A Class of Bidiagonal Schemes for Solving the Euler Equations

F. Casier,* H. Deconinck,* and Ch. Hirsch†
Vrije Universiteit Brussel, Brussels, Belgium

A class of implicit bidiagonal schemes is developed for the conservative Euler equations. The schemes are based on a predictor-corrector approach for the characteristic variables, adapted according to the sign of the characteristic speeds of information. They are second-order accurate in space and time and have intrinsic dissipative properties. A new fully implicit boundary treatment is developed for this class of schemes, allowing high CFL numbers in all regions of the flow domain. Results are presented for the quasi-one-dimensional flow in a nozzle.

Introduction

In recent years, substantial progress has been made in the development of numerical methods for solving conservative systems of hyperbolic equations such as the Euler equations describing inviscid fluid flow.[1-4]

One of the main achievements has been the development of implicit methods based on tridiagonal solution procedures.[2] The use of explicit-implicit artificial viscosity[5] and of implicit boundary conditions[6,7] has made these algorithms very robust and efficient. Recently, an implicit method based on a bidiagonal solution technique has been introduced by MacCormack.[1]

The first part of this paper is concerned with further developments related to the bidiagonal solution technique. A general class of second-order accurate bidiagonal schemes is proposed, containing a class of implicit schemes and one of semiexplicit schemes. It is shown that these schemes are intrinsically dissipative so that no artificial viscosity is necessary for linear stability.

The second part of the paper introduces an implicit boundary treatment for these bidiagonal schemes that preserves the stability properties of the interior scheme. Indeed, while a number of successful boundary procedures have been proposed for explicit and tridiagonal implicit schemes, this stability remained an unsolved problem in earlier applications of bidiagonal techniques.[8,9]

Numerical experiments for the Euler flow in a nozzle confirm the unconditional linear stability, allowing the use of very high CFL numbers.

Development of the Class of Schemes

Consider a scalar hyperbolic equation in conservation law form

$$\frac{\partial u}{\partial t}+\frac{\partial f}{\partial x}=0 \tag{1a}$$

or in quasilinear form (with Jacobian $a=\partial f/\partial u$)

$$\frac{\partial u}{\partial t}+a\frac{\partial u}{\partial x}=0 \tag{1b}$$

First, a generalized "forward step" discretization will be considered. From an analysis of the domain of dependence of the solution procedure, it will be seen that a second type of discretization (a "backward step") has to be introduced for stability. Combining these discretizations according to the signs of the characteristic speeds will finally lead to the class of central bidiagonal schemes.

Forward Step

A conservative two-parameter family of finite difference discretizations of Eq. (1a) is determined by

$$\left[\frac{\partial u}{\partial t}+\frac{\partial f}{\partial x}\right]^{t=n+\theta}_{x=i+\frac{1}{2}+\xi}=0 \tag{2}$$

The discretization is carried out with the help of the mesh points i and $i+1$ at time levels n and $n+1$ (compact differencing). The temporal discretization of Eq. (2) is determined by

$$\left[\frac{\partial u}{\partial t}\right]^{n+\theta}=\frac{\Delta u^n}{\Delta t}+\left(\theta-\frac{1}{2}\right)\Delta t\frac{\partial^2 u}{\partial t^2}+0(\Delta t^2) \tag{3}$$

$$\left[\frac{\partial f}{\partial x}\right]^{n+\theta}=\frac{\partial}{\partial x}\left(f^n+\theta a^n\Delta u^n\right)+0(\Delta t^2) \tag{4}$$

which leads to

$$\left\{1+\theta\Delta t\frac{\partial}{\partial x}a^n\right\}\Delta u^n=-\Delta t\frac{\partial f^n}{\partial x}+\left(\frac{1}{2}-\theta\right)\Delta t^2\frac{\partial^2 u}{\partial t^2}+0(\Delta t^3) \tag{5}$$

where $\Delta u^n=u^{n+1}-u^n$.

For the spatial discretization, one has

$$\left[\frac{\partial u}{\partial t}\right]_{i+\frac{1}{2}+\xi}=\left(\frac{1}{2}+\xi\right)\frac{\partial u}{\partial t}\bigg|_{i+1}+\left(\frac{1}{2}-\xi\right)\frac{\partial u}{\partial t}\bigg|_{i}+0(\Delta x^2) \tag{6}$$

$$\left[\frac{\partial f}{\partial x}\right]_{i+\frac{1}{2}+\xi}=\frac{1}{\Delta x}\left(f_{i+1}-f_i\right)+\xi\Delta x\frac{\partial^2 f}{\partial x^2}+0(\Delta x^2) \tag{7}$$

which leads to

$$(\mu+\xi\delta)\left(\frac{\partial u}{\partial t}\right)_{i+\frac{1}{2}}+\frac{1}{\Delta x}\delta f_{i+\frac{1}{2}}=-\xi\Delta x\frac{\partial^2 f}{\partial x^2}+0(\Delta x^2) \tag{8}$$

where

$$\mu[\ ]_{i+\frac{1}{2}}=\tfrac{1}{2}\{[\ ]_{i+1}+[\ ]_i\}$$

$$\delta[\ ]_{i+\frac{1}{2}}=[\ ]_{i+1}-[\ ]_i$$

Presented as Paper 83-0126 at the AIAA 21st Aerospace Sciences Meeting, Reno, Nev., Jan. 10-13, 1983; submitted Feb. 9, 1983; revision received Dec. 27, 1983.

*Research Assistant, Department of Fluid Mechanics.
†Professor, Department of Fluid Mechanics. Member AIAA.

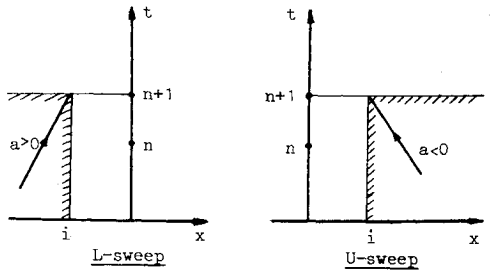


Fig. 1 Domain of dependence for forward-marching discretization.

