JOURNAL OF COMPUTATIONAL PHYSICS 59, 232–246 (1985)

Experiments with Implicit Upwind Methods for the Euler Equations

WIM A. MULDER

University Observatory, Leiden, The Netherlands

AND

BRAM VAN LEER

Department of Mathematics and Informatics, University of Technology, Delft, The Netherlands

Received December 28, 1983; revised August 28, 1984

A number of implicit integration schemes for the one-dimensional Euler equations with conservative upwind spatial differencing are tested on a problem of steady discontinuous flow. Fastest convergence (quadratic for the first-order "backward Euler" scheme) is obtained with the upwind switching provided by van Leer's differentiable split fluxes, which easily linearize in time. With Roe's nondifferentiable split flux-differences the iterations may get trapped in a limit-cycle. This also happens in a second-order scheme with split fluxes, if the matrix coefficients arising in the implicit time-linearization are not properly centered in space. The use of second-order terms computed from split fluxes degrades the accuracy of the solution, especially if these are subjected to a limiter for the sake of monotonicity preservation. Second-order terms computed from the characteristic variables perform best.

1. INTRODUCTION

Implicit time-dependent methods for computing steady inviscid flows have been around for some time [1], but there have been few applications with strong shocks. In particular, the influence of a shock wave on the convergence to a steady state is not well known. The aim of this paper is to study the performance of a number of implicit schemes for time-integration of the one-dimensional Euler equations using conservative upwind-biased spatial differencing for a proper treatment of shocks.

The implicit method, "backward Euler," is formulated in Section 2. It is tested on a one-dimensional model of the isothermal flow in a spiral galaxy, described in Section 3. Computations of this flow with explicit methods were made by van Albada *et al.* [2]. It makes a good test problem, because of the strong shock (Mach 2.5) and the cyclic space-coordinate that prevents disturbances to leave the computational domain.

0021-9991/85 $3.00

The performance of schemes with first-order spatial accuracy is described in Section 4. First, Roe's [3] approximate solution of the Riemann problem for neighboring cells is used to provide the upwind switching, with a modification to exclude expansion shocks. Then we use flux-vector splitting [11], with continuously differentiable flux-vectors according to van Leer [4].

The latter method is also used in the schemes with second-order spatial accuracy tested in Section 5. Such accuracy can be achieved by assuming a set of independent state quantities to have linear distributions in each cell, rather than the uniform distributions of first-order schemes. The slope of a linear distribution can be found through a difference-averaging procedure constrained by the monotonicity condition [5]. The averaging was applied to differences of various sets of quantities: the split flux-vectors, variables related to the characteristic variables, and the correct characteristic variables.

Section 6 contains the main conclusions of this paper.

2. Formulation of the Method

Let w denote the vector of M conserved state quantities, $f(w)$ the corresponding flux-vector and $s(w, x)$ the source-term vector. The space-coordinate is x, time is t. We wish to find a stationary solution for w of the hyperbolic equations

$$\frac{\partial w}{\partial t} = -\frac{\partial f}{\partial x} + s \equiv g, \tag{1}$$

which is equivalent to the problem of determining w such that the residual g vanishes.

To discretize Eq. (1) we define a grid with N zones centered at the points x_i, separated by Δx. Time at iteration step n is denoted by t^n, the time step by Δt^n, so $t^{n+1} = t^n + \Delta t^n$. Differences with respect to time are written as $\Delta_t w = w^{n+1} - w^n$ and with respect to space as $\Delta_{i+1/2} w = w_{i+1} - w_i$. Here w_i represents the zone-averaged value of w. The vector g is discretized by upwind differencing.

Let W be the vector consisting of $N \times M$ elements w_{ki} ($k = 1, 2, \ldots, M$ and $i = 1, 2, \ldots, N$), and let G be defined analogously with elements g_{ki}. A class of implicit schemes is given by

$$L^n \Delta_t W \equiv \left[\frac{1}{\Delta t^n} - \alpha M^n\right] \Delta_t W = G^n, \tag{2}$$

where $M^n = M(W^n)$ is a linear operator providing numerical stability for arbitrary values of the timestep Δt^n. For $\alpha = 1$ and $M = \partial G / \partial W$ (the Jacobian of G with respect to W), we have the backward Euler or implicit Euler scheme. The latter reduces to Newton's method for finding a root if $\Delta t^n \to \infty$, in which case quadratic convergence is obtained. If the Jacobian $\partial G / \partial W$ is costly to form, it may be

replaced by a simpler approximation. As a consequence, the number of iterations required to obtain a converged solution is likely to increase, but the lower cost per iteration may compensate for this.

The right-hand side G^n of Eq. (2) contains all the physics of the problem and defines the accuracy of the solution. The matrix L^n controls the convergence process.

