Note the form of $G_{p,q}$, see (1.15),

$$G_{p,q} = \frac{1 + \sigma^2 \xi_p^2 \eta_q^2 + (1-\alpha)\,\sigma(\xi_p^2 + \eta_q^2)}{(1+\sigma\xi_p^2)(1+\sigma\eta_q^2)}. \tag{2.1}$$

To bound $G_{p,q}$ away from unity, for large σ, we propose the following modification to (2.1),

$$G_{p,q}(\gamma) = \frac{1 + (1-\alpha)\,\sigma(\xi_p^2 + \eta_q^2) + \sigma^2 \xi_p^2 \eta_q^2 (1 - \gamma(\xi_p^2 + \eta_q^2))}{(1+\sigma\xi_p^2)(1+\sigma\eta_q^2)}, \tag{2.2}$$

where for stability $0 \leqslant \gamma \leqslant 1$. It remains to be seen whether one can predetermine an optimum value of γ, in the sense that ρ_n^2 is minimized for large σ. Before that, however, we would like to explain the choice of the functional form of the correction term, $(\xi_p^2 + \eta_q^2)$. This is best done by going from the Fourier space to the physical one. The new algorithm now differs from the original one, by an additional explicit term on the right-hand side of (1.3), as follows:

$$(1 - \lambda\delta_x^2)(1 - \lambda\delta_y^2)(v_{j,k}^{(n+1)} - v_{j,k}^{(n)}) = \alpha\lambda(\delta_x^2 + \delta_y^2)\, v_{j,k}^{(n)} + \frac{\gamma}{4}\,\lambda^2 \delta_x^2 \delta_y^2 (\delta_x^2 + \delta_y^2)\, v_{j,k}^{(n)}. \tag{2.3}$$

Now, the virtue of the delta form of (1.3) is that upon convergence, i.e., when $v_{j,k}^{(n+1)} \to v_{j,k}^{(n)}$, the vanishing of the right-hand side implies $(\delta_x^2 + \delta_y^2)\, v_{j,k}^{(n)} \to 0$; i.e., the steady-state equation is satisfied. This feature is maintained in the new algorithm (2.3) since the operator $I + \lambda\delta_x^2\delta_y^2$ is invertible and its inverse has a norm less than 1.

To estimate γ, we start by asking what value of γ will maximize the rate of reduction of the residual in one time-step or, equivalently, what γ will minimize ρ_1^2 for large σ.

$$\rho_1^2 \xrightarrow[\sigma\geqslant 1]{} \frac{4}{N^2}\sum_{\substack{p=1\\ \text{odd}}}^{N-1}\sum_{\substack{q=1\\ \text{odd}}}^{N-1}\left\{1-\gamma\left[\sin^2\left(\frac{\pi p}{2N}\right)+\sin^2\left(\frac{\pi q}{2N}\right)\right]\right\}^2$$

$$=\frac{1}{m^2}\sum_{k=1}^{m-1}\sum_{j=1}^{m-1}[1-\gamma(\sin^2\theta_k+\sin^2\theta_j)]^2 \tag{2.4}$$

where

$$\theta_k=\frac{(2k-1)\,\pi}{4m},\qquad \theta_j=\frac{(2j-1)\,\pi}{4m},\qquad m=\frac{N}{2}.$$

Equation (2.4) may be rewritten as

$$\rho_1^2 \xrightarrow[\sigma\geqslant 1]{} 1-4\gamma_1 S_1+2\gamma^2(S_1^2+S_2), \tag{2.5}$$

where

$$S_1=\frac{1}{m}\sum_{k=1}^{m}\sin^2\theta_k=\frac{1}{m}\sum_{k=1}^{m}\frac{1}{2}(1-\cos 2\theta_k)=\frac{1}{2}-\frac{1}{2\mathrm{m}}\sum_{k=1}^{m}\cos(2k-1)\frac{\pi}{2m}=\frac{1}{2}.$$

$$S_2=\frac{1}{m}\sum_{k=1}^{m}\sin^4\theta_k=\frac{1}{m}\sum_{k=1}^{m}\frac{1}{8}(3-4\cos 2\theta_k+\cos 4\theta_k)=\frac{3}{8}.$$