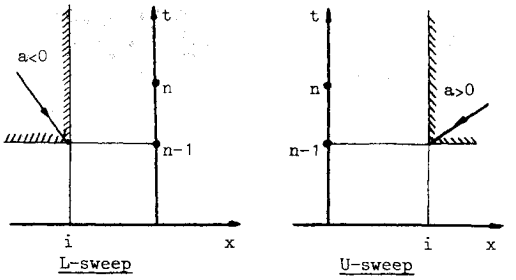


Fig. 2 Domain of dependence for backward-marching discretization.

Combining the spatial and temporal discretizations yields the following bidiagonal finite difference equations for each mesh cell $(i, i+1)$:

$$\{(\mu+\xi\delta)+\theta\sigma\delta a^n_{i+\frac{1}{2}}\}\Delta u^n_{i+\frac{1}{2}} = -\sigma\delta f^n_{i+\frac{1}{2}} \tag{9}$$

where $\sigma = \Delta t/\Delta x$.

This discretization is first-order accurate, except for the special choice $\theta = \frac{1}{2}$ (second-order temporal accuracy) or $\xi = 0$ (second-order spatial accuracy). The correction Δu^n obtained from Eq. (9) is determined by marching forward in time from level n to level $n+1$. For this reason, Eq. (9) is called a "forward step" (F step).

Solution Procedure: L and U Sweeps

Consider the compact forward step [Eq. (9)]. The application of this equation in each mesh cell leads to a set of equations with a bidiagonal coefficient matrix. One additional condition is necessary to close the system. If this condition is imposed at the left side of the domain, a lower bidiagonal coefficient matrix is obtained and the solution can be computed by a sweep from left to right ("L sweep").

If the condition is imposed at the right side, an upper bidiagonal coefficient matrix results and the solution can be computed by a right-to-left sweep ("U sweep").

However, the choice of the sweep type (L and U) is not arbitrary, but rather depends on the domain of dependence analysis. Indeed, it is well known that the characteristics of the hyperbolic differential equation must be included in the numerical domain of dependence of the difference scheme. This domain is the left side of the discretization point for an L sweep and the right side for a U sweep (Fig. 1). As a result, when solved with an L sweep Eq. (9) will be stable for only positive signal speeds a. Similarly, the solution with a U sweep is stable for only negative a. Since the scheme will have to deal with mixed or variable sign signal speeds, a two-step procedure is to be considered, with an L sweep in the first step and a U sweep in the second (or vice versa). To stabilize the first step (L sweep) for a negative signal speed and the second step (U sweep) for a positive signal speed, a modification is necessary.

Backward Step

The modification consists of using a backward time integration for the unstable characteristic, such that the domain of dependence requirement is again satisfied (Fig. 2). The backward discretization is determined by the following discretization point used for integration from level n to $n-1$:

$$\left[\frac{\partial u}{\partial t}+\frac{\partial f}{\partial x}\right]^{n-\theta}_{i+\frac{1}{2}+\xi} = 0$$

With the temporal discretization

$$\left.\frac{\partial u}{\partial t}\right|^{n-\theta} = \frac{\Delta u^{n-1}}{\Delta t} + \left(\frac{1}{2}-\theta\right)\Delta t\frac{\partial^2 u}{\partial t^2} + 0(\Delta t^2) \tag{10}$$

and using the spatial discretization [Eq. (8)], the following bidiagonal equation is obtained, in which the solution at level n is assumed to be known (backward or B step):

$$\left.\frac{\partial f}{\partial x}\right|^{n-\theta} = \frac{\partial}{\partial x}(f^n - \theta a^n \Delta u^{n-1}) + 0(\Delta t^2) \tag{11}$$

$$\{(\mu+\xi\delta)-\theta\sigma\delta a^n_{i+\frac{1}{2}}\}\Delta u^{n-1}_{i+\frac{1}{2}} = -\sigma\delta f^n_{i+\frac{1}{2}} \tag{12}$$

The Two-Step L-U Scheme

The class of two-step central bidiagonal schemes that satisfy the domain of dependence requirements for arbitrary sign of the signal speed is then formulated as follows.

First or predictor step, L sweep:

$$L\begin{cases}\{(\mu+\xi\delta)+\theta\sigma\delta|a^n_{i+\frac{1}{2}}|\}\overline{\Delta u}_{i+\frac{1}{2}} = -\sigma\delta f^n_{i+\frac{1}{2}} \\ u^{\overline{n+1}} = u^n + \overline{\Delta u}\end{cases} \tag{13}$$

Second or corrector step, U sweep:

$$U\begin{cases}\{(\mu-\xi\delta)-\theta\sigma\delta|a^{\overline{n+1}}_{i+\frac{1}{2}}|\}\overline{\overline{\Delta u}}_{i+\frac{1}{2}} = -\sigma\delta f^{\overline{n+1}}_{i+\frac{1}{2}} & \text{(14a)} \\ u^{n+1} = u^n + \frac{1}{2}(\overline{\Delta u}+\overline{\overline{\Delta u}}) = u^n + \Delta u^n & \text{(14b)}\end{cases}$$

The two-step U-L scheme is obtained in a similar way by replacing θ with $-\theta$ and ξ with $-\xi$.

By taking the absolute value of a in the first step, Eq. (13) is a forward time integration for positive a [reducing to Eq. (9)] and a backward time integration for negative a [reducing to Eq. (12)]. Thus, for negative a, the predictor solution is obtained by a linear extrapolation from levels n and $n-1$,

$$u^{\overline{n+1}} = u^n + \overline{\Delta u}^{n-1} \approx 2u^n - u^{n-1}$$

For the corrector step [Eqs. (14)], on the other hand, the absolute value amounts to a backward step from level $\overline{n+1}$ to $\bar{n}$ for positive signal speed and to a forward step from level $\overline{n+1}$ to $\overline{\overline{n+2}}$ if a is negative. Hence, the updating after two steps [Eq. (14b)] is in fact given by

$$u^{n+1} = u^n + \tfrac{1}{2}(\overline{\Delta u}^n + \overline{\overline{\Delta u}}^n) \qquad \text{for } a>0 \tag{15a}$$

$$u^{n+1} = u^n + \tfrac{1}{2}(\overline{\Delta u}^{n-1} + \overline{\overline{\Delta u}}^{n+1}) \qquad \text{for } a<0 \tag{15b}$$

This interpretation of the time level of the corrections is important for the specification of the consistent implicit boundary conditions discussed in the next section.