Newton's method will work properly if W is sufficiently close to the stationary solution. Therefore, the scheme is started with a small Δt^n to let the iteration process mimic an explicit, time-accurate integration. Once W starts feeling the attraction of the final solution, Δt^n may become large and, if $\alpha = 1$, Newton's method starts working. The stage of convergence can be monitored by

$$RES^n = \max_{k,i} \left(\frac{|g_{ki}^n|}{|w_{ki}^n| + h_{ki}^n} \right), \qquad k = 1, 2, \ldots, M,\ i = 1, 2, \ldots, N, \tag{3}$$

where h_{ki} is some positive constant that prevents division by zero. The time-step is derived from this quantity according to

$$\Delta t^n = \varepsilon / RES^n. \tag{4}$$

Equation (4) guarantees that in the *explicit* case ($\alpha = 0$) the relative change of the state quantities per time-step will nowhere exceed ε:

$$\frac{|\Delta_t w_{ki}|}{|w_{ki}^n| + h_{ki}^n} \leqslant \varepsilon, \qquad k = 1, 2, \ldots, M \text{ and } i = 1, 2, \ldots, N. \tag{5}$$

3. The Test Problem

The test problem, due to Woodward [12] and used previously for a comparison [2] of explicit methods, is briefly summarized below. It describes the flow of isothermal gas along an almost circular path through the stellar gravitational field of a rotating two-armed spiral galaxy, as sketched in Fig. 1. The vector of conserved quantities in this case is

$$w = \begin{pmatrix} \rho \\ \rho u \\ \rho v \end{pmatrix}, \tag{6a}$$

where ρ is the density, u the velocity perpendicular to and v the velocity along the spiral arms. The space-coordinate x measures distance perpendicular to the arms; the vector of fluxes in this direction is

$$f = \begin{pmatrix} \rho u \\ \rho(u^2 + c^2) \\ \rho u v \end{pmatrix}, \tag{6b}$$

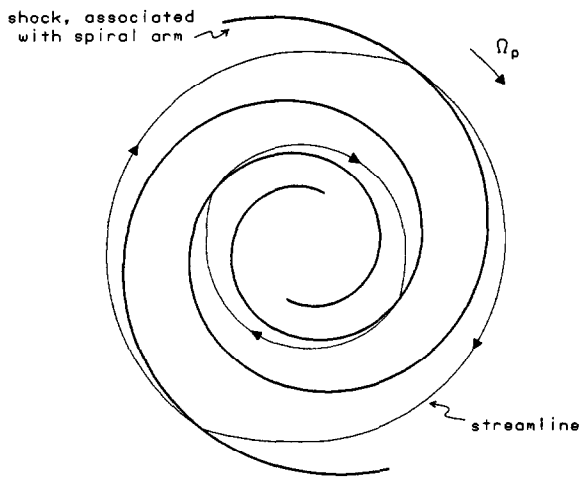

FIG. 1. Two typical streamlines in the gravitational field of a rotating spiral galaxy.

with c the effective sound speed, a constant. The source-term vector is

$$s = \begin{pmatrix} 0 \\ 2\Omega(v - v_0)\,\rho + \dfrac{2}{\alpha r}\,\rho A \sin\eta \\ -\dfrac{\kappa^2}{2\Omega}\,(u - u_0)\,\rho \end{pmatrix}. \tag{6c}$$

Here $\Omega = \Omega(r)$ is the angular velocity at a radius r with respect to the center of the galaxy. The spiral arms are assumed to be tightly wound, with small pitch angle θ_s. The unperturbed velocities are approximately $v_0 = r(\Omega - \Omega_p)$ and $u_o = \alpha r(\Omega - \Omega_p)$, where Ω_p is the pattern speed of the rapidly rotating spiral pattern and $\alpha = \sin\theta_s \ll 1$. The epicycle frequency κ is given by $\kappa^2 = (2\Omega/r)(d/dr)(r^2\Omega)$, A is the amplitude of the spiral perturbation on the gravitational field, and η is the spiral phase defined by $\eta = 2x/\alpha r$. The source term with A is due to the gravitational field, the other source terms are due to the Coriolis force. The problem is periodic: $w(\eta + 2\pi, t) = w(\eta, t)$. A procedure to find the stationary solution by integrating the stationary differential equations is given by Roberts [6]. For sufficiently large A this solution is transonic and discontinuous.

The parameters are chosen as is thought to be appropriate for the solar neighborhood: $\Omega = 25$ km sec^{-1} kpc^{-1}, $\kappa = 31.3$ km sec^{-1} kpc^{-1}, $\Omega_p = 13.5$ km sec^{-1} kpc^{-1}, $c = 8.56$ km sec^{-1}, $r = 10$ kpc, $\alpha = \sin(6°.7)$, and $A = 72.92$ (km sec^{-1})^2.