Thus, for σ large

$$\rho_1^2\sim 1-2\gamma+\tfrac{5}{4}\gamma^2$$

and ρ_1^2 is minimized at $\gamma=0.8$. Note that this result is independent of the mesh size and the value α. One would then expect the efficacy of the correction to be insenstive to changes in mesh size, to grid stretchings, etc. Numerical evidence confirms this property. This value of $\gamma_{\text{opt}}=0.8$ was obtained on the basis of analyzing the improved behaviour of ρ_1^2. Numerical experimentarion with ρ_n^2 (2.2), as well as with the finite-difference algorithm (2.3) showed γ_{opt} for a wide range of N and n to be $0.8\leqslant\gamma_{\text{opt}}\leqslant 0.9$ with very little sensitivity in this range. We thus recommend the use of the predicted value $\gamma_{\text{opt}}=0.8$. Figure 3 shows the variation in time-steps versus λ for various mesh sizes N, both for the uncorrected AF-solver and for the modified algorithm.

Numerical experimentation has shown that further improvement is obtained by running the modified algorithm with two alternating time-steps: one corresponds to the desired λ, and the other is for $\lambda=O(1)$. Details are presented in Section 3.

3

Numerical Results

The purpose of the numerical calculations to be presented in this section is threefold:

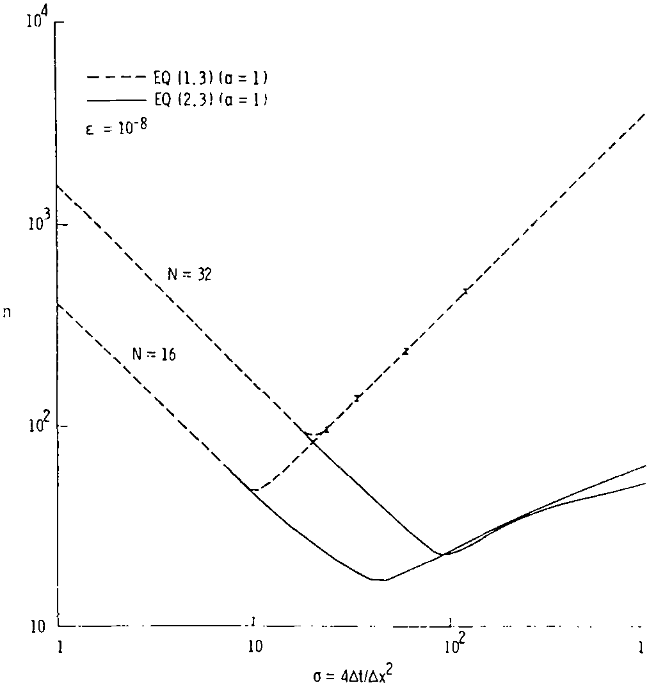

FIG. 3. Comparison of performance of standard backward Euler ADI scheme (Eq. (1.3), $\alpha = 1$) with new algorithm (Eq. (2.3), $\alpha = 1$).

(a) To test the validity of the analysis by comparing the number of iterations (time-steps) to converge to steady state n as a function of σ as predicted by the theory—to the actual results obtained by the finite-difference scheme (1.3).

(b) To test the quality of the improvement predicted for the new corrected algorithm (2.3).

(c) To test how this new algorithm will perform when applied to situations not covered by the analysis such as stretched grids and mixed Dirichlet–Neumann boundary conditions.

Figure 2 shows a log–log graph of n, the number of iterations required to reduce the L_2^2-norm of the initial residual by a factor of 10^8, versus $\sigma = 4\Delta t/\Delta x^2$. This is done for various mesh sizes, $N = 8, 16, 32, 64, 128$. The full lines are the theoretical predictions given by (1.34). The symbols are the results of individual runs using the AF finite-difference scheme (1.3) run to the same tolerance of 10^{-8} of the ratio of the residual norms. The initial condition used was a regular pyramid whose Fourier coefficients are given by (1.21); the boundary conditions were homogeneous. Note that this initial condition is different from the one implied by (1.34). We did this to demonstrate the insensitivity of the prediction to different initial conditions as long as they all introduce many frequencies into the problem. In fact, when the scheme (1.3) is run with initial conditions corresponding to (1.22), rather than (1.21), the numerical results lie on top of the predicted curves. Further experimentation with the initial conditions (such as truncating the pyramid by 30%) and with the boundary conditions (such as replacing the zero boundary values with linear segments with discontinuous slopes) gave essentially the same answers as shown on the graph. The agreement shown in Fig. 2 between the predicted and actual finite-difference runs seems to substantiate the hypothesis underlying the present analysis; namely, that if the initial conditions and/or the boundary condition introduce many frequencies into the problem, then the representation of the residual attenuation, given by Eq. (1.34) is appropriate. All of the above was done for the $\alpha = 1$ case (Douglas–Gunn); for $\alpha = 2$ similar results were obtained.