The scheme of Eqs. (13) and (14) is second-order accurate (or central) in both space and time. To achieve the second-order space accuracy, the discretization has been chosen at $i+\frac{1}{2}-\xi$ in the corrector step. Indeed, the first-order truncation error terms to be added to Eqs. (13) and (14), respectively, are then given by

$$\Delta t\left\{(\tfrac{1}{2}-\theta)\Delta t\frac{\partial^2 u}{\partial t^2} - \xi\Delta x\frac{\partial^2 f}{\partial x^2} + \Delta t 0(\Delta t^2, \Delta x^2)\right\} \tag{16}$$

$$\Delta t\left\{(\theta-\tfrac{1}{2})\Delta t\frac{\partial^2 u}{\partial t^2}+\xi\Delta x\frac{\partial^2 f}{\partial x^2}+\Delta t0(\Delta t^2,\Delta x^2)\right\} \tag{17}$$

such that the first-order truncation errors cancel after determining the global correction explicitly from Eqs. (13) and (14).

The stability of the finite difference system of Eqs. (13) and (14) can be studied with a Von Neumann analysis. The amplification factor is given by

$$G=1-\frac{j\sigma\ at\ +\tfrac{1}{2}\sigma^2a^2t^2}{1+(\xi+\theta\sigma|a|)^2t^2} \tag{18}$$

where $t=2\tan(k\Delta x/2)$ with k the wave number.

The system is linearly stable if the amplitude $|G|$ is less than unity, assuming ξ and θ are positive,

$$\zeta\equiv\frac{\xi}{\sigma|a|}+\theta\geq\frac{1}{2} \tag{19}$$

For $\zeta=\tfrac{1}{2}$, a nondissipative scheme is obtained. For larger values of ζ, intrinsically dissipative schemes are obtained. The higher frequencies are especially affected by the choice of ζ. Complete dissipation of the $2\Delta x$ waves (highest frequency in a mesh) is obtained for the choice $\zeta=\sqrt{2}/2$.

The study also shows that the CFL number (σa) particularly affects the lower frequencies. It follows that the value of ζ can be used to control the amount of dissipation for the higher frequencies, while the value of CFL can be used to control the amount of dissipation for the middle and low frequencies.

Conservative System of Coupled Equations

Consider the conservative system of coupled hyperbolic equations

$$\frac{\partial U}{\partial t}+\frac{\partial F}{\partial x}=0 \tag{20}$$

or in quasilinearized form ($A=\partial F/\partial U=$ the Jacobian)

$$\frac{\partial U}{\partial t}+A\frac{\partial U}{\partial x}=0 \tag{21}$$

This linearized system can be decoupled by the transformation to characteristic variables W. This gives

$$\partial W=L\partial U \qquad A=L^{-1}DL \tag{22}$$

where D is the diagonalized Jacobian containing the speeds of the associated characteristics. The L-U scheme of Eqs. (13) and (14) is then applied to each equation of the decoupled system and the finite difference system is transformed back to conservative variables.

This leads to the following class of central (second-order accurate) bidiagonal schemes for a coupled system of hyperbolic equations in conservation form:

$$L\begin{cases}\{(\mu+\xi\delta)I+\theta\sigma\delta|A|^n_{i+\frac{1}{2}}\}\overline{\Delta U}_{i+\frac{1}{2}}=-\sigma\delta F^n_{i+\frac{1}{2}}\\ U^{n+1}=U^n+\overline{\Delta U}\end{cases}\quad\text{(predictor)} \tag{23a}$$

$$U\begin{cases}\{(\mu-\xi\delta)I-\theta\sigma\delta|A|^{\overline{n+1}}_{i+\frac{1}{2}}\}\overline{\overline{\Delta U}}_{i+\frac{1}{2}}=-\sigma\delta\overline{F}^{n+1}_{i+\frac{1}{2}}\\ U^{n+1}=U^n+\tfrac{1}{2}(\overline{\Delta U}+\overline{\overline{\Delta U}})\end{cases}\quad\text{(corrector)} \tag{23b}$$

where I is the identity matrix (δ_{ij}) and

$$|A|=L^{-1}|D|L, \qquad |D_{ij}|=|a_i|\delta_{ij} \tag{24}$$

The schemes are stable if, for all characteristic speeds a_i,

$$\zeta_i=\left(\theta+\frac{\xi}{\sigma|a_i|}\right)\geq\frac{1}{2} \tag{25}$$

independently of the sign of the characteristics.

Two special subclasses are:

1) $\xi=0$ implicit schemes. These schemes are unconditionally stable under the condition $\theta\geq\tfrac{1}{2}$. For $\theta=\tfrac{1}{2}$ a nondissipative scheme is obtained and for $\theta=\sqrt{2}/2$ the scheme completely damps the propagation of the $2\Delta x$ waves (the highest frequencies in the discretized solution). As nonlinearities usually generate high-frequency oscillations, the choice of $\theta=\sqrt{2}/2$ will be very interesting in these regions. For the special choice of $\theta=1$, one obtains a scheme which is closely related to the bidiagonal implicit scheme of MacCormack,[1] except for the boundary treatment.

The parameter λ introduced in the MacCormack scheme [see Ref. 1, Eq. (5)] is related to θ in the following way:

$$\theta=\frac{1}{|a|}\left(\lambda+\frac{1}{2\sigma}\right) \tag{26}$$

Further in the paper MacCormack [Ref. 1, Eq. (8)] uses the choice

$$\lambda=\max\left\{|a|-\frac{1}{2\sigma},\ 0\right\} \tag{27}$$

which is equivalent to

$$\begin{cases}\theta=1 & \text{if } |a|>\dfrac{1}{2\sigma}\\[2ex] \theta=\dfrac{1}{2\sigma|a|} & \text{if } |a|<\dfrac{1}{2\sigma}\end{cases} \tag{28}$$

2) $\theta=0$ semiexplicit schemes. These schemes have the form

$$L\begin{cases}(\tfrac{1}{2}+\xi)\overline{\Delta U}_{i+1}=-\sigma\delta F^n_{i+\frac{1}{2}}-(\tfrac{1}{2}-\xi)\overline{\Delta U}_i\\ U^{n+1}=U^n+\overline{\Delta U}\end{cases}\quad\text{(predictor)}$$

$$U\begin{cases}(\tfrac{1}{2}+\xi)\overline{\overline{\Delta U}}_i=-\sigma\delta\overline{F}^{n+1}_{i+\frac{1}{2}}-(\tfrac{1}{2}-\xi)\overline{\overline{\Delta U}}_{i+1}\\ U^{n+1}=U^n+\tfrac{1}{2}(\overline{\Delta U}+\overline{\overline{\Delta U}})\end{cases}\quad\text{(corrector)} \tag{29}$$

and are conditionally stable under the restriction

$$|\mathrm{CFL}|\leq 2\xi \tag{30}$$

For the special choice $\xi=\tfrac{1}{2}$, the well-known explicit scheme of MacCormack is obtained. For choices of $\xi>\tfrac{1}{2}$, larger CFL numbers can be used at the computational cost of the explicit schemes.