All calculations started from uniform initial values:

$$\rho_i^0 = 1, \qquad u_i^0 = u_0, \qquad v_i^0 = v_0, \qquad i = 1, 2, \ldots, N. \tag{7a}$$

The constants h_{ki} in Eq. (3) are chosen to be

$$h_{1i}=0, \qquad h_{2i}=\rho_i c, \qquad h_{3i}=\rho_i c, \qquad i=1, 2,..., N. \tag{7b}$$

Finally, we mention that the source terms are linear in w:

$$s=Bw, \tag{8a}$$

$$B(x)=\begin{pmatrix} 0 & 0 & 0 \\ -2\Omega v_0+\dfrac{2A}{\alpha r}\sin\eta & 0 & 2\Omega \\ \dfrac{\kappa^2 u_0}{2\Omega} & \dfrac{-\kappa^2}{2\Omega} & 0 \end{pmatrix}. \tag{8b}$$

4. First-Order Accurate Results

Up to this point we have discretized Eq. (1) in time, yielding Eq. (2), and we have specified the source term. We now have to discretize in space in order to evaluate $\partial f/\partial x$. This is done by upwind differencing in two different ways: (i) using Roe's [3] linearized Riemann solution and (ii) using van Leer's [4] flux-vector splitting. For a comprehensive review of upwind-differencing techniques the reader is referred to Harten, Lax, and van Leer [7].

4.1. *Roe's Approximate Riemann Solver*

An upwind method of discretizing the flux derivative $\partial f/\partial x$ has been described by Roe [3, 8]. We included Roe's own modification to suppress expansion shocks, reported elsewhere [9]. The right-hand side of Eq. (2) becomes

$$g_i^n(w_{i-1}, w_i, w_{i+1})=B_i w_i-\frac{1}{\Delta x}(A^+_{i-1/2}\Delta_{i-1/2}\, w+A^-_{i+1/2}\Delta_{i+1/2} w), \qquad i=1, 2,..., N, \tag{9}$$

where $A^+_{i-1/2}(w_{i-1}, w_i)$ and $A^-_{i+1/2}(w_i, w_{i+1})$ are the positive and negative parts of the discrete approximations to $A(w)\equiv\partial f/\partial w$. Roe's approximate Riemann solver includes a simple recipe to calculate these matrices but their derivatives, needed on the left-hand side of (2), are costly to compute and do not even exist in shocks and around sonic points. We therefore freeze $A^{\pm}$ at the old time-level when deriving $\partial g/\partial w$. Eq. (2) becomes:

$$\left[-\frac{\alpha}{\Delta x} A^{+}_{i-1/2}\right] \Delta_t w_{i-1}$$
$$+\left[\frac{1}{\Delta t^n} - \alpha B_i + \frac{\alpha}{\Delta x}\left(A^{+}_{i-1/2} - A^{-}_{i+1/2}\right)\right] \Delta_t w_i$$
$$+\left[\frac{\alpha}{\Delta x} A^{-}_{i+1/2}\right] \Delta_t w_{i+1} = g^n_i, \qquad i = 1, 2, \ldots, N, \tag{10}$$

giving the matrix L^n a cyclic block-tridiagonal structure. At this point we remark that if the matrix M^n_i is given by

$$M^n_i = B_i - \frac{1}{\Delta x}(A^{+}_{i-1/2}\Delta_{i-1/2} + A^{-}_{i+1/2}\Delta_{i+1/2}), \tag{11}$$

then Eq. (10) can be rewritten as

$$\left[\frac{1}{\Delta t^n} - \alpha M^n_i\right] \Delta_t w = M^n_i w^n_i, \qquad i = 1, 2, \ldots, N. \tag{12}$$

The convergence behavior of this scheme is given in Fig. 2a, for $\alpha = 1$, $N = 64$, and $\varepsilon = 0.5$. Plotted is $^{10}\log(RES^n)$ as defined in Eqs. (3) and (7b) versus the number of iteration steps n. After some almost explicit searching the convergence process sets in, but it suddenly stops. Closer inspection reveals that the solution jumps symmetrically up and down across the correct solution: the convergence process is locked in a two-vector limit-cycle. An analysis of Eq. (12) makes this comprehensible. For $\alpha = 1$ and large Δt^n we have

$$-M^n \Delta_t w = M^n w^n, \tag{13}$$

implying that $M^n w^{n+1} = 0$, but not necessarily $M^{n+1} w^{n+1} = 0$ which is actually desired. For the two solutions in the limit-cycle we find the following, almost exact relation

$$M^{n+1} w^{n+1} = -M^n w^n. \tag{14}$$

Neglecting terms of the order $O((\Delta_t w)^2)$ we can write the left-hand side as

$$M^{n+1} w^{n+1} = M^{n+1} w^n + M^{n+1} \Delta_t w = M^{n+1} w^n + M^n \Delta_t w$$
$$= M^{n+1} w^n - M^n w^n = (\Delta_t M)\, w^n. \tag{15}$$

Thus we are confronted with the fact that the time derivative of $A^{\pm}$, hence of M, was neglected earlier.