In Fig. 3, we compare the performance of the standard AF-scheme, (1.3), with that of the corrected algorithm (2.3) with $\gamma = 0.8$. This is done for two meshes, 16×16 and 32×32. The dashed lines are the uncorrected results, and the solid lines are computed with the new algorithm.

First, note the improved performance as σ is increased. Another interesting feature of the results for the corrected scheme is the shallowness of the n versus σ curve for $\sigma > \sigma_{\text{opt}}$. Indeed note how for large $\sigma (\sigma \rightarrow 10^3)$ the corrected algorithm

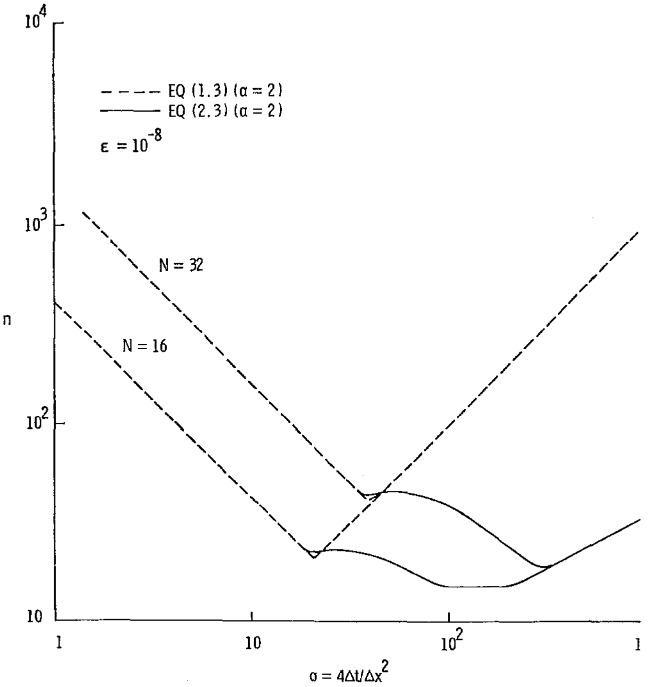

FIG. 4. Comparison of performance of the Peaceman–Rachford scheme, (Eq. (1.3), $\alpha = 2$) with new algorithm, (Eq. (2.3), $\alpha = 2$).

tends to converge to steady state in the same number of steps independently of the mesh size.

So far we have shown the various comparisons for the Douglas–Gunn algorithm ($\alpha = 1$). In Fig. 4, we repeat the numerical experiment of Fig. 3 for the $\alpha = 2$ case (Peaceman–Rachford). It is seen that the corrected algorithm (2.3) with $\alpha = 2$ gives the same kind of improvement as for the $\alpha = 1$ case as predicted by the theory for large σ.

Next, the corrected algorithm (2.3), with $\alpha = 1$, was applied to problem with mixed Neumann Dirichlet data. A Neumann condition $u_x = 0$ was imposed on one edge. Dirichlet data of linear segments were given on the opposite side. On the other two edges were imposed parabolas, which joined the Neumann edge smoothly. The results are shown in Fig. 5. Note that both the uncorrected and corrected schemes perform better than under the fully Dirichlet boundary conditions. Also recall that the theory is based on vanishing deviations on all boundaries and is therefore presumably applicable only to the Dirichlet problem.

It is well known that the most efficient way of reaching steady state, on a uniform mesh via the standard AF-schemes, is to use a cyclical sequence of time-steps [10]. It might be interesting to check this point for the new corrected algorithm (2.3). Since, as we have seen, the corrected algorithm displays a convergence behaviour