Boundary Treatment

As is shown in the above subsection on the solution procedure, a starting value of the boundary correction $\overline{\Delta u}_0$ ($\overline{\overline{\Delta u}}_I$) is necessary to start the L sweep (U sweep) solution procedure. The number and nature of these boundary conditions can be determined by applying the theory of characteristics: the information entering the computational

domain should be replaced by prescribed natural boundary conditions, while the information leaving the domain should be predicted from inside the domain (numerical boundary conditions).

In particular, the treatment of the numerical boundary conditions has been a serious obstacle in achieving unconditional stability in earlier applications of bidiagonal techniques[8,9]: the extrapolation techniques (which have been used with success in explicit and tridiagonal implicit algorithms) led to serious stability problems [CFL restriction to $O(1)$] when applied in bidiagonal implicit algorithms.

In this section, a new time-accurate boundary treatment for the numerical boundary conditions is proposed, based on the time level interpretation of the correction resulting from the backward or forward time integration. This boundary treatment uses no space extrapolation and preserves the linear stability properties of the interior scheme, as will be shown in the numerical experiments.

Boundary Conditions for the L-U Scheme for $a>0$

For characteristics with positive signal speeds, natural boundary conditions have to be imposed to start the L sweep (point 0) and numerical conditions have to be imposed to start the U sweep (point I). Consider the numerical conditions at the linking boundary (I). The corrections $\overline{\Delta u}$ are both consistent approximations of the correction Δu^n [cf. Eq. (15a)].

$$\begin{aligned}\overline{\Delta u} &= \Delta u^n + \Delta t 0(\Delta t)\\ \overline{\overline{\Delta u}} &= \Delta u^n + \Delta t 0(\Delta t)\end{aligned} \tag{31}$$

In the best case, both corrections are equal in all points and also at the boundary. $\overline{\Delta u}_I$ is determined in the previous sweep, so that $\overline{\overline{\Delta u}}_I$ can be predicted for the numerical boundary conditions by

$$\overline{\overline{\Delta u}}_I^n = \overline{\Delta u}_I^n \tag{32}$$

As a consequence, the second-order accurate solution will be determined by

$$u_I^{n+1} = u_I^n + \tfrac{1}{2}(\overline{\Delta u}_I^n + \overline{\overline{\Delta u}}_I^n) = u_I^n + \overline{\Delta u}_I^n = u_I^{n+1} \tag{33}$$

The natural boundary condition can now be imposed in a consistent way by requiring

$$u_0^{n+1} = u_0^* \quad \text{or} \quad \overline{\Delta u}_0^n = u_0^* - u_0^n \tag{34}$$

where u_0^* is the natural boundary condition.

Boundary Conditions for the L-U Scheme for $a<0$

For negative characteristics, numerical conditions have to be imposed to start the L sweep (point 0) and natural conditions to start the U sweep (point I). Consider first the linking boundary. Here, the second-order accurate solution can be determined by the natural conditions

$$u_I^{n+1} = u_I^* \tag{35}$$

where u_I^* is the natural boundary condition.

As the value $\overline{\Delta u}_I$ has already been determined in the previous sweep, the starting correction at the linking boundary is simply obtained from Eq. (15b),

$$\overline{\overline{\Delta u}}_I^{n+1} = 2(u_I^* - u_I^n) - \overline{\Delta u}_I^{n-1} \tag{36}$$

Several possibilities are now open for introducing numerical boundary conditions in point 0 consistent with the second-order accurate solution. For example, one can consider the solution obtained in the previous iterations as the exact one,

$$\overline{\Delta u}_0^{n-1} = 0 \tag{37}$$

or one can try to find more time-accurate predictions by considering results of previous steps,

$$\overline{\Delta u}_0 = \tfrac{1}{2}(\overline{\Delta u}_0 + \overline{\overline{\Delta u}}_0)^{\text{previous iteration}} \tag{38}$$

Comparing this to Eq. (15b) shows that this is equivalent to

$$\overline{\Delta u}_0^{n-1} = \tfrac{1}{2}(\overline{\Delta u}_0^{n-2} + \overline{\overline{\Delta u}}_0^n) = \Delta u_0^{n-1} \tag{39}$$

Another valid choice is

$$\overline{\Delta u}_0 = (\overline{\overline{\Delta u}}_0)^{\text{previous iteration}} \tag{40}$$

which is equivalent to

$$\overline{\Delta u}_0^{n-1} = \overline{\overline{\Delta u}}_0^n$$

Boundary Conditions for a Conservative System of Coupled Hyperbolic Equations

Consider the conservative system of coupled equations (20). The locally linearized characteristics are defined by Eq. (22) as

$$\partial W = L \partial U$$

The natural boundary conditions are imposed as "primitive" variables (for instance, pressure, density, and velocity), defined by

$$\partial \tilde{U} = T \partial W \tag{41}$$

Both transformations can be linearized, which yields

$$\Delta W = L \Delta U \text{ and } \Delta \tilde{U} = T \Delta W \tag{42}$$

The characteristic corrections can now be divided into those associated with the characteristics entering domain (ΔW_{in}) and leaving domain (ΔW_{out}). In the same way, the primitive corrections can be divided into natural ($\Delta \tilde{U}^I$) and numerical ($\Delta \tilde{U}^{II}$) primitive corrections. This leads to

$$\begin{bmatrix}\Delta W_{\text{in}}\\ \Delta W_{\text{out}}\end{bmatrix} = \begin{bmatrix}T_{\text{in}}^{-1}\\ T_{\text{out}}^{-1}\end{bmatrix} \Delta \tilde{U} = \begin{bmatrix}T_{\text{in}}^{-II} & T_{\text{in}}^{-III}\\ T_{\text{out}}^{-II} & T_{\text{out}}^{-III}\end{bmatrix} \begin{bmatrix}\Delta \tilde{U}^I\\ \Delta \tilde{U}^{II}\end{bmatrix} \tag{43}$$

and

$$\begin{bmatrix}\Delta W_{\text{in}}\\ \Delta W_{\text{out}}\end{bmatrix} = \begin{bmatrix}L_{\text{in}}\\ L_{\text{out}}\end{bmatrix} \Delta U \tag{44}$$