It is worth mentioning that a run with fixed boundary values, taken from the exact solution, converged properly (Fig. 2b). Clearly the boundaries provide enough damping to avoid the limit-cycle.

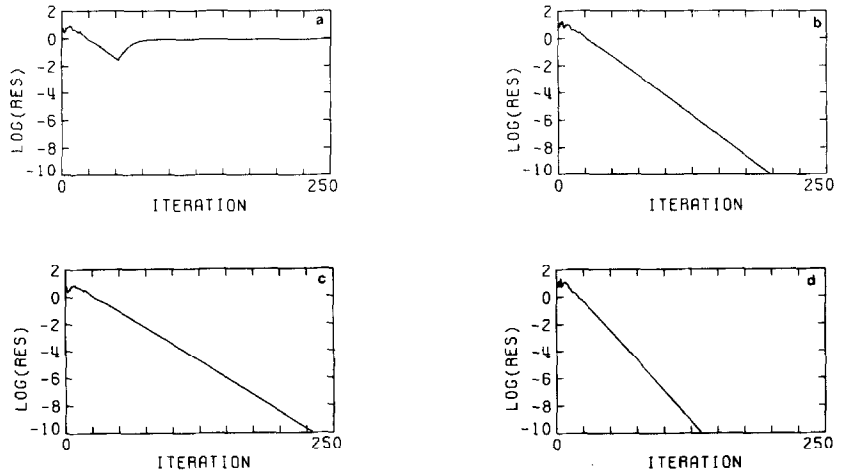

FIG. 2. (a) Convergence history for scheme (10), based on Roe's approximate Riemann solver. The iteration process ends up in a two-vector limit-cycle. Parameters: $\alpha = 1$, $\varepsilon = 0.5$. (b) As in (a), but with fixed boundary values. (c) As in (a), but with $\alpha = 1.2$. (d) Convergence history for the β-scheme as given in Eq. (15). Parameters: $\alpha = 1$, $\varepsilon = 0.5$, and $\beta = 0.5$.

Being aware of Eq. (14), we experimented with two different amendments. The most obvious remedy is to make $\alpha > 1$, implying under-relaxation. It turned out that for $\alpha \gtrsim 1.2$ the scheme converged: the convergence slowed down with increasing α. The behavior for $\alpha = 1.2$ is shown in Fig. 2c.

Another way out is to repair the inconsistency of Eq. (13). A consistent formulation would be $-M^{n+1}\Delta_t w = M^{n+1}w^n$, ensuring that $M^{n+1}w^{n+1} = 0$. On the left-hand side we may replace M^{n+1} by M^n without reproach. To estimate M^{n+1} for the right-hand side we devised a predictor-corrector method that will be referred to as the "β-scheme":

Step 1

$$\left[\frac{1}{\Delta t^n} - \alpha M^n\right] \tilde{\Delta_t} w = M^n w^n, \tag{16a}$$

from which follow $\tilde{w}^{n+1} = w^n + \tilde{\Delta_t} w$ and $\tilde{M}^{n+1}$;

Step 2

$$\left[\frac{1}{\Delta t^n} - \alpha M^n\right] \Delta_t w = [(1-\beta) M^n + \beta \tilde{M}^{n+1}] w^n, \tag{16b}$$

after which follows the final update $w^{n+1} = w^n + \Delta_t w$. For $\beta = 0$ we have the old version of Eq. (12). The obvious choice $\beta = 1$, does not lead to convergence. Eq. (14) suggests that $\beta = \frac{1}{2}$ at least will destroy the limit-cycle. It actually does, as can be seen in Fig. 2d, with $\alpha = 1$.

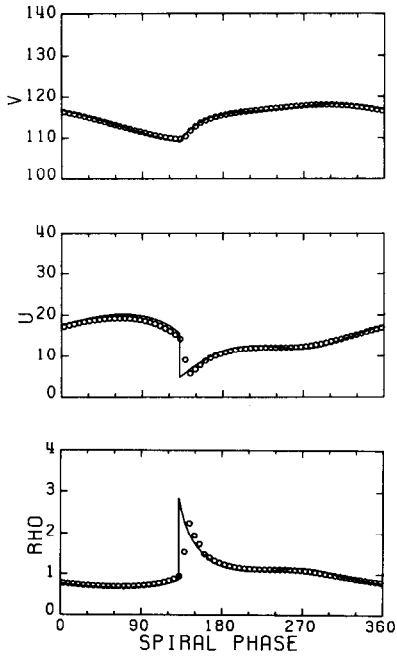

FIG. 3. The converged solution using Roe's approximate Riemann solver.