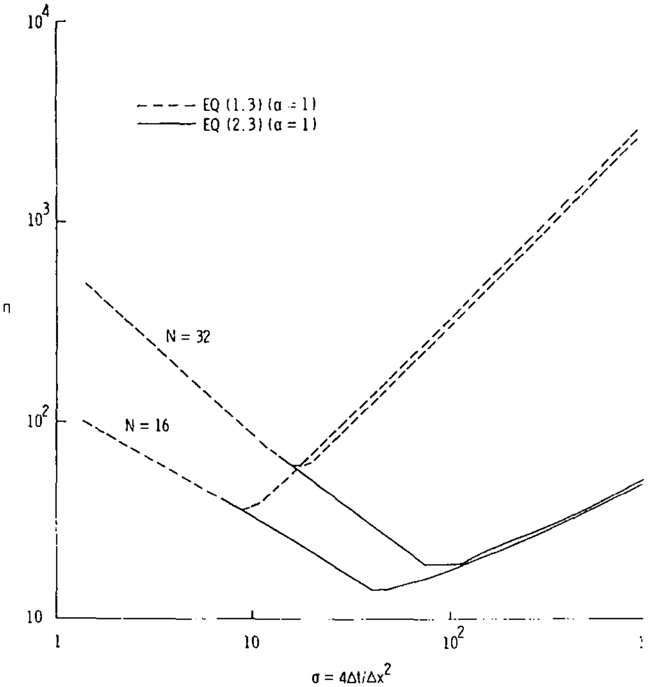

FIG. 5. Effect of mixed Dirichlet–Newmann boundary conditions on performance of standard and new algorithms.

which is insensitive to σ (for large σ), we tried a sequence of only 2 alternating time-steps, one corresponding to small $\sigma(\lambda=1)$ and one to large σ. The results for grids of 16×16 and 32×32 are shown in Table I for the solution of the $\alpha=2$ case. Column 1 shows the number of iterations n necessary to converge to steady state with residual tolerance of 10^{-8}, for the case of the standard Peaceman–Rachford ADI solver with the constant time-step taken at the appropriate optimal σ. Column 2 indicates the improvement when the new algorithm is applied at its optimal σ. (Note, however, from Fig. 4 that the performance of the scheme is insensitive to σ above σ_{opt}.) This improved timing still does not match the cyclical time-step results applied to the standard scheme as given in Column 4. We then tried the alternating time-step sequence mentioned above, see Column 3. While an additional improvement was achieved, it still does not quite match the cyclical time-step method of solution when applied to this problem of uniform mesh.

Finally, we applied the various algorithms and methods of solution to the case of a stretched grid ($N=17$). The problem has Dirichlet data on the square with boundary conditions such that they vanish at the corners and rise linearly to the midpoints of the sides of the square. The discontinuities of the slopes assure the presence of many frequencies in the problem. The grid was stretched through the transformations,

$$\begin{aligned} x_j &= \pi(\tfrac{1}{2}z_1)[z_1+v-(z_1^2+v^2)^{1/2}], \qquad 0 \leqslant j \leqslant N, \\ y_k &= \pi(\tfrac{1}{2}z_2)[z_2+v-(z_2^2+v^2)^{1/2}], \qquad 0 \leqslant k \leqslant N, \end{aligned} \tag{3.1}$$

where

$$\begin{aligned} z_1 &= \tanh^{-1} X_j, \\ z_2 &= \tanh^{-1} Y_k. \end{aligned} \tag{3.2}$$

Note that $1 \geqslant X,\ Y \geqslant -1$ for $0 \leqslant x,\ y \leqslant \pi$. One may now view the problem as either being solved in the physical plane with the finite differencing done on a stretched grid, or equivalently, as being solved in the X–Y plane, with the governing equation having variable coefficients and the differencing done on a uniform mesh. We have chosen to use the latter approach in coding the problem. The transformation (3.1) clusters the points exponentially near the lines $x=\pi/2$, $y=\pi/2$, and leads to a mesh

TABLE I

N	Eq. (1.3) ($\alpha=2$)	Eq. (2.3) ($\alpha=2; \gamma=0.8$)	Eq. (2.3) $\alpha=2, \gamma=.8$ Alternating time step sequence	Eq. (1.3) ($\alpha=2$) Cyclical time step sequence
17	22	15	10	7
33	42	19	14	9

TABLE II

ν	Max cell aspect ratio	Eq. (1.3) ($\alpha = 2$)	Eq. (2.3) ($\alpha = 2$)	Eq. (1.3) ($\alpha = 2$) Cyclical time step sequence	Eq. (2.3) ($\alpha = 2$) Cyclical time step sequence	Eq. (1.3) ($\alpha = 2$) λ_{ij} = constant
0.5	2.45	40	16	16	16	25
1	7.91	73	23	33	16	25
1.5	14.72	104	31	74	16	34
2	22.09	133	38	133	21	45
2.5	29.70	160	44	209	24	54
3	37.49	185	49	301	28	62
3.5	45.32	208	55	409	29	70
4	53.20	230	58	533	32	78
4.5	61.08	251	62	677	32	85
5	69.00	271	65	834	34	90
10	148.50	414	77	337	55	141

