JOURNAL OF COMPUTATIONAL PHYSICS 116, 195–211 (1995)

An Implicit–Explicit Eulerian Godunov Scheme for Compressible Flow

J. P. COLLINS*

Information and Mathematical Sciences Branch, Naval Surface Warfare Center, White Oak, Silver Spring, Maryland 20903

P. COLELLA,†

Mechanical Engineering Department, University of California, Berkeley, California 94720

AND

H. M. GLAZ‡

Department of Mathematics, University of Maryland, College Park, Maryland 20742

Received July 5, 1990; revised August 24, 1992

A hybrid implicit–explicit scheme is developed for Eulerian hydrodynamics. The hybridization is a continuous switch and operates on each characteristic field separately. The explicit scheme is a version of the second-order Godunov scheme; the implicit method is only first-order accurate in time but leads to a block tridiagonal matrix inversion for efficiency and is unconditionally stable for the case of linear advection. The methodology is described for the cases of linear advection, for nonlinear scalar problems, and for gas dynamics. An important element of our work is the use of a modified Engquist–Osher flux function in place of the Godunov flux. Several numerical results are presented to demonstrate the properties of the method, especially stable numerical shocks at very high CFL numbers and second-order accurate steady states.

1. INTRODUCTION

In this paper, we develop a hybrid implicit–explicit scheme of a type first discussed in [12] for the case of Lagrangian hydrodynamics. The underlying idea is much older and is discussed, for example, in the book of Richtmyer and Morton [19]; a more thorough and up-to-date literature survey is available in [9]. For these schemes, the difference approximation in time is either implicit or explicit, separately for each family of characteristics and for each cell in the finite difference grid, depending on whether the local CFL number for that family is greater than or less than one. To the extent possible, the hybridization is continuous at CFL number equal to one, and the scheme for the explicit modes is a second-order Godunov method of a type discussed in [6, 7]. Another solution strategy, which is often used for the same class of problems that the hybrid implicit–explicit approach pursued here is used, is based on splitting the equations, i.e., constructing a solution operator which is split into explicit and implicit parts; see [1, 18] for recent work in this area.

This implicit–explicit strategy is intended for problems with spatially and/or temporally localized stiffness in wave speeds. By stiffness, we mean that the high speed modes contain very little energy, yet they determine the explicit time step through the CFL condition. For hydrodynamics, the main example is nearly incompressible flow. Here, the sonic waves will be nearly acoustic and largely decoupled from the particle modes. Traditionally, such problems are handled in one of two ways: at the PDE level by solving instead the incompressible limit of the set of governing equations, or numerically by artificially increasing the temperature of the gas in such a way that the Mach number becomes significant but still low enough that compressibility effects are negligible.

Another example is provided by magnetohydrodynamic

* The support of the Defense Nuclear Agency, MIPR 90-567 is gratefully acknowledged.

† The support of the Applied Mathematical Sciences Program of the U.S. Department of Energy at the Lawrence Livermore National Laboratory under Contract W-7405-Eng-48; of the Army Research Office under Grant DAAL03-88-K-0197; and of a National Science Foundation Presidential Young Investigator Award ACS-8958522, is gratefully acknowledged.

‡ The support of the National Science Foundation under Grant DMS-8703971; of the Applied Mathematical Sciences Program of the U.S. Department of Energy at the Courant Institute of Mathematical Sciences under Grant DE-FG02-88ER25053; and of the Defense Nuclear Agency under Contract DNA001-87-0303 is gratefully acknowledged.

0021-9991/95 $6.00

calculations in certain configurations [4, 20]. Here, if the problem leads to spatially localized regions where a low density plasma is immersed in a strong magnetic field then the Alfvèn wave speed can become very large. Our approach can potentially deal with this problem in a straightforward manner and avoid ad hoc solutions, e.g., modifying the calculation of the electric field at low densities.

A related example is the problem of shock wave–boundary layer interactions. It is often important to simultaneously perform a high resolution calculation of the strong wave interactions in the free stream while also resolving enough of the detail in the boundary layer to accurately model the effect of the no-slip boundary condition on the free stream as well as handling any transport effects between the boundary and the inviscid flow region (which might include particle and temperature transport from the boundary condition). The extension of our approach to the Navier–Stokes equations, also developed here, allows for the smooth transition between a high resolution explicit scheme in the exterior flowfield and a stable method in the boundary layer.

The development of our implicit–explicit scheme has followed the following design principles:

(1) The scheme is in conservation form, both for conservation laws and for their viscous extensions.

(2) The implicit–explicit hybridization is a continuous switch and operates on each characteristic field separately; the method is entirely local in both time and space, depending only on the data in each computational zone.

(3) In the event that all characteristic modes are explicit, the scheme reduces to a version of the second-order Godunov scheme [6, 7].

(4) The implicit advection scheme, which is only first-order accurate in time, satisfies a maximum principle and is unconditionally stable for linear problems.

(5) In the limit of steady state, the scheme is second-order accurate in space.

The first three of these properties are satisfied by the scheme of Fryxell *et al.* [12]. Their implicit scheme is also second-order accurate in time, but it does not satisfy a maximum principle (and, therefore, is not monotone). Additionally, for a system of n equations, their method requires the inversion of a block $2n \times 2n$ tridiagonal system, whereas the current method requires only a block $n \times n$ tridiagonal system. For our intended applications, we do not regard the lack of second-order temporal accuracy to be a serious problem. Waves that we wish to resolve in a flowfield will be treated using the second-order accurate explicit scheme; the others will be rapidly relaxed to equilibrium.

A key component of this work is the introduction of an appropriately smooth numerical flux function for the hyperbolic equations. We use a suitable version of the Engquist–Osher flux [11] as modified for systems by Bell *et al.* [3]. The version used is sufficiently smooth so that the Newton's method linearization is well-behaved; in particular, it converges to steady states even in the presence of strong shocks.

We discuss and implement our method here in the special case of one-dimensional inviscid and viscous compressible flow; a polytropic equation-of-state is assumed for convenience. However, the ideas and implementation methodology are easily extended to more general systems in one space dimension. In particular, it is expected that our results can be extended to gas dynamics with a general equation-of-state using the ideas in [6]. A more difficult question is the extension of these ideas to two or three space dimensions. One starting point is the unsplit second-order Godunov scheme introduced by Colella [5]; results for low Mach number flow may be found in [9, 10].

The first part of Section 2 is devoted to developing our ideas for the case of linear advection; extensions of this linear scalar scheme to handle source terms and viscous extensions follow. In particular, we are able to show that this scheme satisfies (1)–(5) above, thereby fixing our ideas for extension to nonlinear systems. Also in Section 2, we introduce a nonlinear stability constraint for our method in the special case of a scalar conservation law. Section 3 is devoted to describing the scheme for systems of conservation laws and their viscous extensions; the formulae necessary for gas dynamics and compressible flow are included. The development in Section 3 assumes the existence of a sufficiently smooth numerical flux function; its construction is taken up in Section 4, where we also introduce approximations with suitable behavior and which are efficient for practical large-scale applications. Numerical results are presented in Section 5. Some further details and extra computations have appeared in the report [8].

2. AN IMPLICIT–EXPLICIT ALGORITHM FOR ADVECTION

We begin with the linear scalar advection equation for $u = u(x, t)$,

$$\frac{\partial u}{\partial t} + a_0 \frac{\partial u}{\partial x} = 0, \tag{2.1}$$

where $a_0 > 0$ is constant. If $\{u_j^n\}$ represents the zone averages of $u(x, t^n)$ at time level n on a uniform mesh of width Δx, then we calculate $\{u_j^{n+1}\}$ using a conservative difference scheme in the predictor–corrector form,

$$u_j^{n+1} = u_j^n - \sigma(u_{j+1/2} - u_{j-1/2}), \tag{2.2}$$

where $\sigma = a_0 \Delta t/\Delta x$ and $\Delta t = \Delta t^n = t^{n+1} - t^n$ is the time step. The edge states are given by

$$u_{j+1/2} = \begin{cases} u_j^n + \frac{1}{2}(1-\sigma)\,\Delta u_j^n, & \sigma < 1 \\ u_j^n + \left(1 - \frac{1}{\sigma}\right)(u_j^{n+1} - u_j^n), & \sigma \geq 1. \end{cases} \tag{2.3}$$

Here, Δu_j^n is some finite difference approximation to $\partial u/\partial x|_{(x_j, t^n)}\,\Delta x$, to which a monotonicity constraint has been applied [7].

The scheme (2.2)–(2.3) is in conservation form and the hybridization from explicit to implicit form is continuous at $\sigma = 1$. For $\sigma < 1$ the scheme is a second-order upstream weighted scheme of a type considered in [6]. For regions in which $\sigma > 1$ the scheme may be rewritten in the form

$$u_j^{n+1} = \frac{1}{\sigma} u_{j-1}^n + \left(1 - \frac{1}{\sigma}\right) u_{j-1}^{n+1}. \tag{2.4}$$

Using (2.4), an easy induction argument on j shows that the implicit scheme is max norm stable as well as monotonicity-preserving; this implies that all Fourier modes are damped as well. On the other hand, it is only first-order accurate in time.

We now consider extensions of the scheme necessary to handle source terms and viscous effects. The linear scalar equation for $u = u(x, t)$

$$\frac{\partial u}{\partial t} + a_0 \frac{\partial u}{\partial x} = g(x, u) + \mu \frac{\partial^2 u}{\partial x^2}, \tag{2.5}$$

where $a_0, \mu > 0$ are constants is treated. In our applications here, the source term $g(x, u)$ will represent geometric effects, although more general situations are possible. The resulting scheme, analogous to (2.2)-(2.3), is

$$\begin{aligned} u_j^{n+1} = {} & u_j^n + \sigma(u_{j-1/2} - u_{j+1/2}) \\ & + \Delta t\, L((1-\psi)\, u^n + \psi u^{n+1})_j + \Delta t\, g_j^{n+1/2} \end{aligned} \tag{2.6}$$

where σ, Δt are as above, $\frac{1}{2} \leqslant \psi \leqslant 1$, $(Lu)_j = (\mu/(\Delta x)^2)(u_{j-1} - 2u_j + u_{j+1})$, $g_j^{n+1/2}$ is a suitably stable and consistent approximation of the source term, e.g., $g_j^{n+1/2} = \frac{1}{2}(g(x_j, u_j^n) + g(x_j, u_j^{n+1}))$ and

$$u_{j+1/2} = \begin{cases} u_j^n + \frac{1}{2}(1-\sigma)\,\Delta u_j^n + \frac{\Delta t}{2}(g_j^n + (Lu^n)_j), & \sigma < 1 \\ u_j^n + \left(1 - \frac{1}{\sigma}\right)(u_j^{n+1} - u_j^n) + \frac{\Delta x}{2a_0}(g_j^n + (Lu^n)_j), & \\ \qquad \sigma \geq 1. & \end{cases} \tag{2.7}$$