Note that the inversion of the cyclic block-tridiagonal matrix in Step 2 is identical to the one in Step 1. Since this inversion is the most elaborate part of the scheme, the second step adds a relatively small amount of extra computer time. All together the β-scheme is almost twice as efficient as the original scheme with $\alpha = 1.2$.

Attempts to optimize the β-scheme by choosing a linear combination of $\Delta_t w$ ($\beta = 0$) and $\Delta_t w (\beta = \frac{1}{2})$ resulted in a somewhat faster convergence, but no significant decrease in computer time resulted because of the extra work involved.

We end this section with Fig. 3, showing the converged solution of Roe's scheme.

4.2. *Flux-Vector Splitting*

The previous results motivate a version of scheme (2) where M^n is the *exact* linearization of the right-hand side G^n. This requires a Riemann solver that allows continuous differentiation. The simplest one is given by van Leer [4] on the basis of flux-vector splitting, a technique due to Steger and Warming [11]. Another possibility is Osher's [10] solver, which leads to more complicated formulas and will not be considered here.

In flux-vector splitting, the flux f is considered the sum of a forward flux f^+ and a backward flux f^-. Here we use

$$f^+ = \begin{cases} f & , \quad u \geqslant c, \\ \begin{bmatrix} \frac{\rho}{4c}(u+c)^2 \\ \frac{\rho}{2}(u+c)^2 \\ \frac{\rho}{4c}(u+c)^2 v \end{bmatrix} & , \quad |u| < c, \\ 0 & , \quad u \leqslant -c, \end{cases} \tag{17a}$$

$$f^- = f - f^+. \tag{17b}$$

We now can readily derive the Jacobian matrices $E^{\pm} = \partial f^{\pm}/\partial w$, needed in the implicit upwind differencing scheme. Equation (2) becomes

$$\left[-\frac{\alpha}{\Delta x} E_{i-1}^+\right] \Delta_t w_{i-1} + \left[\frac{1}{\Delta t} - \alpha B_i + \frac{\alpha}{\Delta x}(E_i^+ - E_i^-)\right] \Delta_t w_i + \left[\frac{\alpha}{\Delta x} E_{i+1}^-\right] \Delta_t w_{i+1}$$
$$= B w_i - \frac{1}{\Delta x}(f_i^+ - f_{i-1}^+ + f_{i+1}^- - f_i^-), \qquad i = 1, 2, \ldots, N. \tag{18}$$

The matrix L again has a cyclic block-tridiagonal structure.

The improvement in convergence speed with regard to the previously described schemes is dramatic. Figure 4a shows the convergence history: machine zero is reached in 16 steps with quadratic convergence. The solution is displayed in Fig. 4b. Comparison with Fig. 3 shows that Roe's modified scheme is superior in describing the flow near the sonic point.

We end this section with the conclusion that quadratic convergence can be achieved for discontinuous flows using an implicit first-order upwind-difference scheme with a continuously differentiable numerical flux function. We shall use the same numerical flux function in the second-order schemes of the next section.

5. Second-Order Spatial Accuracy

We can achieve second-order accuracy by assuming a set of (not necessarily conserved) quantities q to have a piecewise linear distribution in each computational cell:

$$q^n(\xi) = q_i^n + \xi \overline{\Delta_i q^n}, \quad \xi \equiv \frac{x - x_i}{\Delta x}, \quad |\xi| < \frac{1}{2}. \tag{19}$$

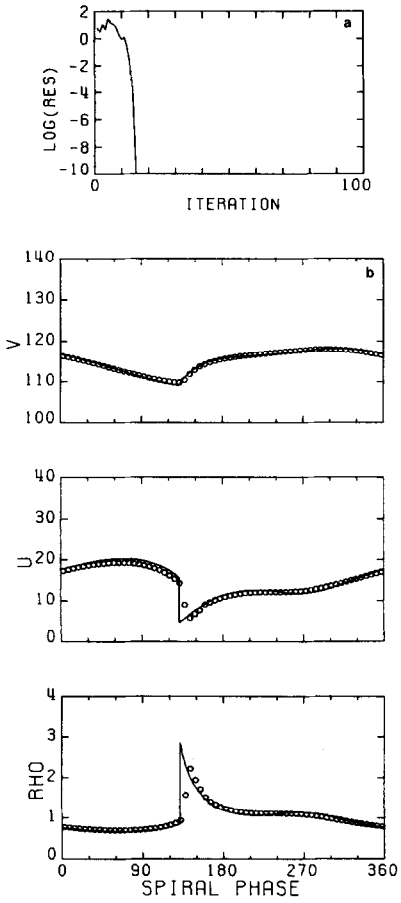

FIG. 4. (a) Convergence history of the first-order accurate solution for scheme (18), using flux-vector splitting. Parameters: $\alpha = 1$, $\varepsilon = 0.5$. (b) Converged solution corresponding with (a).