The natural boundary conditions are determined by $\Delta \tilde{U}^I$ and the numerical boundary conditions by

$$\Delta W_{\text{out}} = L_{\text{out}} \Delta U$$

so that the correction $\Delta \tilde{U}^{II}$ can be found from

$$\Delta \tilde{U}^{II} = [T_{\text{out}}^{-III}]^{-1}(L_{\text{out}} \Delta U - T_{\text{out}}^{-II} \Delta \tilde{U}^I) \tag{45}$$

For the linking boundary, one must insert in Eq. (45)

$$\begin{aligned}\Delta \tilde{U}^I &= 2(\tilde{U}^* - \tilde{U}^n) - \overline{\Delta \tilde{U}} && \text{Eq. (36)}\\ \Delta U &= \overline{\Delta U} && \text{Eq. (32)}\end{aligned} \tag{46}$$

For the nonlinking boundary, one must insert

$$\begin{aligned}\Delta \tilde{U}^I &= \tilde{U}^* - \tilde{U}^n && \text{Eq. (34)}\\ \Delta U &\quad \text{defined by Eq. (37), (38), or (40)}\end{aligned} \tag{47}$$

The starting correction is now given in primitive variables by

$$\Delta\tilde{U}=\begin{bmatrix}\Delta\tilde{U}^{I}\\ \Delta\tilde{U}^{II}\end{bmatrix} \tag{48}$$

so that the starting value ΔU can be easily computed.

Numerical Results

To test the class of schemes and the boundary treatment, a number of numerical experiments have been conducted for the quasi-one-dimensional Euler flow in a Laval nozzle. The quasi-one-dimensional Euler equations are given by

$$S(x)\frac{\partial U}{\partial t}+\frac{\partial F}{\partial x}+H=0 \tag{49}$$

where

$$U=\begin{bmatrix}\rho\\ \rho u\\ e\end{bmatrix}\quad F=S(x)\begin{bmatrix}\rho u\\ p+\rho u^2\\ (e+p)u\end{bmatrix}\quad H=\frac{\partial S}{\partial x}\begin{bmatrix}0\\ -p\\ 0\end{bmatrix} \tag{50}$$

with $p=(\gamma-1)(e-0.5\rho u^2)$ and the area ratio $S(x)$ is defined by[7]

$$S(x)=1+1.5\left(1-\frac{x}{5}\right)^2 \qquad 0\le x\le 5$$

$$S(x)=1+0.5\left(1-\frac{x}{5}\right)^2 \qquad 5\le x\le 10$$

For this steady test case, local time stepping (constant CFL number throughout the flow) was used

$$\sigma_i=\mathrm{CFL}/[u+c]_i$$

where c_i is equal to the local speed of sound and $[u+c]_i$ to the local maximum eigenvalue of the system of equations.

The schemes considered in the numerical experiments are the following *U-L* schemes:

1) A subclass of implicit schemes with $\xi=0$ and $\theta>\frac{1}{2}$.

2) A subclass of semiexplicit schemes with $\theta=0$ and $\xi>\frac{1}{2}$ CFL.

In order to exclude all effects other than those caused by the basic scheme and the boundary treatment, no additional terms (artificial damping) have been introduced.

To test the accuracy of the steady-state results, a mass flux error is introduced, as

$$\mathrm{ER}=\frac{(\rho uS)_{\mathrm{num}}-(\rho uS)_{\mathrm{exact}}}{(\rho uS)_{\mathrm{exact}}}\cdot 100\%$$

For all test cases, a maximum error (ERM) and a mean value of the mass flux error (ERG) are computed.

To test the convergence of the algorithms, a mean residual (RMS) and a maximum residual (RSM) are defined as

$$\mathrm{RMS}=\sqrt{\sum_i(\rho_i^{n+1}-\rho_i^n)^2\Big/\sum_i(\rho_i^1-\rho_i^0)^2}$$

$$\mathrm{RSM}=\frac{\max_i|\rho_i^{n+1}-\rho_i^n|}{\max_i|\rho_i^1-\rho_i^0|}$$

All experiments were conducted up to machine accuracy ($\mathrm{RMS}\approx 10^{-12}$) and the convergence criterium was defined by $\mathrm{RMS}\le 10^{-10}$.

The natural boundary conditions are obtained from the exact solution, which is determined by the total temperature (300 K), the total pressure (10^5 Pa), the critical area S^* at the inlet, and the shock position x_s. The initial condition is a linear interpolation of ρ, u, and p between the exact boundary values. The test cases are described in Table 1.

Cases I, II, and IV were designed to test the boundary treatment and cases IV and V to study the influence of the sonic point. Cases V and VI were selected to study the influence of a discontinuity, depending on its position in the mesh: the shock was positioned on a mesh point in case V and midway between two mesh points in case VI. For all experiments, a mesh of 81 points was used.

Table 2 shows results for the implicit subclass ($\xi=0$). Results for $\theta=0.7$ and 1.0 have been retained. The former choice leads to complete dissipation of the shortest waves in the solution. The latter choice was made for its resemblance to the implicit scheme of MacCormack.[1] In this scheme, if no explicit-implicit switching is used, no damping term is included and, when the implicit boundary treatment of the previous section added, one obtains the $\theta=1$ scheme.

Table 1 Test case definiton (Laval nozzle, $P_t=10^5$ Pa, $T_t=300$ K)

Case	Flow type	S^*	x_{shock}	Natural boundary condition: Inlet	Natural boundary condition: Outlet
I	Fully supersonic	0.8	—	ρ,u,p	—
II	Fully subsonic	0.8	—	ρ,p	p
IV	Subsonic-supersonic	1.0	—	ρ,p	—
V	Transonic + shock[a]	1.0	7.0000	ρ,p	p
VI	Transonic + shock[b]	1.0	7.0625	ρ,p	p

[a] On a mesh point. [b] Midway between two mesh points.