As in the first example, $u_{j+1/2}$ is continuous as a function of σ at $\sigma = 1$. The explicit predictor formula may be derived by a Taylor series argument:

$$\begin{aligned} u_{j+1/2} &\sim u_j^n + \frac{\Delta x}{2}\frac{\partial u}{\partial x}\bigg|_{(x_j, t^n)} + \frac{\Delta t}{2}\frac{\partial u}{\partial t}\bigg|_{(x_j, t^n)} \\ &= u_j^n + \frac{\Delta x}{2}\frac{\partial u}{\partial x}\bigg|_{(x_j, t^n)} \\ &\quad - \frac{\Delta t}{2}\left(a_0 \frac{\partial u}{\partial x}\bigg|_{(x_j, t^n)} - g(x_j, u_j^n) - \mu \frac{\partial^2 u}{\partial x^2}\bigg|_{(x_j, t^n)}\right) \\ &\sim u_j^n + \frac{1}{2}(1-\sigma)\,\Delta u_j^n + \frac{\Delta t}{2}(g_j^n + (Lu^n)_j). \end{aligned}$$

Our motivation for the form of $u_{j+1/2}$ when $\sigma > 1$ is that we want any discrete steady solution to (2.6) to be a second-order accurate approximation to the solution of

$$a_0 \frac{\partial u}{\partial x} = g(x, u) + \mu \frac{\partial^2 u}{\partial x^2}. \tag{2.8}$$

If we have a discrete steady state, $u_j^{n+1} = u_j^n = u_j$ and the formula for $u_{j+1/2}$ can be obtained from a Taylor expansion in x:

$$\begin{aligned} u_{j+1/2} &\sim u_j + \frac{\Delta x}{2}\frac{\partial u}{\partial x} \\ &\sim u_j + \frac{\Delta x}{2a_0}(g(x_j, u_j) + L(u)_j). \end{aligned}$$

This is sufficient to ensure that steady solutions to (2.6) are second-order accurate approximations to solutions of (2.8). We remark here that schemes have been proposed which construct piecewise steady profiles in each computational zone, even far from steady state, and use the Riemann problem solution to resolve nonlinear wave interactions [15, 16, 24]. Our method may be viewed as an approximate version of this idea, at least near steay state.

The appearance of the viscous and source terms in the predictor, lagged to the old time level, is similar to the treatment of the advective terms in a recent version of the projection algorithm for incompressible flow [2]. A standard Fourier argument shows that the stability properties of the advection scheme remain for this treatment of the viscous terms. We do not have any corresponding stability results in the case of general source terms. We expect that lagging g_j^n in the calculation of $u_{j+1/2}$ will be stable, provided that the time scale of the source term is comparable to or greater than the discrete advection time scale, i.e., $\Delta x/a_0 \lesssim |\partial g/\partial u|^{-1}$. This condition holds, for example, for the geometric source terms discussed in Section 3. In that

case, we have also used $g_j^{n+1/2} = \frac{1}{2}(g(x_{j-1/2}, u_{j-1/2}) + g(x_{j+1/2}, u_{j+1/2}))$ in (2.6) and found it to be stable.

This method can be extended to the case of a convex scalar conservation law

$$\frac{\partial u}{\partial t} + \frac{\partial f}{\partial x} = 0, \tag{2.9}$$

$d^2f/du^2 > 0$. The special case $a(u) = df/du > 0$ is considered here; the discussion of sonic point behavior, as well as viscous and source terms for nonlinear problems is deferred to Section 3. The straightforward extension of the algorithm (2.2)–(2.3) to this case is given by

$$u_j^{n+1} = u_j^n - \frac{\Delta t}{\Delta x}(f(u_{j+1/2}) - f(u_{j-1/2})), \tag{2.10}$$

where the edge states are given by

$$u_{j+1/2} = \begin{cases} u_j^n + \frac{1}{2}(1-\sigma_j)\,\Delta u_j^n & \text{if} \quad \sigma_j < 1 \\ u_j^n + \left(1 - \frac{1}{\sigma_j}\right)(u_j^{n+1} - u_j^n) & \text{if} \quad \sigma_j \geq 1, \end{cases} \tag{2.11}$$

with $\sigma_j = a(u_j^n)\,\Delta t/\Delta x$. We now have a system of nonlinear equations for $\{u_j^{n+1}\}$. Their solution may be approximated by setting $u_j^{n+1} = u_j^n + \delta u_j^n$ and expanding (2.10) to terms of first order in δu_j^n. The result, for regions in which $\sigma_j > 1$, is a bidiagonal system of linear equations for δu_j^n:

$$\sigma_j \delta u_j^n - (\sigma_{j-1} - 1)\,\delta u_{j-1}^n = -\frac{\Delta t}{\Delta x}(f(u_j^n) - f(u_{j-1}^n)). \tag{2.12}$$

It is easy to see that there are nonlinear stability problems with this method. For example, let $f(u) = \frac{1}{2}u^2$ and consider the initial data

$$u_j^n = \begin{cases} u_L, & \text{if} \quad j < j_0 \\ 1, & \text{if} \quad j \geq j_0. \end{cases} \tag{2.13}$$

If $u_L > 1$ and $\Delta t \geq \Delta x$, then the method is implicit everywhere. For $j < j_0$ we have

$$\sigma_j \delta u_j^n - (\sigma_{j-1} - 1)\,\delta u_{j-1}^n = 0, \tag{2.14}$$

and steady boundary conditions imply that $\delta u_j^n = 0$ for $j < j_0$. If $j = j_0$, then we have

$$\delta u_{j_0}^n = -\frac{\Delta t}{\Delta x}\frac{[f(u_{j_0}^n) - f(u_{j_0-1}^n)]}{\sigma_{j_0}} = (u_L - 1)\frac{(u_L + 1)}{2}. \tag{2.15}$$

The exact solution is the initial discontinuity moving to the right at speed $(1 + u_L)/2$; therefore, δu_j^n should never exceed $(u_L - 1)$. As can be seen from (2.15), the numerical solution will produce a discontinuity moving to the right; however, the magnitude of the jump will increase as it moves.

The above instability is a result of the linearization about the old time level, especially the evaluation of σ in (2.11) using $a(u_j^n)$. The linearization is computed by expanding about the solution at time t^n with entries depending on the wave speed. If the wave speed changes by a large amount during the course of the time step, then the approximation of (2.10) by (2.12) will not be very accurate. If the large relative change in $a(u)$ is negative, so that our linearization uses an overestimate of the wave speed, this leads to a large diffusive error. If the large relative change in $a(u)$ is positive, so that our linearization uses an underestimate of the wave speed, the solution will oscillate and be nonlinearly unstable as in the above example. Our approach to eliminate this problem is to impose a constraint on the time step Δt such that the relative change in wave speed in each zone during the time step is not too great. That is, we impose

$$\sigma_{\text{imp}} > \frac{1}{a}\frac{\partial a}{\partial t}\Delta t = \frac{1}{a}\frac{da}{du}\frac{\partial u}{\partial t}\Delta t \sim \frac{1}{a}\frac{da}{du}(u_j^{n+1} - u_j^n), \tag{2.16}$$

where σ_{imp} is a dimensionless constant analogous to the CFL number for explicit methods; typically, $\sigma_{\text{imp}} < 1$. Given Δt^n, the constraint on the next time step is implemented as

$$\Delta t^{n+1} = \min_j \left\{ \max \times \left[\frac{\sigma_{\text{imp}}\,\Delta t^n}{\max\{((1/a)(da/du))_j^n (u_j^{n+1} - u_j^n), \varepsilon\}}, \frac{\Delta x_j}{a_j^n} \right] \right\}, \tag{2.17}$$

where $\varepsilon > 0$ is some arbitrarily small number. This can be viewed as a heuristic procedure used to prevent a shock from moving more than a fraction of one computational zone per time step. Note that there is a distinguished sign in this procedure; the time step is only limited at points where

a is increasing, which is where failure to do so could lead to an instability. Also, at steady state, the time step is bounded only by the arbitrary limit $\sigma_{\text{imp}}\,\Delta t^n/\varepsilon$.

3. CONSERVATION LAWS AND THEIR VISCOUS EXTENSIONS

To illustrate the extension of the ideas in the previous section to a system of equations in one space dimension, we first consider hyperbolic systems of the following form:

$$\frac{\partial U}{\partial t}+\frac{\partial A F^{\text{ad}}(U)}{\partial \Lambda}+\frac{\partial H(U)}{\partial x}=0,$$
$$U(x,t)=U\colon \mathscr{R}\times[0,T]\to\mathscr{R}^N. \tag{3.1}$$

Here $\Lambda=\Lambda(x)$ is a volume coordinate, with $A=A(x)=d\Lambda/dx>0$ the cross-sectional area associated with $\Lambda(x)$. F^{ad} and H are functions of U with F^{ad} representing the advective portion of the flux. For the purpose of performing a characteristic analysis of the above equations, we also use a nonconservative form, including a possible nonlinear change of variables:

$$\frac{\partial V}{\partial t}+\mathbf{A}(V)\frac{\partial V}{\partial x}=G(V,x). \tag{3.2}$$

Here $V=V(U)$ is an invertible function of U,

$$\mathbf{A}=(\nabla_V U)^{-1}\,\nabla_U(F^{\text{ad}}+H)(\nabla_V U),$$

and

$$G=-(\nabla_V U)^{-1}\frac{d}{dx}(\ln A)\,F^{\text{ad}}(U(V)).$$

We assume that $\mathbf{A}$ has $\bar N$ ($\bar N\leqslant N$) eigenvalues $\lambda^1\leqslant\cdots\leqslant\lambda^{\bar N}$ of fixed multiplicity corresponding to a complete set of linearly independent left and right eigenvectors, $[l^\nu, r^\nu]_{\nu\in(1,2,\ldots,N)}$: $l^\nu\mathbf{A}=\lambda^\nu l^\nu$, $\mathbf{A}r^\nu=\lambda^\nu r^\nu$; $l^\nu\cdot r^{\nu'}=\delta_{\nu\nu'}$.

The example of system (3.1) to be studied here is the equations of unsteady gas dynamics for the case of a one-dimensional duct of varying cross section; here, $N=3$ and Λ can be any smooth function with $A>0$. The dependent variables U and V and the fluxes F^{ad} and H are given by

$$U=\begin{pmatrix}\rho\\ \rho u\\ \rho E\end{pmatrix},\qquad V=\begin{pmatrix}\rho\\ u\\ p\end{pmatrix},$$
$$F^{\text{ad}}(U)=\begin{pmatrix}\rho u\\ \rho u^2\\ \rho u E+up\end{pmatrix},\qquad H(U)=\begin{pmatrix}0\\ p\\ 0\end{pmatrix},$$

where ρ is the density, u the velocity, E the total energy per unit mass, and p the pressure obtained from the equation of state: $p=p(\rho,e)$, $e=E-u^2/2$. In our work here, it is assumed that $p=(\gamma-1)\rho e$ with $\gamma=1.4$ constant. For this example, the system (3.1) may be written in the equivalent weak conservation form

$$\frac{\partial U}{\partial t}+\frac{\partial F(U)}{\partial x}=-\frac{A'(x)}{A(x)}F^{\text{ad}}(U), \tag{3.1'}$$

where $F=F^{\text{ad}}+H$.