The undivided gradient $\overline{\Delta_i q^n}$, or simply $\bar{\Delta}$, is found by averaging the differences $\Delta_{i+1/2} q^n$ and $\Delta_{i-1/2} q^n$, or Δ_+ and Δ_-. In this way information of 3 zones is used, leading to the desired increase in accuracy. In regions where q varies smoothly an obvious way of averaging is $\bar{\Delta} = \delta \equiv \frac{1}{2}(\Delta_+ + \Delta_-)$. However, this is not justified at the head or foot of a discontinuity; some modification is recommended then to limit $\bar{\Delta}$ to $O(\Delta x)$. A clear account on this subject is given by van Leer [5] on the basis of the monotonicity condition. A smooth approximation of the switches given in [5], Eq. (66) has been proposed by van Albada [2]:

$$\bar{\Delta} = \frac{2\Delta_+ \Delta_- + 2\varepsilon_a^2}{\Delta_+^2 + \Delta_-^2 + 2\varepsilon_a^2}\,\delta, \tag{20}$$

where ε_a is a small bias of order $O(\Delta x)$, preventing clipping of smooth extrema.

The averaging-limiting procedure is valid for a given single convection equation, implying that we should choose the quantities q to be the characteristic variables of our equations. These variables are obtained by the diagonalization of Eq. (1) without the source terms; they are listed in Eq. (24). In the following we will consider averaging on (i) the components of the split flux-vectors $f^{\pm}$ as given by Eq. (17), (ii) variables related to the characteristic variables, as used in [2], and (iii) the true characteristic variables.

We start with (i). To justify the choice of the split flux-vectors we mention that the $f^{\pm}$ themselves are the characteristic variables that follow from the

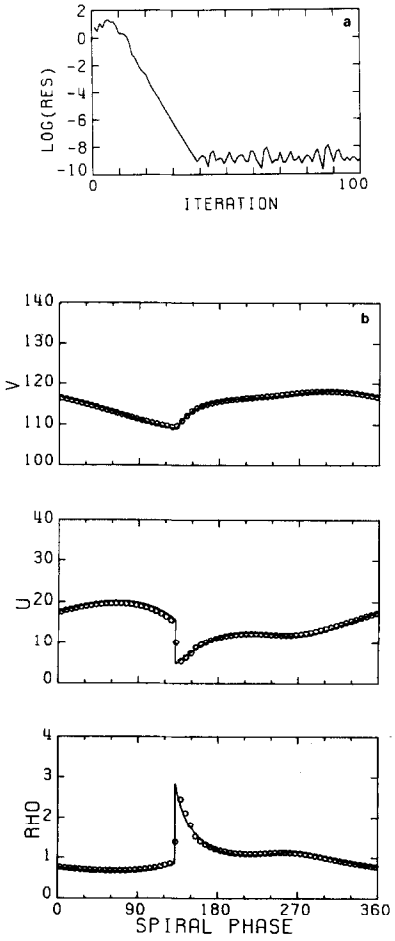

FIG. 5. (a) Convergence history of the second-order accurate solution for the flux-vector splitting scheme (21) with averaging procedure (20) based on the split fluxes ($\alpha = 1$, $\varepsilon = 0.5$). (b) Converged solution corresponding with (a).

diagonalization of $(\partial w/\partial t)+(\partial f^{\pm}/\partial x)=0$, albeit only for $|u|<c$. It is tempting to write down the following second-order version of scheme (18):

$$\left[-\frac{\alpha}{\Delta x}E_{i-1}^{+}\right]\Delta_t w_{i-1}+\left[\frac{1}{\Delta t}-\alpha B_i+\frac{\alpha}{\Delta x}(E_i^{+}-E_i^{-})\right]\Delta_t w_i$$
$$+\left[\frac{\alpha}{\Delta x}E_{i+1}^{-}\right]\Delta_t w_{i+1}=Bw_i-\frac{1}{\Delta x}\{(f_i^{+}+\tfrac{1}{2}\overline{\Delta_i f^{+}})-(f_{i-1}^{+}+\tfrac{1}{2}\overline{\Delta_{i-1}f^{+}})$$
$$+(f_{i+1}^{-}-\tfrac{1}{2}\overline{\Delta_{i+1}f^{-}})-(f_i^{-}-\tfrac{1}{2}\overline{\Delta_i f^{-}})\}. \tag{21}$$

The only change with respect to (18) is the explicit addition of the proper second-order terms on the right-hand side; the left-hand side is the same as in (18) in order to preserve the block-tridiagonal structure. The effect of the second-order terms is that f_i^{+} is evaluated at the right boundary of cell i and f_i^{-} at the left.