that has some very high "aspect-ratio" ($\Delta x_j/\Delta y_k$) cells; for example, with $\nu = 10$, $\Delta x_N/\Delta y_N \simeq 148$. Table II summarizes the results for the $\alpha = 2$ case. Column 1 gives the iteration count for the standard ADI Peaceman-Rachford scheme for the optimum Δt which was found experimentally. In Column 2, the tabulation is for corrected algorithm with $\gamma = 0.8$ and an optimal time-step, also found experimentally, although the results were much less sensitive to changes in Δt in this case. Column 3 gives the results for a cyclic time sequence chosen on the basis of a uniform grid applied to the standard $\alpha = 2$ scheme. As can be seen, the results deteriorate rapidly with the stretch ratio. The same time-step sequence, although not optimal for the nonuniform mesh when applied to the new corrected algorithm (2.3) with $\alpha = 2$, leads to substantially improved results; see Column 4. Since some investigators choose to converge to steady state not by using a constant Δt over the field, but by employing a constant $\lambda = \Delta t_{jk}/\Delta x_j \Delta y_k$, we tried this approach coupled with the cyclic time-step sequence; see Column 5. First, it should be observed that the results quoted in this column are not as efficacious as those given in Column 4, although they do represent an improvement over those given in Column 3. Moreover, this approach of using constant λ, was found experimentally not to be unconditionally stable and we were not able to get any results for $\nu > 20$. This is true for both the standard and the new corected algorithm.

SUMMARY

The following main results were obtained in this paper:

The convergence to steady state of parabolic ADI solvers, such as the Douglas–Gunn or the Peaceman–Rachford algorithms, is analyzed in terms of the

L_2-norm of the residual. This approach, which assumes the presence of many frequencies and averages over their spectrum, turns out to be successful in predicting the dependence of the number of iterations needed to converge to steady state on the Courant number.

A new corrected ADI algorithm has been devised which has the following properties:

(a) Its construction necessitates only the addition of the same explicit term to all existing AF-codes.

(b) It is robust in the sense that it need not be fine-tuned for different mesh sizes, different grid stretchings, mixed Dirichlet–Neumann boundary conditions, etc.

(c) The rate of convergence to steady state is substantially improved and is insensitive to the Courant number for a large range of σ.

(d) The correction term appears to largely remove the effect of grid stretching, allowing straightforward application of uniform grid techniques (e.g., cyclic time-steps) to stretched grid problems.

(e) The method of derivation is easily extended to the three-dimensional case.

REFERENCES

1. S. S. ABARBANEL, D. L. DWOYER, AND D. GOTTLIEB, "Improving the Convergence Rates of Parabolic ADI Methods," ICASE Report, No. 82–28, 1982.
2. S. ABARBANEL, D. L. DWOYER, AND D. GOTTLIEB, "Stable Implicit Finite Diference Methods for Three-Dimensional Hyperbolic Systems" ICASE Report, No. 82-39, 1982.
3. R. W. BEAM AND R. F. WARMING, *J. Comput. Phys.* **22**, (1976), 87.
4. A. BRANDT, *Math. Comput.* **31**, (1977), 333.
5. W. R. BRILEY, AND H. MCDONALD, Solution of the three-dimensional compressible Navier–Stokes equations by an implicit technique, *in* "Fourth Intl. Conf. Numerical Methods Fluid Dynamics," Lecture Notes in Physics, No. 35, p. 105, Springer-Verlag, Berlin/New York, 1974.
6. W. R. BRILEY AND H. MCDONALD, *J. Comput. Phys.* **34**, (1980), 54.
7. J. DOUGLAS AND J. E. GUNN, *Numer. Math.* (1964), 428.
8. D. L. DWOYER AND F. C. THAMES, "Temporal and Spatial Inconsistencies of Time-Split Finite-Difference Schemes," NASA TP-1790, 1981.
9. R. HIRSH, Private communication.
10. A. R. MITCHELL AND D. F. GRIFFITHS, "The Finite-Difference Method in Partial Differential Equations, "Wiley–Interscience, New York, 1980.
11. D. W. PEACEMAN AND H. H. RACHFORD, *J. Soc. Ind. Appl. Math.* **3**, (1955), 28.
12. R. D. RICHTMYER AND K. W. MORTON, "Difference Methods for Initial Value Problems," 2nd ed., Interscience, New York, 1967.
13. W. T. THOMPKINS, JR AND R. H. BUSH, *J. Comput. Phys.* **48** (1982), 302.