Let $\{x_{j+1/2}\}$ be the edges of a finite difference mesh and Δt, a time increment. We will consider conservative finite difference approximations of the form

$$U_j^{n+1}=U_j^n-\frac{\Delta t}{\Delta\Lambda_j}(A_{j+1/2}F^{\text{ad}}_{j+1/2}-A_{j-1/2}F^{\text{ad}}_{j-1/2})$$
$$-\frac{\Delta t}{\Delta x_j}(H_{j+1/2}-H_{j-1/2}), \tag{3.3}$$

where $\Delta x_j=x_{j+1/2}-x_{j-1/2}$, $\Delta\Lambda_j=\Lambda(x_{j+1/2})-\Lambda(x_{j-1/2})$, $A_{j+1/2}=A(x_{j+1/2})$ and $F^{\text{ad}}_{j+1/2}$, $H_{j+1/2}$ are some approximation of the time averages of F^{ad} and H at $x_{j+1/2}$. Our strategy for computing $F^{\text{ad}}_{j+1/2}$, $H_{j+1/2}$ follows the general predictor–corrector formalism in [6, 7]. We compute the left and right states at each cell edge, $V_{j+1/2,L}$ and $V_{j+1/2,R}$, using the characteristic form of Eq. (3.2). These left and right states are then used to produce single-valued fluxes, using a variation of the Engquist–Osher procedure described in [3]. The details of this procedure will be described in the next section; for the moment, we will assume the existence of a flux function $(V_L, V_R)\to(F^{\text{ad}}(V_L,V_R), H(V_L,V_R))$ which is a C^1 function of its arguments in the absence of transonic shocks; it should also reduce to the flux evaluated along the ray $x/t=\xi=0$ of some approximate Riemann problem solver.

The predictor step for obtaining $V_{j+1/2,L}$, $V_{j+1/2,R}$ is derived, using a straightforward extension of the formalism for explicit methods in [6], by applying the advection scheme given above for a linear problem to the propagation of signals along each characteristic family. In particular, when all the characteristics have CFL numbers less than 1, we obtain an explicit second-order Godunov method similar to the one described in [6]. If any of the characteristics are implicit, we obtain a system of nonlinear equations for U^{n+1}. We approximate these in the usual fashion by linearizing them about U^n, leading to a block tridiagonal linear system.

Our basic predictor step is given as follows. We define, for example,

$$V_{j+1/2,L} = V_j^n + RD_L^S R^{-1} G_j^n + \frac{1}{2} RD_L^E R^{-1} \left(I - \mathbf{A}(V_j^n) \frac{\Delta t}{\Delta x} \right) \Delta V_j^n + RD_L^I R^{-1} (V_j^{n+1} - V_j^n), \tag{3.4}$$

where R is the matrix whose columns are the eigenvectors r^ν, $\nu = 1 \cdots N$, $V_j^n = V(U_j^n)$, and $G_j^n = G(x_j, V_j^n)$. The matrices D_L^S, D_L^E, D_L^I are diagonal matrices with entries

$$(D_L^S)_{\nu\nu} = \begin{cases} \dfrac{\Delta t}{2}, & \text{if } |\sigma_j^\nu| < 1, \\ \dfrac{\Delta t}{2\sigma_j^\nu}, & \text{otherwise;} \end{cases} \tag{3.5a}$$

$$(D_L^I)_{\nu\nu} = \begin{cases} 1 - \dfrac{1}{|\sigma_j^\nu|}, & \text{if } \sigma_j^\nu > 1, \\ 0, & \text{otherwise;} \end{cases} \tag{3.5b}$$

$$(D_L^E)_{\nu\nu} = \begin{cases} 1, & \text{if } 0 < \sigma_j^\nu < 1, \\ 0, & \text{otherwise.} \end{cases} \tag{3.5c}$$

Here $\sigma_j^\nu = (\Delta t/\Delta x_j)\, \lambda^\nu(V_j^n)$ is the local CFL number for the νth characteristic in the jth zone. The slopes ΔV_j^n are a finite difference approximation to $\partial V/\partial x|_{(x_j, t^n)} \Delta x_j$ to which monotonicity constraints have been applied, in the same fashion as for the explicit second-order Godunov methods in [6, 7]. Similarly, we can define

$$V_{j+1/2,R} = V_{j+1}^n + RD_R^S R^{-1} G_{j+1}^n - \frac{1}{2} RD_R^E R^{-1} \left(I + \mathbf{A}(V_{j+1}^n) \frac{\Delta t}{\Delta x} \right) \Delta V_{j+1}^n + RD_R^I R^{-1} (V_{j+1}^{n+1} - V_{j+1}^n), \tag{3.6}$$

where R is evaluated at V_{j+1}^n and D_R^S, D_R^E, D_R^I are defined analogously to (3.5), with the substitution $\sigma_j^\nu \to -\sigma_{j+1}^\nu$. It should be noted that the matrices D_L^S, D_R^S are not continuous functions of σ at $\sigma = -1$, $\sigma = 1$, respectively. However, the components which jump correspond to characteristics leading away from the interface and can be expected to be a higher order effect in the subsequent flux computation; consequently, the resulting fluxes remain continuous to leading order in the jump $|V_j - V_{j+1}|$.

There are several distinguished limits contained in the above procedure. First, if all the σ_j^ν have absolute value less than 1, then $D_L^S = D_R^S = (\Delta t/2) I$, $D_L^I = D_R^I = 0$ and we obtain a second-order accurate explicit scheme for Eqs. (3.1) similar to that in [6]. For those characteristics for which $|\sigma_j^\nu|, |\sigma_{j+1}^\nu| > 1$, we are applying, using the terms which have in them D_L^I, D_R^I, the implicit scheme to the characteristic variables to compute the left and right states. Finally, in the limit that all the characteristics are implicit (so that D_L^E, $D_R^E = 0$) and the solution is a steady state (so that $V^{n+1} = V^n$), then the calculation of the left and right states is obtained from a difference approximation to the steady state of the nonconservative form of Eqs. (3.2) which has a second-order local truncation error:

$$V_{j+1/2,L} = V_j^n + (\mathbf{A}(V_j^n))^{-1} G(V_j^n) \frac{\Delta x_j}{2}, \tag{3.7a}$$

$$V_{j+1/2,R} = V_{j+1}^n - (\mathbf{A}(V_{j+1}^n))^{-1} G(V_{j+1}^n) \frac{\Delta x_{j+1}}{2}. \tag{3.7b}$$

Thus, if the solution has reached a discrete steady state, then the difference equation it satisfies is a second-order accurate approximation to the steady equations

$$\frac{\partial AF}{\partial A} + \frac{\partial H}{\partial x} = 0. \tag{3.8}$$

We have found that the numerical steady state can be unstable at the sonic state of a transonic region if the computational zone with $\lambda_j^\nu \sim 0$ is explicit. Specifically, the instability arises in the situation for which $|\sigma_j^\nu| < 1$ and $|\sigma_{j\pm1}^\nu| > 1$ for some ν. As already noted, (3.7) is a second-order accurate extrapolation formula for the predicted edge values. However, in the case under consideration, the explicit predictor (3.4) and (3.6) will be in effect; since the sign of the correction in the explicit predictor from D_L^S and D_R^S is of the same sign for both the left and right edges, one of these corrections must be qualitatively wrong (compare with (3.7) and the implicit predictor). If the transonic region has a sufficiently strong jump, this error can lead to predicted values with wave speeds of the incorrect sign; the flux function (especially the Engquist–Osher flux) may not be able to correct this error and the instability is driven. Transonic instabilities of this type may be expected on theoretical grounds; see Liu [15, 16], where the problem is resolved in the context of a generalized random choice method by analyzing the possible stable asymptotic states and forcing the solution towards the correct one without computing the details of the unstable wave interactions. Such an approach does not appear feasible for finite difference methods, so we have made a modification which does not allow the instability to form in the first place. In the given situation, two changes are made to the predictor algorithm. First, the slopes are set to zero in the jth zone and, second, the explicit predictor is modified in the jth zone to avoid the sign error:

$$(D_L^S)_{\nu\nu} = \text{sign}(\lambda_j^\nu) \frac{\Delta t}{2}, \qquad (D_R^S)_{\nu\nu} = -\text{sign}(\lambda_j^\nu) \frac{\Delta t}{2}. \tag{3.9}$$

This procedure is locally qualitatively correct, while avoiding a division by the vanishing wave speed that would

be involved in using the steady equations. We note that the procedure described here lowers the order of accuracy to first order locally; however, it should be emphasized that this loss of accuracy does not occur near all sonic points—only those points satisfying all of the conditions enumerated above are affected.

Finally, the overall magnitude of the correction due to the source terms in the predictor step (3.4), (3.6) is limited to be less than the magnitude of the local gradient in the characteristic variables. To be precise, let $\delta q_{j+1/2,L}$, $q=\rho, u, p$ represent one of the components of the correction to $V_{j+1/2,L}$ in (3.4) due to the source terms. Then, $\delta q_{j+1/2,L}$ is limited in the following fashion:

$$\delta q_{j+1/2,L} \leftarrow \min(|\delta q_{j+1/2,L}|, \alpha \cdot |q_j - q_{j+1}|) \cdot \operatorname{sgn}(\delta q_{j+1/2,L}). \tag{3.10}$$

Here, α is set to be equal to a problem-dependent parameter α_0, satisfying the condition $0.6 \leqslant \alpha_0 \leqslant 1.0$ in the calculations presented below, except in the transonic explicit zones described above for which we use the more restrictive value $\alpha = 0.5$. The motivation for this change comes from the fact that, in the limit of a steady state, we are using the steady equations to extrapolate from the cell centers to the cell edges. This is analogous to the use to which the slopes are put in the explicit method; as in that case, we limit the extrapolated values so as not to exceed limits defined by the values in adjacent cell centers. The formula for the correction to $V_{j+1/2,R}$ is analogous to (3.10) with the subscript $j-1$ replacing $j+1$. For the geometric source terms considered here, the source terms typically satisfy the constraints given above, except in the neighborhood of a sonic point, so that at most points we are using Eqs. (3.4), (3.6) without modification to compute the solution. Similar considerations led van Leer in [24] to use a modified version of extrapolation of the steady-state equations in his explicit scalar algorithm.

Equations (3.3) are a system of nonlinear equtions which define the solution at the new time implicitly. We replace them in the usual fashion with a system of linear equations obtained by linearizing (3.3) about U^n while preserving the conservation form of the equations. This leads to a block tridiagonal system of equations for $\delta U_j = U_j^{n+1} - U_j^n$,

$$\begin{aligned} &\mathscr{A}_j \delta U_j + \mathscr{B}_j \delta U_{j+1} + \mathscr{C}_j \delta U_{j-1} \\ &\quad = -\frac{\Delta t}{\Delta \Lambda_j}(A_{j+1/2}\tilde{F}^{\mathrm{ad}}_{j+1/2} - A_{j-1/2}\tilde{F}^{\mathrm{ad}}_{j-1/2}) \\ &\qquad -\frac{\Delta t}{\Delta x_j}(\tilde{H}_{j+1/2} - \tilde{H}_{j-1/2}), \end{aligned} \tag{3.11}$$

where

$$\mathscr{A}_j = I + \frac{\Delta t}{\Delta \Lambda_j}(A_{j+1/2}\nabla_{U_j^{n+1}}F^{\mathrm{ad}}_{j+1/2} - A_{j-1/2}\nabla_{U_j^{n+1}}F^{\mathrm{ad}}_{j-1/2}) + \frac{\Delta t}{\Delta x_j}(\nabla_{U_j^{n+1}}H_{j+1/2} - \nabla_{U_j^{n+1}}H_{j-1/2}), \tag{3.12a}$$

$$\mathscr{B}_j = \frac{\Delta t}{\Delta \Lambda_j}A_{j+1/2}\nabla_{U_{j+1}^{n+1}}F^{\mathrm{ad}}_{j+1/2} + \frac{\Delta t}{\Delta x_j}\nabla_{U_{j+1}^{n+1}}H_{j+1/2}, \tag{3.12b}$$

$$\mathscr{C}_j = -\frac{\Delta t}{\Delta \Lambda_j}A_{j-1/2}\nabla_{U_{j-1}^{n+1}}F^{\mathrm{ad}}_{j-1/2} - \frac{\Delta t}{\Delta x_j}\nabla_{U_{j-1}^{n+1}}H_{j-1/2}, \tag{3.12c}$$

with $\tilde{F}^{\mathrm{ad}}_{j+1/2}$, $\tilde{H}_{j+1/2}$ obtained by setting $U^{n+1} = U^n$ in (3.4) and (3.6) and with all the Jacobians being evaluated at U^n.