The convergence history and solution for scheme (21) with $N=64$ are given in Figs. 5a and b. The solution is not as accurate as we expect from a second-order scheme, neither in the supersonic nor in the subsonic region (cf. Fig. 6b). We can think of two possible explanations for this. First, in the supersonic regions the fluxes $f^{\pm}$ are not the characteristic variables of the equations. Second, in a sonic point the matrices $E^{\pm}=\partial f^{\pm}/\partial w$ are continuous but not continuously differentiable, as required in a second-order method. The effect of this can be seen between in the plot of u in Fig. 5b: the small jump that is typical for flux-vector splitting around the sonic point in a first-order scheme has not disappeared in the second-order scheme.

The bias ε_a^2 was given a fairly large value in the experiment of Fig. 5, implying weak limiting. For smaller values, or stronger limiting, the solution becomes non-unique, namely, different for different initial values. This, again, is an effect of the nonsmoothness of $\partial f^{\pm}/\partial w$ worsened by the limiter. Our opinion on the basis of these results is that second differences of split fluxes, limited or not limited, should be avoided.

Continuing with (ii) we choose

$$q=\begin{pmatrix}\ln\rho\\ u/c\\ v/c\end{pmatrix},\qquad \Delta_{i+1/2}q=\begin{pmatrix}\dfrac{\Delta_{i+1/2}\rho}{\frac{1}{2}(\rho_i+\rho_{i+1})}\\ \dfrac{\Delta_{i+1/2}u}{c}\\ \dfrac{\Delta_{i+1/2}v}{c}\end{pmatrix},\qquad \begin{pmatrix}\overline{\Delta_i\rho}\\ \overline{\Delta_i u}\\ \overline{\Delta_i v}\end{pmatrix}=\begin{pmatrix}\rho_i\overline{\Delta_i q_1}\\ c\overline{\Delta_i q_2}\\ c\overline{\Delta_i q_3}\end{pmatrix}. \tag{22}$$

We first evaluate ρ, u and v at both cell boundaries, then use these values to compute f_i^{-} at the left boundary and f_i^{+} at the right. It is necessary to evaluate the matrices $E_i^{\pm}$ on the basis of the same boundary values for proper convergence. The

use of mid-zone values of $E_i^{\pm}$ as in (21), causes the convergence process to end up in a limit-cycle. The deviation from the steady state is largest on the subsonic side of the sonic point, suggesting that the limit-cycle is caused by the nonsmoothness of $E^{\pm}$ around the sonic point.

Convergence for the scheme with $E^{\pm}$ evaluated at the same position as $f^{\pm}$ was achieved within 90 steps for $N=64$ and $\varepsilon_a^2=0.008$, as in [2]. The resulting solution is almost indistinguishable from solution obtained with the explicit time-accurate scheme of [2]. Comparison with the first-order solution in Fig. 4 shows that the kink across the sonic point has disappeared, a result of the second-order approach.

We have also tried a different averaging procedure, somewhat closer to the switches in Eq. (66) of [5] than Eq. (20):

$$\bar{\Delta} = \frac{4\Delta_+\Delta_- + 2\varepsilon_a^2}{(|\Delta_+| + |\Delta_-|)^2 + 2\varepsilon_a^2}\,\delta. \tag{23}$$

This does not appear to significantly affect the convergence or the accuracy of the scheme.

Finally, we discuss the averaging (iii) applied to differences of the true characteristic variables:

$$q = \begin{bmatrix} \dfrac{u}{c} - \ln \rho \\ \dfrac{u}{c} + \ln \rho \\ \dfrac{v}{c} \end{bmatrix}, \qquad \Delta_{i+1/2} q = \begin{bmatrix} \dfrac{\Delta_{i+1/2} u}{c} - \dfrac{\Delta_{i+1/2}\rho}{\frac{1}{2}(\rho_i + \rho_{i+1})} \\ \dfrac{\Delta_{i+1/2} u}{c} + \dfrac{\Delta_{i+1/2}\rho}{\frac{1}{2}(\rho_i + \rho_{i+1})} \\ \dfrac{\Delta_{i+1/2} v}{c} \end{bmatrix},$$

$$\begin{bmatrix} \overline{\Delta_i \rho} \\ \overline{\Delta_i u} \\ \overline{\Delta_i v} \end{bmatrix} = \begin{bmatrix} \dfrac{\rho_i}{2}(\overline{\Delta_i q_2} - \overline{\Delta_i q_1}) \\ \dfrac{c}{2}(\overline{\Delta_i q_2} + \overline{\Delta_i q_1}) \\ c\overline{\Delta_i q_3} \end{bmatrix} \tag{24}$$

We use $\varepsilon_a^2 = 0.008$ in the averaging procedure (20), as in [2], and evaluate $f_i^{\pm}$ and $E_i^{\pm}$ both from the same boundary values of ρ, u, and v in zone i. Results are shown in Fig. 6. Convergence is significantly faster than for the previous second-order schemes. The solution is a notch better in the regions where the flow is smooth. Apparently, interpolating the characteristic variables yields the most efficient second-order scheme. This scheme is also more efficient than the first-order schemes of Figs. 3 and 4b, which require a mesh-refinement of about a factor $\frac{1}{6}$ to match the accuracy of Fig. 6b (see [2, Table 1]).