Table 2 Numerical results for the implicit subclass ($\xi=0$)

Case	θ	CFL	Iterations[b]	ERG, %	ERM, %	Figs.
I	0.7	10^{12}[a]	10	10^{-12}	2×10^{-12}	3a,b
II	0.7	10^{12}	35	0.025	0.025	4a,b
IV	0.7	20	100	0.020	0.020	5a,b
IV	1.0	20	140	0.034	0.034	5a,b
V	0.7	60	110	0.043	1.0	6a,b
VI	0.7	39	160	0.015	0.016	8a,b

[a] CFL = 100 for first two iterations. [b] Needed to obtain $\mathrm{RMS}=10^{-10}$.

Table 3 Numerical results for the semiexplicit subclass ($\theta = 0$)

Case	ξ	CFL	Iterations[a]	ERG, %	ERM, %	Figs.
I	10^6	1.8×10^6	200	10^{-12}	2×10^{-12}	3c
I	200	360	200	10^{-12}	2×10^{-12}	3c
II	40	70	360	0.025	0.025	4c
II	10^{12}	1.8×10^{12}	320	0.025	0.025	4c
IV	20	36	Fig. 5c	0.016	0.021	5c
V	20	39	Fig. 7c	0.009	1.1	7a,b
VI	20	39		0.005	0.023	8a,b

[a] Needed to obtain RMS = 10^{-10}.

Table 3 shows results for the semiexplicit subclass ($\theta = 0$) for various values of ξ.

Stability

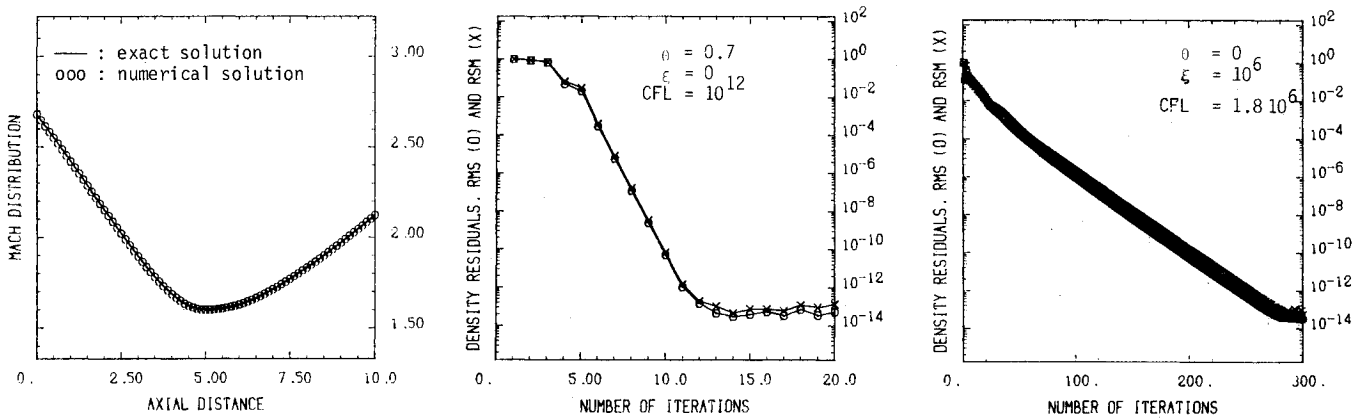
Boundary Treatment

If no strong nonlinearities are present (cases I and II, Figs. 3 and 4), the stability of the interior scheme as determined from Eq. (25) is fully preserved. For the implicit subclass unlimited CFL numbers can be used, and for the semiexplicit subclass the CFL number is restricted only by the choice of ξ [Eq. (30)]. In all cases, the rate of convergence was not slowed by the boundary treatment and machine convergence is reached (Figs. 3b, 3c, 4b, and 4c). This shows that the boundary treatment can be considered as unconditionally stable and fully consistent.

Nonlinearities

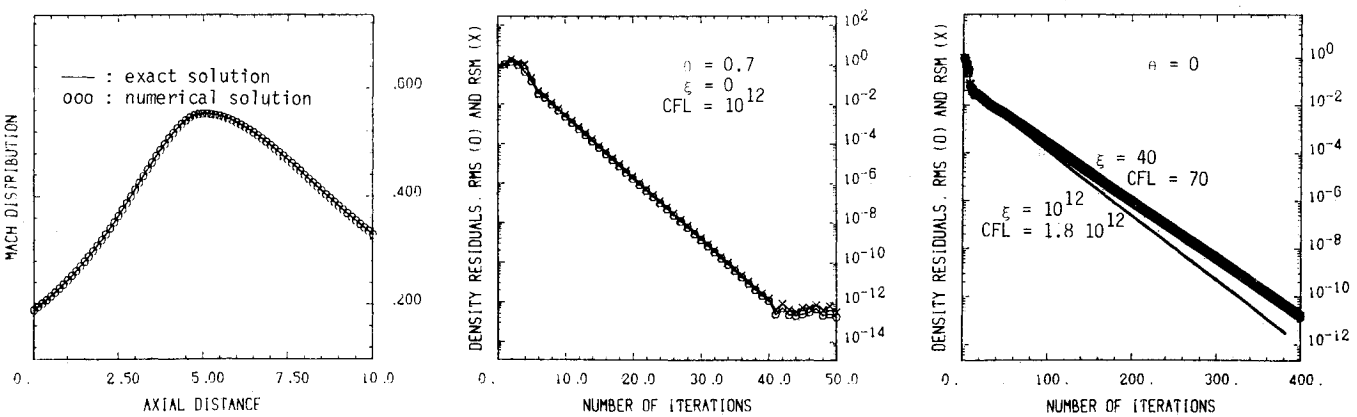
Schemes based on Taylor expansions are inconsistent in the vicinity of discontinuities. So large errors can be generated at shocks and contact surfaces, destabilizing the solution. Indeed, the numerical experiments with the implicit subclass showed a CFL restriction for cases V and VI (CFL $\approx$ 100) with optimal convergence at CFL $\approx$ 60 (Fig. 6). These stability problems can be reduced by the use of shock-handling techniques such as flux splitting and artificial viscosity.