The Jacobians of $F^{\mathrm{ad}}_{j+1/2}$, $H_{j+1/2}$ are evaluated using the chain rule; for example,

$$\nabla_{U_j^{n+1}}F^{\mathrm{ad}}_{j+1/2} = \nabla_{V_{j+1/2,L}}F^{\mathrm{ad}}_{j+1/2} \cdot \nabla_{V_j^{n+1}}V_{j+1/2,L} \cdot \nabla_{U_j^{n+1}}V_j^{n+1}. \tag{3.13}$$

The third term is simply the Jacobian of the change of variables transformation, while the second term is $RD_L^I R^{-1}$. The first term is the Jacobian of the flux function defined in the next section with respect to the input left state; it can either be computed analytically, or obtained by numerical differentiation. In the calculations presented in Section 5, we have done the latter.

One can also use these ideas in the context of parabolic extensions of the system (3.1). In that case, the equations we wish to solve are of the form

$$\frac{\partial U}{\partial t} + \frac{\partial}{\partial \Lambda}(A(F^{\mathrm{ad}}(U) - b(U, U_x)) + \frac{\partial H}{\partial x} = 0, \tag{3.14}$$

where b is a nonlinear parabolic term and a function of U and $U_x = \partial U/\partial x$. For the one-dimensional duct of varying cross section b is given by

$$b(U, U_x) = \begin{pmatrix} 0 \\ \mu u_x \\ \mu(\frac{1}{2}u^2)_x + \kappa T_x \end{pmatrix}. \tag{3.15}$$

Here T is the temperature, which is some function of (p, e); μ and κ are the viscosity and thermal conductivity, which can also depend on (p, e). The explicit functional form of b in terms of U and U_x is obtained by substituting the expressions in conserved variables defining u and T and by expanding the derivatives.

We discretize b in space by applying the standard difference and averaging operators to U:

$$b^n_{j+1/2} = b(U^n_{j+1/2}, (D_x U)^n_{j+1/2}), \tag{3.16a}$$

$$U^n_{j+1/2} = \tfrac{1}{2}(U^n_j + U^n_{j+1}), \tag{3.16b}$$

$$(D_x U)^n_{j+1/2} = \frac{(U^n_{j+1} - U^n_j)}{\tfrac{1}{2}(\Delta x_j + \Delta x_{j+1})}. \tag{3.16c}$$

The presence of viscous terms introduces changes in both the predictor and corrector steps. In the predictor, the only change is to modify the source term appearing in the discretized characteristic equations (3.4) and (3.6) defining the left and right states:

$$G^n_j \leftarrow G^n_j + ((\nabla_V U)^n_j)^{-1} \frac{1}{\Delta A_j}(A_{j+1/2} b^n_{j+1/2} - A_{j-1/2} b^n_{j-1/2}). \tag{3.17}$$

In viscous calculations, the source term limiter correction (3.10) is applied to the geometric and viscous source terms together, that is, to G^n_j given by (3.17). With these modifications, which depend only on the solution at the old time, the calculation of $F^{ad}_{j+1/2}$, $H_{j+1/2}$ proceeds as before.

The inviscid conservative differencing step (3.3) is modified by the addition of appropriately time-centered viscous terms,

$$\begin{aligned} U^{n+1}_j = U^n_j &- \frac{\Delta t}{\Delta A_j}(A_{j+1/2} F^{ad}_{j+1/2} - A_{j-1/2} F^{ad}_{j-1/2}) \\ &- \frac{\Delta t}{\Delta x_j}(H_{j+1/2} - H_{j-1/2}) \\ &+ (1-\psi)\frac{\Delta t}{\Delta A_j}(A_{j+1/2} b^n_{j+1/2} - A_{j-1/2} b^n_{j-1/2}) \\ &+ \psi \frac{\Delta t}{\Delta A_j}(A_{j+1/2} b^{n+1}_{j+1/2} - A_{j-1/2} b^{n+1}_{j-1/2}). \end{aligned} \tag{3.18}$$

Here ψ is a weighting factor; for $\psi = \frac{1}{2}$ we obtain Crank–Nicholson differencing, while $\psi = 1$ corresponds to backward Euler differencing, for the viscous terms only. The latter, $\psi = 1$, may be more efficient for applications involving convergence to steady state. This system can be linearized about U^n to obtain a block tridiagonal system for $U^{n+1} - U^n$. The viscous flux is linearized separately with respect to U and U_x:

$$\begin{aligned} b^{n+1}_{j+1/2} &= b(U^{n+1}_{j+1/2}, (D_x U)^{n+1}_{j+1/2}) \\ &\approx b(U^n_{j+1/2}, (D_x U)^n_{j+1/2}) \\ &\quad + (\nabla_U b)\cdot(U^{n+1}_{j+1/2} - U^n_{j+1/2}) \\ &\quad + (\nabla_{U_x} b)\cdot((D_x U)^{n+1}_{j+1/2} - (D_x U)^n_{j+1/2}). \end{aligned} \tag{3.19}$$

Here the derivatives $\nabla_U b$, $\nabla_{U_x} b$ are evaluated at $U^n_{j+1/2}$, $(D_x U)^n_{j+1/2}$. To obtain full second-order accuracy for the viscous terms, it would be necessary to iterate once on this procedure to obtain time-centered values for $\nabla_U b$, $\nabla_{U_x} b$; we have not done so for the results presented here. In any case, the introduction of viscous stresses adds only terms arising from the linearization of (3.19) to the block tridiagonal system (3.11); in particular, the coefficients in the linear system coming from the linearization of $F^{ad}_{j+1/2}$ and $H_{j+1/2}$ are the same as in the inviscid case. As in the inviscid case, the overall method is unconditionally stable and is second-order accurate in space for steady state solutions.

4. RIEMANN PROBLEMS AND FLUX FUNCTIONS

We now turn to the construction of the flux function

$$(V_L, V_R) \rightarrow (F^{ad}(V_L, V_R), H(V_L, V_R)),$$

required in the previous section. Our approach is a variation on the approximate Engquist–Osher flux function described in [3]. For the purpose of the present exposition, we will make an additional structural assumption on the system (3.1). We assume that each family of eigenvectors and eigenvalues satisfy either $r^\nu \cdot \nabla_V \lambda^\nu \neq 0$ or $r^\nu \cdot \nabla_V \lambda^\nu \equiv 0$, i.e., that the field is either genuinely nonlinear or linearly degenerate [14]. Note that these conditions are independent of any change of variables $U \rightarrow V(U)$. This restriction will simplify the discussion which follows; the more general case can be dealt with using the techniques in [3]. We denote by $\Sigma^\nu(V_S, \varepsilon)$ the simple wave curve through V_S, i.e., the solution to the ordinary differential equation,

$$\frac{d\Sigma^\nu}{d\varepsilon} = r^\nu(V), \qquad \Sigma^\nu(V_S, 0) = V_S \tag{4.1}$$

for $-\varepsilon_0 < \varepsilon < \varepsilon_0$. If the νth family is genuinely nonlinear, we assume a choice of sign for r^ν so that

$$\frac{d}{d\varepsilon}(\lambda^\nu(\Sigma^\nu(V_S, \varepsilon))) = \nabla\lambda^\nu \cdot r^\nu > 0. \tag{4.2}$$

The system (3.1) can also have weak solutions which are bounded, but discontinuous. If $V(x, t)$ consists of two constant states V_L, V_R, separated by a discontinuity propagating at speed s, then s, V_L, V_R satisfy the Rankine–Hugoniot relations

$$F(U(V_L)) - F(U(V_R)) = s(U(V_L) - U(V_R)). \tag{4.3}$$

Given V_S, then for each $\nu = 1, ..., N$ there exists a one-parameter family of states $\Gamma^\nu(V_S, \varepsilon)$, such that

$$F(U(\Gamma^\nu(V_S, \varepsilon))) - F(U(V_S)) = s^\nu(\varepsilon)(U(\Gamma^\nu(V_S, \varepsilon)) - U(V_S)), \tag{4.4}$$

with $d\Gamma^\nu/d\varepsilon|_{\varepsilon=0} = r^\nu(V_S)$, $\Gamma^\nu(V_S, 0) = V_S$. Thus the Hugoniot curves $\Gamma^\nu(V_S, \varepsilon)$ are tangent to the simple wave curves $\sum^\nu(V_S, \varepsilon)$ at U_S; in fact, the parameterization can be chosen in such a way that the second derivatives agree at $\varepsilon = 0$, so that

$$\Gamma^\nu(V_S, \varepsilon) = \sum^\nu(V_S, \varepsilon) + O(\varepsilon^3). \tag{4.5}$$

Finally, if the νth family of waves is linearly degenerate, then parameterizations can be chosen so that the simple wave curves and Hugoniot curves coincide.

We will consider the Riemann problem for (3.1), defined to be the initial value problem with initial data,

$$U(x, 0) = \begin{cases} U_L, & \text{if } x < 0 \\ U_R, & x \geqslant 0. \end{cases} \tag{4.6}$$

We will assume that the Riemann prolem has a unique piecewise continuous weak solution, which depends only on the similarity variable $x/t = \xi$ and which satisfies the Lax entropy conditions, i.e., has no rarefaction shocks. This solution consists of $N+1$ constant states, separated by N waves each associated with one of the N modes of wave propagation for the linearized system:

$$V(t, \xi t) = \begin{cases} V^0 = V_L, & \xi < \xi_L^1 \\ V^\nu, & \xi_R^\nu < \xi < \xi_L^{\nu+1}, \quad \nu = 1, ..., N-1 \\ V^N = V_R, & \xi > \xi_R^N. \end{cases} \tag{4.7}$$

If the νth family of waves is genuinely nonlinear, then there are two possible cases. If $\lambda^\nu(V^{\nu-1}) > \lambda^\nu(V^\nu)$, then the two states are connected by a Hugoniot curve, so that they satisfy (4.4), with shock speed $s^\nu = \xi_L^\nu = \xi_R^\nu$ satisfying $\lambda^\nu(V^{\nu-1}) > s^\nu > \lambda^\nu(V^\nu)$. If $\lambda^\nu(V^{\nu-1}) < \lambda^\nu(V^\nu)$, then the states $V^{\nu-1}$, V^ν are connected by a simple wave curve; $\lambda^\nu(V^{\nu-1}) = \xi_L^\nu$, $\lambda^\nu(V^\nu) = \xi_R^\nu$, and the values taken on by V for $\xi_L^\nu < \xi < \xi_R^\nu$ lie on the simple wave curve connecting the two states, with $\lambda^\nu(V(t, \xi t)) = \xi$. If the νth family of waves is linearly degenerate, then the pair $(V^{\nu-1}, V^\nu)$ is connected by a Hugoniot/simple wave curve, so that the νth wave is a discontinuity propagating at speed $s^\nu = \lambda^\nu(V^{\nu-1}) = \lambda^\nu(V^\nu)$ with $\xi_L^\nu = \xi_R^\nu = s^\nu$.