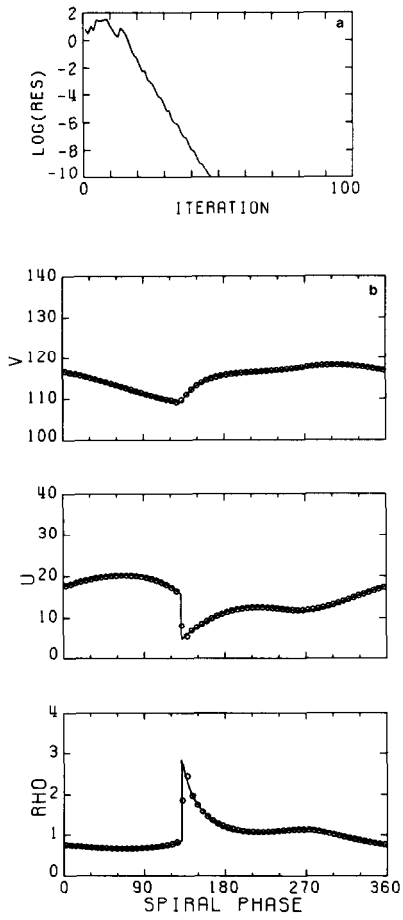

FIG. 6. (a) Convergence history of the solution with averaging procedure (20) based on the characteristic variables (24) ($\alpha = 1$, $\varepsilon = 0.5$, $\varepsilon_a^2 = 0.008$). (b) Second-order accurate solution corresponding with Fig. 6(a).

6. CONCLUSIONS

Our test of implicit upwind-differencing schemes for the Euler equations, including two different Riemann solvers lead to the following conclusions.

(i) Quadratic convergence to a steady state incorporating a shock can be achieved with first-order upwind differencing.

(ii) Incomplete linearization in time of the implicit scheme may cause the iteration process to end up in a limit-cycle. One economic way to avoid this is our β-scheme. Another, more obvious remedy is to use an exact linearization. This

requires a Riemann solver that allows continuous differentiation. Such a solver is provided by [4] (tested) or [10] (not tested).

(iii) If the second-order terms, needed in second-order upwind schemes, are computed directly as differences of split fluxes, the accuracy of the solution degrades considerably in our test problem. The best results are obtained from differences of the characteristic variables.

(iv) The decrease in convergence speed for a second-order upwind scheme with an incomplete linearization in time is amply compensated by the increase in accuracy of the solution. The convergence speed is comparable to that of the non-conservative schemes of Napolitano and Dadone [13].

The extension of the present method to more than one dimension poses the problem of inverting the linear system (2). Several approximative techniques, including multigrid, will be discussed in two forthcoming papers [14, 15].

REFERENCES

1. R. M. BEAM, R. F. WARMING, *J. Comput. Phys.* **22** (1976), 87.
2. G. D. VAN ALBADA, B. VAN LEER, AND W. W. ROBERTS, *Astron. Astrophys.* **108** (1982), 76.
3. P. L. ROE, *J. Comput. Phys.* **43** (1981), 357.
4. B. VAN LEER, *in* "Lecture Notes in Physics," Vol. 170, pp. 507–512, Springer-Verlag, Berlin, 1982.
5. B. VAN LEER, *J. Comput. Phys.* **23** (1977), 276.
6. W. W. ROBERTS, *Astrophys. J.* **158** (1969), 123.
7. A. HARTEN, P. D. LAX, AND B. VAN LEER, *SIAM Rev.* **25** (1983), 35.
8. P. L. ROE, *in* "Lecture Notes in Physics," Vol. 141, pp. 354–359, Springer–Verlag, Berlin, 1980.
9. B. VAN LEER, *SIAM J. Sci. Stat. Comput.* **5** (1984), 1.
10. S. OSHER AND S. CHAKRAVARTHY, *J. Comput. Phys.* **50** (1983), 447.
11. J. L. STEGER AND R. F. WARMING, *J. Comput. Phys.* **40** (1981), 263.
12. P. R. WOODWARD, *Astrophys. J.* **195** (1975), 61.
13. A. DADONE AND M. NAPOLITANO, AIAA Paper No. 82-0972, AIAA-ASMF 3rd Joint Thermophysics, Fluids, Plasma and Heat Transfer Conference, St. Louis, Mo., June 1982.
14. B. VAN LEER AND W. A. MULDER, *in* "Proceedings, INRIA Workshop on Numerical Methods for the Euler Equations for Compressible Fluids, Le Chesnay, France, December 1983," SIAM, Philadelphia, in press.
15. W. A. MULDER, *J. Comput. Phys.*, in press.