As the entropy condition is not contained in the scheme (no natural or artificial damping terms), unphysical expansion shocks are mathematically acceptable solutions.[10] The nonuniqueness of the solution can therefore lead to instability. This problem was met in the numerical experiments for case IV (Fig. 5) when using the implicit subclass. Here, the CFL number was restricted to $O(50)$ for stability and to $O(25)$ to obtain a unique solution. This uniqueness problem can be removed for example by including a damping term that vanishes as steady state is reached.[10] For the semiexplicit subclass, the nonlinear effects lead to a stagnation of the convergence (rms $\approx 10^{-5}$), indicating a low-scale instability (Figs. 5c and 7b).



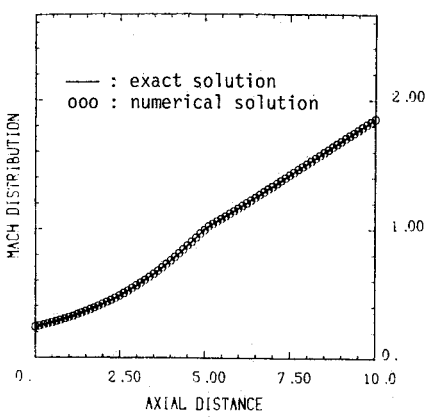
a) Mach number distribution. b) Convergence for implicit subclass. c) Convergence for semiexplicit subclass.

Fig. 3 Case I, fully supersonic flow.

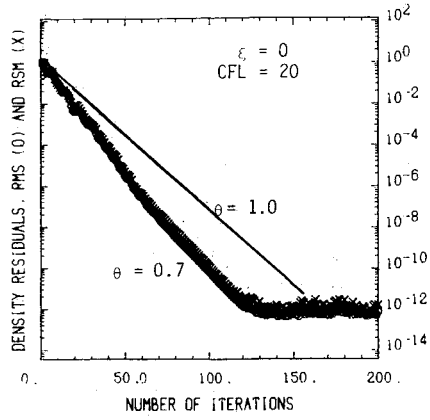


a) Mach number distribution. b) Convergence for implicit subclass. c) Convergence for semiexplicit subclass.

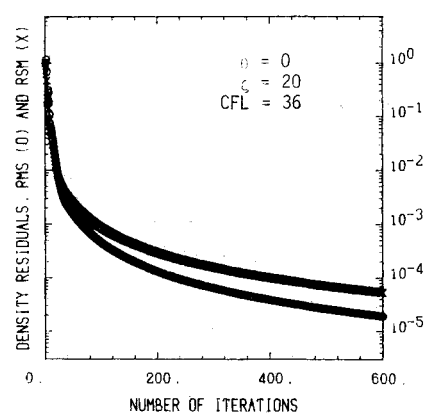
Fig. 4 Case II, fully subsonic flow.



a) Mach number distribution.



b) Convergence for implicit subclass.



c) Convergence for semiexplicit subclass.

Fig. 5 Case IV subsonic-supersonic flow.

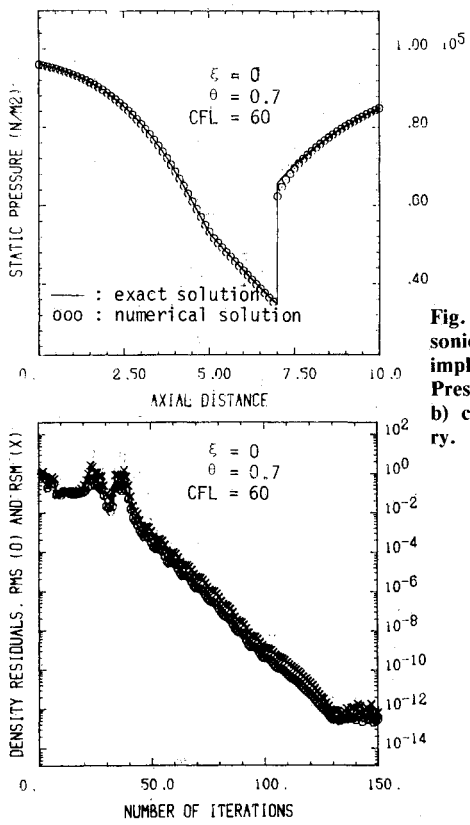


Fig. 6 Case V transonic flow with shock, implicit subclass. a) Pressure distribution; b) convergence history.

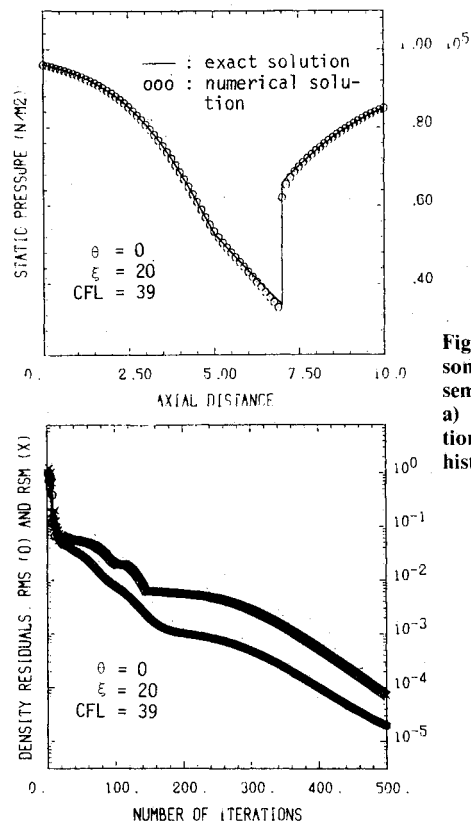


Fig. 7 Case V transonic flow with shock, semiexplicit subclass. a) Pressure distribution; b) convergence history.

Initial Conditions

The large corrections computed in the first few iterations can sometimes cause instability. This problem is easily solved by using a gradual start (a lower CFL number for the first few steps). This was necessary for one of the experiments as shown in Table 2.