The solution to the Riemann problem may be separated into two parts: the solution in phase space, defined to be the calculation of the states V^ν, $\nu = 0, ..., N$; and the solution in physical space, defined to be the value of the solution at a point $(t, \xi t)$. This is a useful division of the problem since entropy conditions, to leading order in the strength of the waves, affect only the solution in physical space; it follows from (4.5) that the solution in phase space is independent of any entropy conditions up to terms of third order in the strength of the genuinely nonlinear waves. This means that in constructing approximate solutions to the Riemann problem for finite difference methods it is possible to ignore entropy conditions in constructing the approximate solution in phase space and only take them into account in the evaluation of the solution in physical space [6, 17].

We can define approximate phase space solutions to the Riemann problem. Let $\bar{\mathbf{R}}(V_L, V_R) = (\bar{r}^1, ..., \bar{r}^N)$ be an $N \times N$ matrix-valued function of the left and right states V_L, V_R. We assume that $\bar{R}$ is nonsingular and that $V_L - V_R = 0$ implies that $\bar{\mathbf{R}} = \mathbf{R}(V) = (r^1, ..., r^N)$, the matrix of right eigenvectors of $\mathbf{A}(V)$ evaluated at $V = V_L = V_R$. $\bar{\mathbf{R}}$ defines an approximate phase space solution to the Riemann problem via

$$V^\nu = V_L + \sum_{\nu' \leqslant \nu} \alpha^{\nu'} \bar{r}^{\nu'}, \qquad \boldsymbol{\alpha} = \mathbf{R}^{-1} \cdot (V_R - V_L). \tag{4.8}$$

Then V^ν differs from the exact phase space solution by $O(\|V_L - V_R\|^2)$. One such approximate phase space solution is due to Osher and Solomon [17] and is constructed by connecting the V^ν by simple waves. We construct a related version due to Bell *et al.* [3] next.

Define the auxiliary function

$$W(V_L, \boldsymbol{\varepsilon}), \qquad \boldsymbol{\varepsilon} \in \mathscr{R}^N$$

by recursion in ν. If we denote by ε_ν the νth entry in $\boldsymbol{\varepsilon}$, and $\boldsymbol{\varepsilon}|_\nu = (\varepsilon_1, ..., \varepsilon_\nu, 0, ..., 0)$; then we have

$$\begin{aligned} W^1(V_L, \boldsymbol{\varepsilon}|_1) &= \Sigma^1(V_L, \varepsilon_1), \\ W^\nu(V_L, \boldsymbol{\varepsilon}|_\nu) &= \Sigma^\nu(W^{\nu-1}(V_L, \boldsymbol{\varepsilon}|_{\nu-1}), \varepsilon_\nu), \\ W(V_L, \boldsymbol{\varepsilon}) &= W^N(V_L, \boldsymbol{\varepsilon}). \end{aligned} \tag{4.9}$$

For $\boldsymbol{\varepsilon}$ sufficiently small, there is a function $\boldsymbol{\varepsilon}(V)$ such that $W(V_L, \boldsymbol{\varepsilon}(V)) = V$ [14]. Then the states given by $V^\nu = W^\nu(V_L, \boldsymbol{\varepsilon}_\nu(V_R))$ define an approximate phase space solution. The generalized right eigenvectors are defined as

$$\bar{r}^\nu = \begin{cases} \dfrac{V^\nu - V^{\nu-1}}{\varepsilon_\nu(V_R)}, & \text{if } \varepsilon_\nu(V_R) \neq 0 \\ r^\nu(V^\nu), & \text{if } \varepsilon_\nu(V_R) = 0. \end{cases} \tag{4.10}$$

We can use this approximate phase space solution to compute fluxes. The starting point for this will be the

Engquist–Osher (E–O) flux function for systems of conservations laws [11, 17], e.g., (3.1′) with zero right-hand side. We start with the simple wave approximate phase space solution to the Riemann problem. Then the E–O flux is defined to be

$$F(V_L, V_R) = F(V_L) + \sum_{\nu=1}^{N} \int_0^{\varepsilon^\nu} \lambda^\nu_-(\varepsilon')\, r^\nu(\varepsilon')\, d\varepsilon'$$

$$= F(V_R) - \sum_{\nu=1}^{N} \int_0^{\varepsilon^\nu} \lambda^\nu_+(\varepsilon')\, r^\nu(\varepsilon')\, d\varepsilon'. \quad (4.11)$$

Here $a_+ = \max(a, 0)$, $a_- = \min(a, 0)$, and $\lambda^\nu(\varepsilon)$, $r^\nu(\varepsilon)$ are the values of λ^ν, r^ν along the νth simple wave curve: $(\lambda^\nu, r^\nu)(\varepsilon) = (\lambda^\nu, r^\nu)(\Sigma^\nu(V^{\nu-1}, \varepsilon))$. We can use the identity

$$\lambda^\nu(\varepsilon)\, r^\nu(\varepsilon) = A(\varepsilon)\, r^\nu(\varepsilon) = \frac{d}{d\varepsilon} F(\Sigma^\nu(V^{\nu-1}, \varepsilon))$$

to evaluate the integrals in (4.11) analytically:

$$\begin{aligned} F(V_L, V_R) = F(V_L) &+ \sum_{\nu:\, \lambda^\nu(V^{\nu-1}),\, \lambda^\nu(V^\nu) < 0} F(V^\nu) - F(V^{\nu-1}) \\ &+ \sum_{\nu:\, \lambda^\nu(V^{\nu-1}) < 0 < \lambda^\nu(V^\nu)} F(V^{\nu,\text{sonic}}) - F(V^{\nu-1}) \\ &+ \sum_{\nu:\, \lambda^\nu(V^{\nu-1}) > 0 > \lambda^\nu(V^\nu)} F(V^\nu) - F(V^{\nu,\text{sonic}}). \end{aligned} \quad (4.12)$$

Here $V^{\nu,\text{sonic}}$ is the unique (by genuine nonlinearity) point along the wave curve $\Sigma^\nu(V^{\nu-1}, \cdot)$ at which λ^ν vanishes. From the point of view of the scheme discussed in the previous section, $F(V_L, V_R)$ defined by (4.12) is the flux function of choice for our implicit–explicit method. In the absence of transonic waves, it reduces to a suitably accurate approximation to the Riemann problem. In the presence of transonic waves, it is dissipative and does not lead to entropy-violating shocks. Finally, it is a C^1 function of V_L, V_R, as shown in [17].

The principal difficulties with the E–O flux for systems as defined above are the expense of the computation of the simple wave phase space solution and the computation of sonic points $V^{\nu,\text{sonic}}$. It is natural to generalize the E–O flux for any approximate phase space solution by using (4.12) as the definition; given an approximate phase space solution (4.8), a particularly simple approximation for $V^{\nu,\text{sonic}}$ is given by linear interpolation:

$$V^{\nu,\text{sonic}} = V^{\nu-1} + \frac{\lambda^\nu(V^{\nu-1})}{\lambda^\nu(V^{\nu-1}) - \lambda^\nu(V^\nu)} (V^\nu - V^{\nu-1}). \quad (4.13)$$

With this definition, F_{E-O} is a continuous piecewise C^1 function of its arguments, reduces to a suitable approximation to the Riemann problem in the absence of transonic shocks, and is dissipative near transonic shocks. In general, it fails to be continuously differentiable; however, it differs from the exact E–O flux by terms which are $O(|V_L - V_R|^2)$. This is in contrast to the flux obtained from evaluating the exact solution to the Riemann problem, which differs from the exact E–O flux by $O(|V_L - V_R|)$. We have found our approximate E–O flux to be sufficiently smooth to give well-behaved Jacobians and convergence to steady state in the implicit–explicit methods described in the previous section.

An approximate phase space solution suitable for gas dynamics is now constructed. Define

$$\bar{\mathbf{R}}(V_L, V_R) = \begin{pmatrix} 1 & 1 & 1 \\ -\tau_L c_L & 0 & \tau_R c_R \\ c_L^2 & 0 & c_R^2 \end{pmatrix}, \quad (4.14)$$

where $\tau = \rho^{-1}$ is the specific volume and c is the sound speed, $c^2 = \gamma p/\rho$. This leads to

$$\alpha = \begin{pmatrix} \rho_{*L} - \rho_L \\ \rho_{*R} - \rho_{*L} \\ \rho_R - \rho_{*R} \end{pmatrix} \quad (4.15)$$

and

$$V^1 = \begin{pmatrix} \rho_{*L} \\ u_* \\ p_* \end{pmatrix}, \qquad V^2 = \begin{pmatrix} \rho_{*R} \\ u_* \\ p_* \end{pmatrix}, \quad (4.16)$$

where u_*, p_* satisfy the simultaneous linear equations

$$\begin{aligned} p_* - p_L &= -C_L(u_* - u_L) \\ p_* - p_R &= +C_R(u_* - u_R), \end{aligned} \quad (4.17)$$

C is the Lagrangian sound speed, $C = \rho c$, and

$$\rho_{*L,R} = \rho_{L,R} + \frac{p_* - p_{L,R}}{c_{L,R}^2}. \quad (4.18)$$

Geometrically, this procedure is equivalent to constructing wave curves through $U_{L,R}$ in the (p, u)-plane and intersecting the two tangent lines at $U_{L,R}$ to obtain (p_*, u_*); see [6].

A modification to this local linearization procedure has proven necessary in practice. Given $\rho_{*L,R}$ from (4.18) above, we compute $C^2_{*L,R} = \gamma p_* \rho_{*L,R}$, set

$$W^2_{L,R} = \begin{cases} C^2_{*L,R}, & \text{if } p_* > p_{L,R}, \\ C^2_{L,R}, & \text{otherwise,} \end{cases} \quad (4.19)$$

and solve for u_* using $W_{L,R}$, instead of $C_{L,R}$; that is,

$$u_* = \frac{W_L u_L + W_R u_R + p_L - p_R}{W_L + W_R}. \tag{4.20}$$

Finally, since u_* is obtained from (4.20), in the implementation of (4.17) only p_* is calculated and the provisional value of u_* is not needed:

$$p_* = \frac{C_L p_R + C_R p_L - C_L C_R (u_R - u_L)}{C_L + C_R}. \tag{4.17'}$$

The local linearization (4.14)–(4.18) with the nonlinear correction (4.19)–(4.20) was used in conjunction with (4.12) and the approximate sonic point formula (4.13) for the computations presented in Section 5 below. This approach may be contrasted to that suggested in [6] for which the wave curves in the Riemann problem are replaced everywhere by Hugoniot curves leading to a less expensive iteration than that resulting from (4.8)–(4.10). Indeed, if one replaces $C_{L,R}$ by $W_{L,R}$ in (4.14) and (4.17) above and replaces (4.18) by

$$\tau_{*L,R} = \tau_{L,R} - \frac{p_* - p_{L,R}}{W_{L,R}^2},$$

then the nonlinear iteration of [6] is obtained where $W_{L,R}$ are nonlinear functions of p_*; this method is effective even when only a few iterations are used and convergence is not attempted [6]. Still, this is more expensive than (4.14)–(4.20) and does not improve the results for the calculations of Section 5. Other approximate phase space solutions (4.8), e.g., Roe's method [22], could also be used with our method but we have not done so here.

The flux obtained from (4.12)–(4.20) is modified by adding a quadratic artificial viscosity of a type similar to that used in [3]. Specifically,

$$F(V_L, V_R) \leftarrow F(V_L, V_R) + \nu(V_L, V_R)(U_L - U_R), \tag{4.21}$$

where

$$\nu(V_L, V_R) = \varepsilon \cdot \max(u_L - u_R, 0). \tag{4.22}$$

We set the constant $\varepsilon = 0.1$, although in many calculations the additional dissipation provided by (4.21) is not required and $\varepsilon = 0$ is used.

Finally, we can easily extend the E–O flux to the case where the flux is split into advective and gradient terms. In that case, we have

$$\begin{aligned} F^{\mathrm{ad}}(V_L, V_R) = F(V_L) &+ \sum_{\nu:\lambda_L^\nu, \lambda_R^\nu < 0} F^{\mathrm{ad}}(V^\nu) - F^{\mathrm{ad}}(V^{\nu-1}) \\ &+ \sum_{\nu:\lambda_L^\nu < 0 < \lambda_R^\nu} F^{\mathrm{ad}}(V^{\nu,\mathrm{sonic}}) - F^{\mathrm{ad}}(V^{\nu-1}) \\ &+ \sum_{\nu:\lambda_L^\nu > 0 > \lambda_R^\nu} F^{\mathrm{ad}}(V^\nu) - F^{\mathrm{ad}}(V^{\nu,\mathrm{sonic}}) \end{aligned} \tag{4.23}$$

and similarly for H.

5. NUMERICAL RESULTS

The scheme's performance for linear advection is assessed by studying the propagation of a slowly moving contact discontinuity. The test problem is defined by setting the density ratio to be 10, the velocity field to be 0.02, and the pressure adjusted in four separate calculations to give sonic CFL numbers of 0.837, 8.37, 83.7, and 837.0. The time step is fixed so that the contact surface moves from its initial position in the fourth zone of a 32-zone mesh to its final location with the leading edge at $x = 0.8$ in exactly 1000 time steps. The results were the same for each calculation with four to five points in the numerical contact surface at the end of the calculation. Figure 1 shows the final density profile for the CFL = 837.0 calculation. The monotone profile illustrates the maximum principle satisfied by the implicit advection scheme, as well as its unconditional stability.

The nonlinear time step control (2.16)–(2.17) is extended to systems in a mode-by-mode fashion. For the acoustic modes ($\nu = 1$ or $\nu = 3$), the constraint is given by

$$\frac{1}{\lambda^\nu} (r^\nu \cdot \nabla_U \lambda^\nu)\, l^\nu \cdot (U_j^{n+1} - U_j^n) < \sigma_{\mathrm{imp}}, \tag{5.1}$$

where λ^ν, r^ν, l^ν, and $\nabla_U \lambda^\nu$ are all evaluated at U_j^n. Note that the analogous control for the linearly degenerate mode is

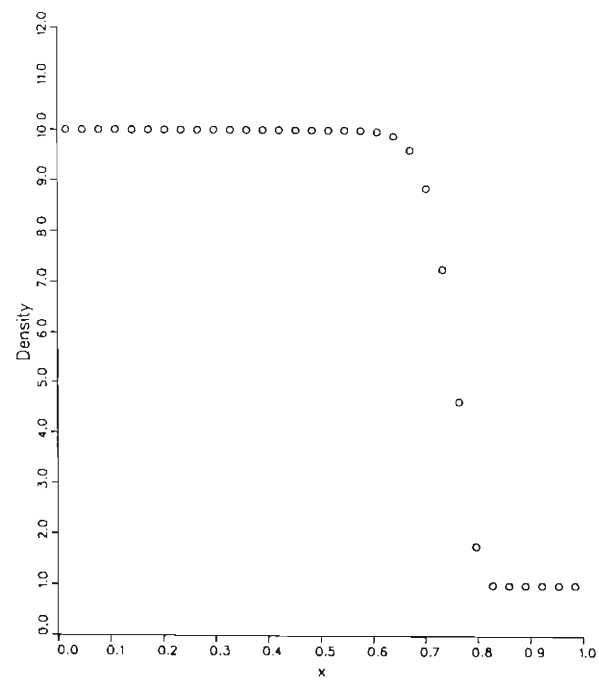

FIG. 1. Density vs. distance for the computed solution of a slowly moving contact surface after 1000 time steps on a 32-zone mesh; CFL = 837.

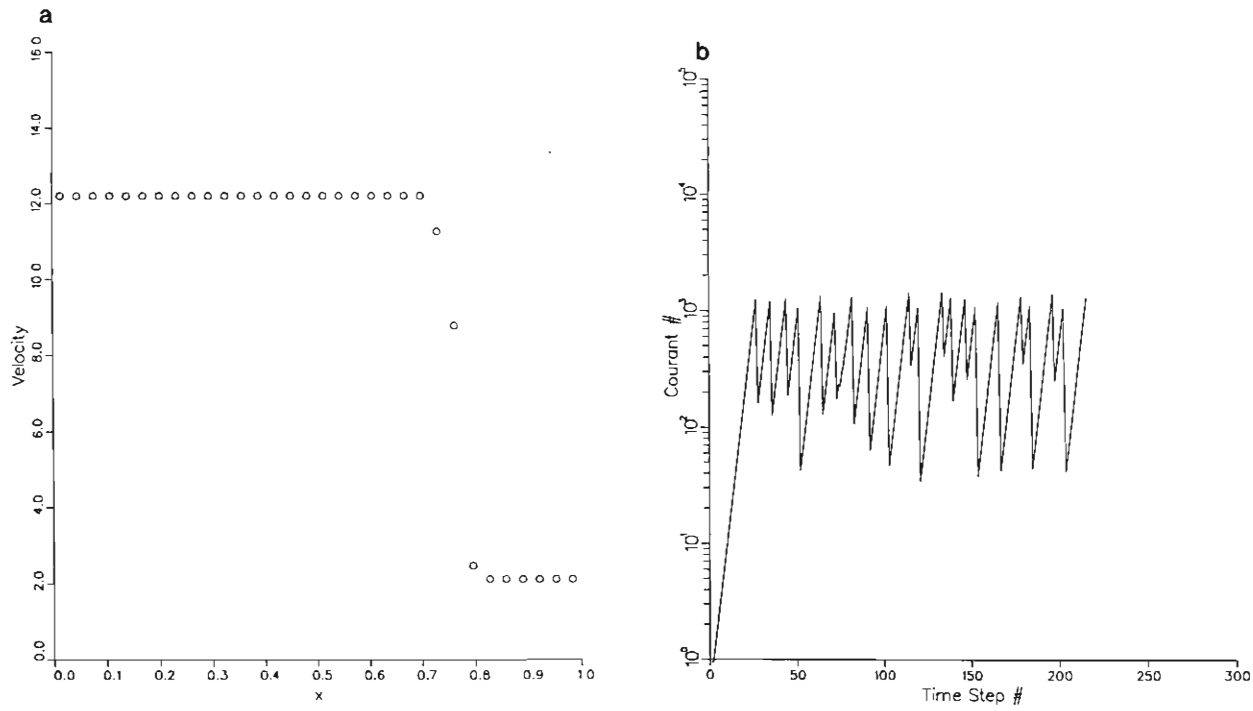

FIG. 2. (a) Velocity vs. distance for the computed solution of a slowly moving shock, shock speed = 0.003, after approximately 220 time steps on a 32-zone mesh; the jump is given by $(\rho, u, p)_L = (1.0, 12.2, 1.0)$ and $(\rho, u, p)_R = (5.73, 2.13, 123.89)$. (b) CFL number vs. time step for the calculation of Fig. 2a.

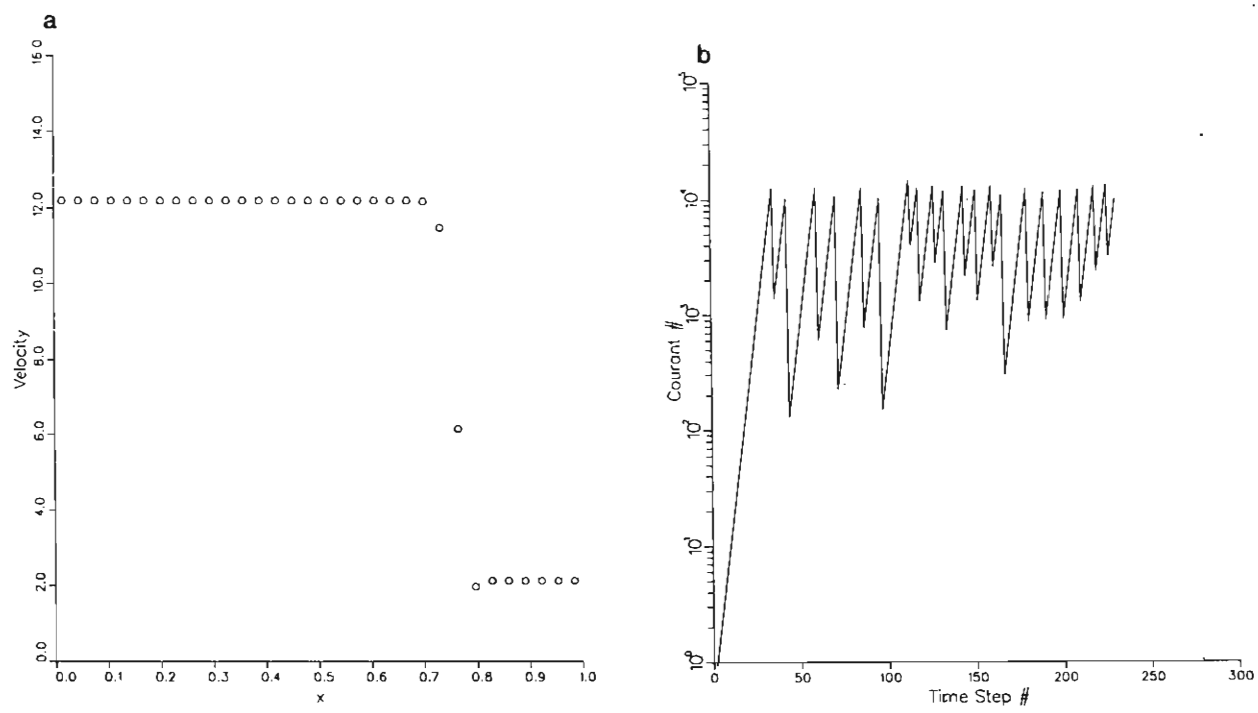

FIG. 3. (a) Velocity vs. distance for the computed solution of a slowly moving shock, shock speed = 0.0003, after approximately 240 time steps on a 32-zone mesh; the left and right states are similar to those of Fig. 2 with the velocities slightly adjusted to obtain the desired shock speed. (b) CFL number vs. time step for the calculation of Fig. 3a.

vacuous since $r \cdot \nabla_U \lambda$ is identically zero in that case. Given Δt^n, this is implemented as

$$\Delta t_{j,\nu}^{n+1} = \begin{cases} \dfrac{\sigma_{\text{imp}} \lambda^\nu \Delta t^n}{(r^\nu \cdot \nabla_U \lambda^\nu)\, l^\nu \cdot (U_j^{n+1} - U_j^n)}, & \text{if positive} \\ \sigma_{\max} \dfrac{\Delta x_j}{\lambda^\nu}, & \text{otherwise} \end{cases} \tag{5.2}$$

$$\Delta t^{n+1} = \min \left\{ \min_{j,\nu} \max \left[\Delta t_{j,\nu}^{n+1}, \frac{\Delta x}{\lambda^\nu} \right], \beta\, \Delta t^n \right\}, \tag{5.3}$$

where $\sigma_{\max}$ is an (arbitrary) overall ceiling for the time step and the factor $\beta > 1$ controls the growth of Δt from time step to time step. For the calculations presented here, we have taken $1.1 < \beta < \frac{4}{3}$ and $\sigma_{\max} = 10^{10}$. The condition (5.1) typically constrains any shock wave present in the problem to not propagate more than some fraction less than one of Δx per time step.