Parameter Choice

Except for the effect of nonlinearities, all results for all schemes seem to indicate that the use of higher CFL numbers improves the convergence of the algorithm. For the implicit subclass, the choice of $\theta = \sqrt{2}/2$ leads to higher efficiency than the choice $\theta = 1$ (Fig. 5b). For the semiexplicit subclass, there seems to be an optimal relation between ξ and CFL, which is approximately given by

$$1.5\xi < \mathrm{CFL} < 2.0\xi$$

Accuracy

Boundary Treatment

It has been shown that the second-order accurate solution is determined by the conditions at the linking boundary (inlet for the *U-L* scheme). For flows with subsonic regions, the final solution is found to be subject to exit conditions slightly different from the imposed exit conditions, as can be seen in the error values. In all cases, this difference was less than 0.05%. If all of the conditions can be imposed at the inlet (case I), machine accuracy is achieved.

Sonic Point

Except for the possibility of expansion shocks, no additional errors are introduced in the sonic region for the implicit subclass (ERM = ERG for case IV). For the semiexplicit subclass, the maximum error was slightly larger than the mean error, indicating some error at the sonic point.

Shock

The comparison of the results for cases V and VI shows that the accuracy at the discontinuities is highly dependent upon the shock position with respect to the mesh points. The largest error is introduced when the shock is positioned on a mesh point (case V), while practically no error is introduced when the shock is positioned in the middle between two mesh points (case VI). See Figs. 7a and 8a.

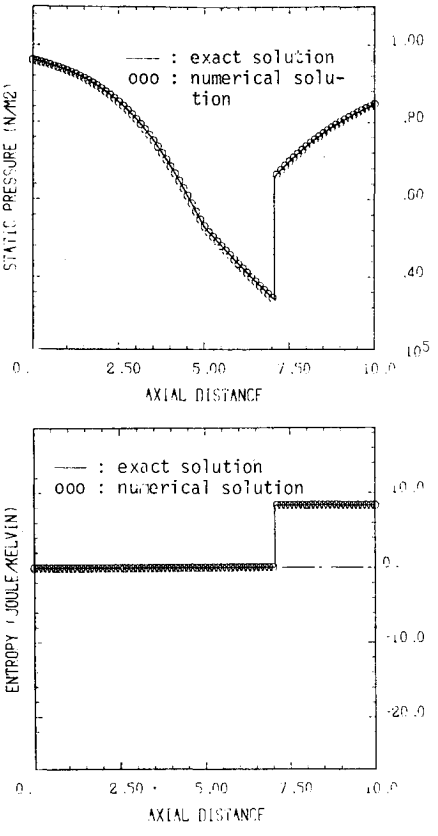


Fig. 8 Case VI transonic flow with shock midway between two mesh points. a) Pressure distribution; b) entropy distribution.

Computational Efficiency

It is obvious that the introduction of techniques stabilizing the strong nonlinear effects (entropy condition, shock-handling techniques, etc.) will strongly improve the efficiency of the schemes described here. For this reason, computational efficiency is discussed only for those cases where no strong nonlinear effects are present (cases I and II).

Without doubt, unconditionally stable schemes with an adequate unconditionally stable boundary treatment (although more complex and three to five times more expensive than the explicit schemes) are extremely efficient. See Figs. 3b and 4b.

If one wishes to sacrifice some efficiency for greater simplicity, the semiexplicit subclass allows a very attractive alternative. These schemes have the same simplicity and the same computational cost as the two-step explicit schemes, but allow the use of very high CFL numbers so that very high rates of convergence are achieved. See Figs. 3c and 4c.

Conclusions

A new class of conservative compact difference schemes has been presented that contains the explicit and implicit MacCormack schemes as special cases.

The schemes are central (second-order accurate) and have dissipative properties that can be controlled by the choice of the discretization parameters (θ and ξ) and by the choice of the time step (CFL number). Two special subclasses have been considered: a fully implicit one that is unconditionally stable for the linear case and a semiexplicit one that is conditionally stable and has the simplicity of explicit schemes.

A fully implicit boundary treatment has been developed that preserves the stability properties of the interior scheme.

Numerical experiments confirm the efficiency and the stability properties of the new method. In particular, very high rates of convergence (CFL of the order of 10^{12}) and high accuracy can be achieved for fully supersonic or fully subsonic flows without artificial viscosity.

Extension and application to two-dimensional problems will be presented in the near future.

References

[1] MacCormack, R. W., "A Numerical Method for Solving the Equations of Compressible Viscous Flow," *AIAA Journal*, Vol. 20, Sept. 1982, pp. 1275-1281.

[2] Beam, R. M. and Warming R. F., "An Implicit Factored Scheme for the Compresible Navier-Stokes Equations," *AIAA Journal*, Vol. 16, April 1978, pp. 393-402.

[3] Van Leer, B., "Flux-Vector Splitting for the Euler Equations," *Proceedings, 8th International Conference in Numerical Methods in Fluid Dynamics, Lecture Notes in Physics*, No. 170, Springer-Verlag, 1982.

[4] Wornom, S. F., "Application of Compact Difference Schemes to the Conservative Euler Equations for One-Dimensional Flows," NASA TM 83262, 1982.

[5] Steger, J. L., "Implicit Finite-Difference Simulation of Flow about Two-Dimensional Geometries," *AIAA Journal*, Vol. 16, July 1978, pp. 679-686.

[6] Chakravarthy, S. R., "Euler Equations—Implicit Schemes and Implicit Boundary Conditions," AIAA Paper 82-0228, 1982.

[7] Yee, H. C., Beam, R. M., and Warming, R. F., "Stable Boundary Approximations for a Class of Implicit Schemes for the One-Dimensional Inviscid Equations of Gas Dynamics," *AIAA Journal*, Vol. 20, Sept. 1982, pp. 1203-1211.

[8] Kordulla, W. and MacCormack, R. W., "Transonic Flow Computation Using an Explicit-Implicit Method," *Proceedings, 8th International Conference on Numerical Methods in Fluid Dynamics, Lecture Notes in Physics*, No. 170, Springer Verlag, 1981.

[9] Von Lavante, E. and Thompkins, W. T., "An Implicit Bidiagonal Numerical Method for Solving the Navier-Stokes Equations," AIAA Paper 82-0063, 1982.

[10] Lax, P. D., "Hyperbolic Systems of Conservation Laws and the Mathematical Theory of Shock Waves," *SIAM Regional Conference Series in Applied Mathematics*, Vol. II, 1973.