An important test of our implicit–explicit strategy and the nonlinear time step control is provided by the test problem of a slowly moving shock wave; i.e., a large number of explicit time steps would be required to move the shock across one computational zone. Ideally, one would like the time step to adjust in such a way that the computational time required to move the shock a fixed distance does not depend strongly on the shock speed and is a considerable reduction from that given by the explicit time step. To verify this behavior, two slowly moving shock calculations were performed with shock speeds of 0.003 and 0.0003, both with a pressure ratio of 123.89, on a 32-zone mesh with the initial shock location in the fourth zone. The shocks here are very strong with a nearly limiting value of the density ratio of 5.73. For these calculations, the artificial viscosity is turned on. Figures 2a and 3a show the velocity profiles at the end of the calculation (defined by the leading edge of the shock reaching $x = 0.8$) for the faster and slower shock, respectively. Figures 2b and 3b are the corresponding plots of the CFL number versus the time step. One notes that the shock profiles are stable, nearly monotone (a small undershoot appears at the leading edge in part of the cycle), and two to three zones wide. Additionally, the slower shock requires only an additional 10% more time steps; indeed, the result is actually better than this when one considers the extra initial rise time to the CFL maximum in Fig. 3b, relative to Fig. 2b. If these calculations had satisfied the explicit CFL condition then 10^6–10^7 time steps would have been required to move these shocks across the given mesh. The sawtooth nature of Figs. 2b and 3b is easily explained by the near periodicity of the discrete traveling wave profile as it traverses the mesh and the effect of the nonlinear time step control (5.1)–(5.3).

Several additional calculations were performed for the slowly moving shock test problem with different variants of the flux function. The most important remark is that if the

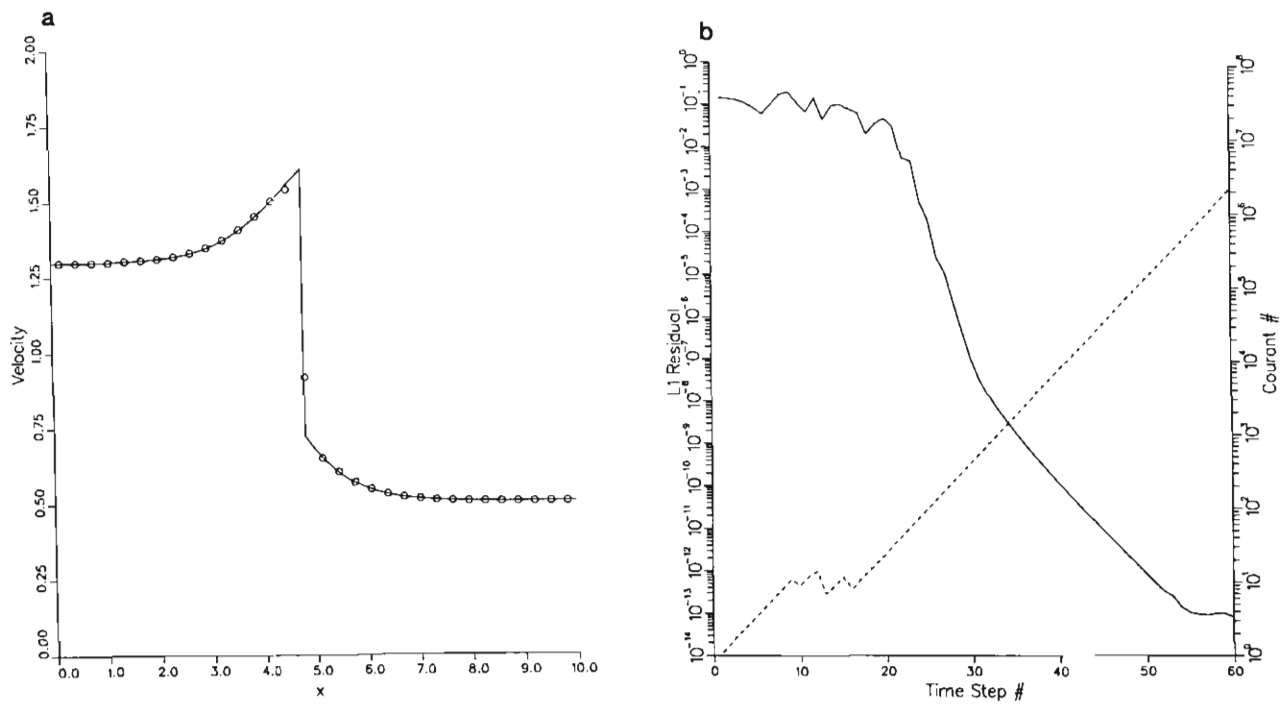

FIG. 4. (a) Velocity vs. distance along the expanding inviscid supersonic–subsonic duct of Ref. [23]; the solid line represents the exact solution and the circles are the converged (steady state) computed solution on a 32-zone mesh. (b) L^1 residual for density (solid line) and CFL number vs. time step for the calculation leading to Fig. 4a (dashed line).

E–O flux (4.12) is replaced by the Godunov flux (even with a nearly converged Riemann problem solver) then the calculations fail catastrophically at high CFL numbers. With respect to the artificial viscosity used for these results, it is interesting to note that if the approximate phase space solution (4.14)–(4.20) is replaced by the Hugoniot iteration used in [6] then the calculations are also successful with results similar to those reported above even if the artificial viscosity is turned off; on the other hand, the above calculations fail without the extra dissipation provided by the artificial viscosity (but they can be made to run successfully by changing the parameters σ_{imp} and β; however, the necessary changes lead to very poor performance in achieving high CFL numbers relative to the results presented here).

A recent study [21] of the application of explicit upwind schemes, including the Godunov, Roe, and E–O fluxes, to slowly moving shocks may be relevant here. In particular, it was found that all of the fluxes generated a long wavelength oscillation behind a slowly moving shock (as noted and discussed in [7]) but that the amplitude was very small when using the E–O flux; this is suggestive of the discussion in the preceding paragraph. However, we caution that the explicit shocks of [21] are weak and relatively fast when compared to the results presented here. More significantly, our calculations are implicit (except for an initial transient) in the nonlinear modes, thereby making a direct comparison of the two papers difficult.

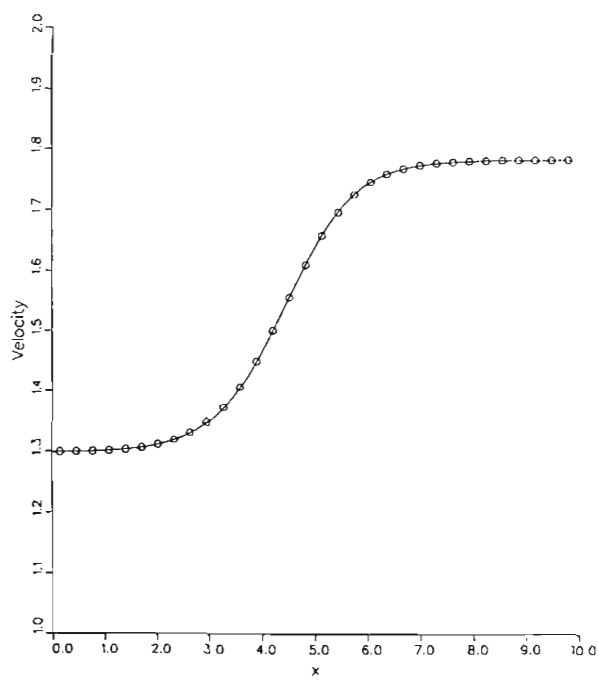

FIG. 5. Velocity vs. distance along the expanding inviscid supersonic–supersonic duct using the inflow conditions of Ref. [23], but changing the exit condition so that the steady flow is smooth; the solid line represents the exact solution and the circles are the converged (steady state) computed solution on a 32-zone mesh.

TABLE I

Convergence Results for the Inviscid Duct

Mesh size	$L1$ error in velocity	Order of convergence
16	0.2376e−01	
32	0.4579e−02	2.375
64	0.1107e−02	2.049
128	0.2731e−03	2.019
256	0.6722e−04	2.023
512	0.1605e−04	2.067
1024	0.3295e−05	2.284

The next calculation was performed on an expanding inviscid duct with supersonic inflow boundary conditions and the exit density specified in such a way that a standing shock is a stable asymptotic solution; the details of the duct shape, boundary conditions, and the solution profiles may be found in [23]. For convenience, the subsonic outflow boundary condition has been implemented numerically by using the known solution at the outflow station as Dirichlet data; the extra information supplied by this condition is spurious because the solution of the Riemann problem at the edge automatically discards it. Initial conditions were linear profiles in density, pressure, and velocity between the inflow and outflow conditions at steady state; this is a very poor initial guess. The numerical parameter α_0 was set equal to 1.0 for this problem in which the transonic region is not smooth. Figure 4a presents a comparison of the converged solution on a 32-zone mesh with the exact solution and Fig. 4b graphs the L^1 residual of the density and the CFL number versus the time step. As can be seen, the solution reaches its asymptotic convergence rate after only about 25 time steps, despite the poor initial guess, and converges to an excellent approximation of the exact steady solution. This behavior illustrates one of the advantages of the present approach, i.e., since the unsteady calculation is done explicitly at the start and is not fully implicit until a steady state is reached, there is less chance of a calculation becoming trapped in a limit cycle or converging to a spurious steady state.

TABLE II

Convergence Results for the Viscous Duct

Mesh size	$L1$ error in velocity	Order of convergence
16	0.2173e+00	
32	0.4821e−01	2.173
64	0.1167e−01	2.046
128	0.2319e−02	2.332
256	0.5122e−03	2.178
512	0.9993e−04	2.358

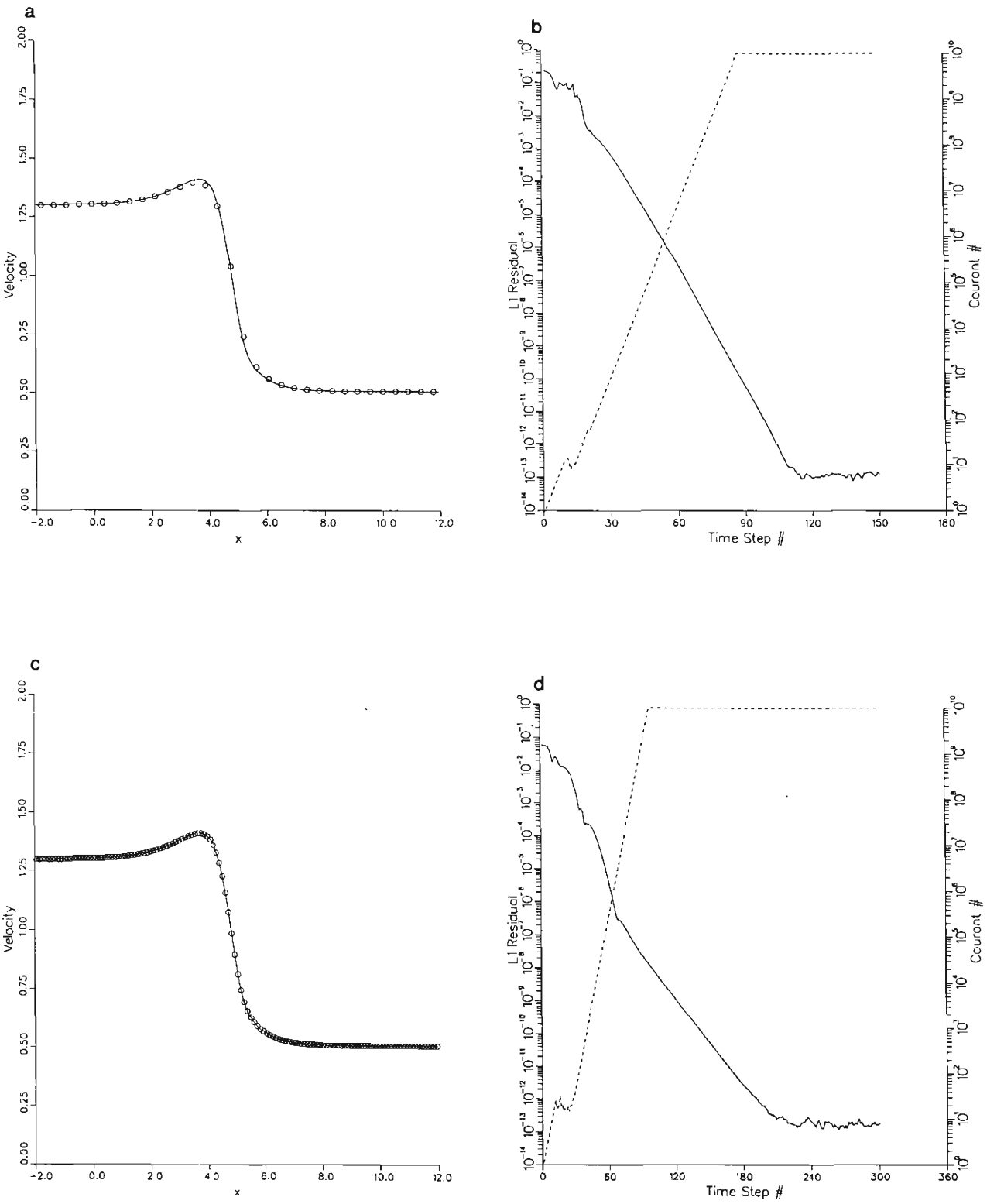

FIG. 6. (a) Velocity vs. distance along the expanding viscous supersonic–subsonic duct with $\varepsilon = 0.1$; the solid line represents the converged (steady state) computed solution using 1024 zones and the circles represent the same for a 32-zone mesh. (b) L^1 residual for density (solid line) and CFL number vs. time step for the 32-zone calculation leading to Fig. 6a (dashed line). (c) Velocity vs. distance along the expanding viscous supersonic–subsonic duct with $\varepsilon = 0.1$; the solid line represents the converged (steady state) computed solution using 1024 zones and the circles represent the same for a 128-zone mesh. (d) L^1 residual for density (solid line) and CFL number vs. time step for the 128-zone calculation leading to Fig. 6c (dashed line).

In order to assess the order of convergence of the inviscid solutions with respect to mesh refinement, we performed a series of calculations using the same duct shape and supersonic inflow boundary conditions as in the above result but changing the exit boundary condition so that the flow is supersonic and smooth throughout. Results of the 32-zone calculation are compared with the exact steady solution (obtained by a quadrature using the inflow conditions as initial data) in Fig. 5. The convergence results are presented in Table I and are based on the L^1 error in velocity; it is clear that the method achieves second-order accurate steady states.

Finally, we discuss our results for a series of viscous duct calculations. The duct shape, inflow boundary conditions, and exit condition are the same as those for the supersonic–subsonic inviscid calculation discussed above, except that the duct has been extended by 20% at each end; this provides a more lengthy section for which the geometric source terms are negligible. The viscosity and heat conductivity are given by $\mu = 1.0$ and $\kappa = 0.4667$. For the results presented, the numerical factor ψ has been set to 1. Calculations using other values for this parameter, such as $\psi = 0.5$, have had trouble reaching steady state, in the sense that the residual could not be driven to machine accuracy. However, the nearly converged solution is indistinguishable from the correct steady state and the scheme performs well in every other respect. No artificial viscosity is used in this series. The L^1 errors are taken with respect to the 1024-zone calculation, taking the place of the exact solution. Figures 6a and 6c compare the 32-zone and 128-zone calculations, respectively, with the "exact" solution; the agreement is excellent for the 32-zone calculation and is essentially perfect for the 128-zone calculation. Figures 6b and 6d graph the L^1 residual of the density and the CFL number versus the time step for the 32-zone and 128-zone calculations, respectively. As in the inviscid case, these results show that the asymptotic convergence rate is reached very quickly. The results of the convergence study are presented in Table II; as before, the attainment of second-order steady states is apparent.

6. CONCLUSIONS

The computational results of the preceding section demonstrate that the implicit–explicit methodology constructed in Sections 3 and 4 satisfy the design criteria proposed in Section 1 for the case of compressible flow, whether inviscid or viscous. In particular, it has been shown that explicit high resolution upwind schemes can be smoothly hybridized on a mode-by-mode basis with an implicit method so that problems involving localized wave speed stiffness and convergence to steady state can be efficiently solved without sacrificing accuracy and resolution for strong wave interactions and energy containing modes. Several of the calculations illustrate the importance of replacing the Godunov flux by the smoother approximate Engquist–Osher flux in the present context. Questions involving the optimal choice of an approximate Riemann problem solver and appropriate dissipative mechanisms have not been completely resolved; however, a very simple and efficient solver augmented by a small amount of artificial viscosity and some additional dissipation at sonic points has been capable of handling our test problems.

We expect that there are many possible applications of the approach taken here. In some of them (e.g., nearly incompressible flow) an obvious competing approach is to resolve the stiffness problem at the level of the governing equations by deriving a new system that is asymptotically valid in the appropriate limit which is not stiff, and developing numerical methods for the limit system (e.g., the incompressible Euler and Navier–Stokes equations). However, it should be noted that it is not known whether or not this approach is valid in all situations; for example, in the case of MHD in the limit as the Alfvèn number approaches zero it does not appear to be possible to derive a limiting set of equations; see [13]. Also, it is usually more natural to derive well-posed initial-boundary value problems and to design numerical boundary conditions for the original set of equations than for the reduced set in the appropriate approximation. Another class of problems for which our approach is an obvious candidate arises in atmospheric flows. Here, in addition to the Mach number varying over a wide range, low amplitude gravity waves are present and restrict the time step considerably; eliminating the sound waves through the implementation of reduced equation sets can lead to numerical difficulties in specifying open boundary conditions.

Using the ideas presented in this paper, we have constructed [9, 10] an unsplit version of the implicit–explicit method for low Mach number multidimensional compressible flow, based on the explicit unsplit algorithm introduced in [5]. In future work, we will study variable Mach number problems such as shock wave–boundary layer interactions (preliminary calculations for this problem using the viscous version of the 1D method and directional operator splitting are already available [8]) using a suitable extension of this work.

ACKNOWLEDGMENT

The authors thank the referees for their interesting comments and corrections.

REFERENCES

1. S. Abarbanel, P. Duth, and D. Gottlieb, *Comput. & Fluids* **17**, 1 (1989).
2. J. B. Bell, P. Colella, and H. M. Glaz, *J. Comput. Phys.* **85**, 257 (1989).

3. J. B. Bell, P. Colella, and J. A. Trangenstein, *J. Comput. Phys.* **82**, 362 (1989).
4. J. U. Brackbill, "Numerical Magnetohydrodynamics for High-Beta Plasmas," in *Methods in Computational Physics*, Vol. 16, edited by J. Killeen (Academic Press, New York, 1976).
5. P. Colella, *J. Comput. Phys.* **87**, 171 (1990).
6. P. Colella and H. M. Glaz, *J. Comput. Phys.* **59**, 264 (1985).
7. P. Colella and P. R. Woodward, *J. Comput. Phys.* **54**, 174 (1984).
8. J. P. Collins, NSWC TR 88-320, Naval Surface Warfare Center, Silver Spring, MD, 1989 (unpublished).
9. J. P. Collins, Ph.D. dissertation, University of Maryland, 1992 (unpublished).
10. J. P. Collins, P. Colella, and H. M. Glaz, "An Unsplit Implicit–Explicit Godunov Method for Compressible Gas Dynamics," in *Computational Fluid Dynamics '92*, Vol. 2, edited by Ch. Hirsch et al. (Elsevier Science Publishers B.V., Amsterdam, 1992).
11. B. Engquist and S. Osher, *Math. Comput.* **36**, 321 (1981).
12. B. A. Fryxell, P. R. Woodward, P. Colella, and K. H. Winkler, *J. Comput. Phys.* **63**, 283 (1986).
13. S. Klainerman and A. Majda, *Commun. Pure Appl. Math.* **34**, 481 (1981).
14. P. D. Lax, *Commun. Pure Appl. Math.* **10**, 537 (1957).
15. T.-P. Liu, *Commun. Math. Phys.* **68**, 141 (1979).
16. T.-P. Liu, *Arch. Rat. Mech. Anal.* **80**, 1 (1982).
17. S. Osher and F. Solomon, *Math. Comput.* **38**, 339 (1982).
18. G. Patnaik, R. H. Guirguis, J. P. Boris, and E. S. Oran, *J. Comput. Phys.* **71**, 1 (1987).
19. R. D. Richtmyer and K. W. Morton, *Difference Methods for Initial-Value Problems* (Wiley, New York, 1957).
20. K. V. Roberts and D. E. Potter, "Magnetohydrodynamic Calculations," in *Methods in Computational Physics*, Vol. 9, edited by B. Alder, S. Fernbach, and M. Rotenberg (Academic Press, New York, 1970).
21. T. W. Roberts, *J. Comput. Phys.* **90**, 141, 1990.
22. P. L. Roe, *J. Comput. Phys.* **43**, 357 (1981).
23. G. R. Shubin, A. B. Stephens, and H. M. Glaz, *J. Comput. Phys.* **39**, 364 (1981).
24. B. van Leer, *SIAM J. Sci. Stat. Comput.* **5**, 1 (1984).