JOURNAL OF COMPUTATIONAL PHYSICS 39, 364–374 (1981)

Steady Shock Tracking and Newton's Method Applied to One-Dimensional Duct Flow

G. R. SHUBIN, A. B. STEPHENS, AND H. M. GLAZ

Applied Mathematics Branch (R44),
Naval Surface Weapons Center, Silver Spring, Maryland 20910

Received February 1, 1980; revised June 9, 1980

A new computational approach combining shock tracking and Newton's method is applied to steady one-dimensional flow through a variable area duct. On a transformed computational grid where the shock is fixed, the physical shock location appears explicitly as an unknown in a set of finite difference equations, and is coupled to the other unknowns. The space-time characteristics for the associated time-dependent problem are used in formulating the boundary conditions. The resulting system is solved by Newton's method. The computed results agree very well with an exact solution, and the Newton iterates converge rapidly in comparison to some explicit shock capturing, time-asymptotic methods.

1. INTRODUCTION

This paper presents a new computational approach to certain steady-state problems of inviscid fluid dynamics which combines shock tracking and Newton's method. This approach is applied to the model problem of one-dimensional flow in a duct of varying cross-sectional area. A common finite difference technique for obtaining steady solutions of such problems is the time-asymptotic method. Explicit time marching methods which capture shocks are relatively easy to implement, but they converge slowly. Increasing the dissipation of a difference scheme by adding an artificial viscosity may speed up convergence but excessively smears the captured shocks. Various ad hoc procedures such as corrected damping are an attempt to accelerate convergence while maintaining a fairly steep shock profile. For one-dimensional duct flow van Hove and Arts [1], Crocco [2], and others have investigated the convergence of the time marching method with respect to different difference schemes. Moretti [3] has implemented a time-dependent shock-tracking procedure for a variety of problems including one-dimensional duct flow.

An alternative to time marching is to properly pose the steady problem, discretize the equations and solve the resulting system using an iterative method. We implement such an approach in this paper. As in one of the time-dependent shock-tracking methods of Moretti, we define a transformation which maps the duct interval (containing the unknown physical shock location) to a computational interval where the shock is fixed. On the computational grid the physical shock location appears as

0021-9991/81/020364-11$02.00/0

an unknown in a set of finite difference equations, and is coupled to the other unknowns which are the values of the primitive flow variables at the mesh points. The resulting nonlinear algebraic system is solved using Newton's method. This approach implements shock tracking in a particularly simple way and avoids the relatively complicated equations and computational logic of time-dependent shock tracking. The explicitly determined standing shock wave is superior to the usual captured shock since it exactly satisfies the Rankine–Hugoniot jump conditions and is perfectly resolved. Newton's method converges quadratically near the solution and proves to be competitive with time-asymptotic method. For the application of Newton's method to shock-free flow in two space dimensions see [4].

2. Problem Formulation

A. *Continuous Problem*

The unsteady one-dimensional flow in a duct of variable cross-sectional area $A(x)$ is described in the physical space $0 \leqslant x \leqslant x_{\max}$ by

$$U_t + F_x + H = 0, \tag{1}$$

where

$$U = \begin{pmatrix} \rho A \\ \rho u A \\ \rho E A \end{pmatrix}, \qquad F = \begin{pmatrix} \rho u A \\ (\rho u^2 + p) A \\ (\rho E + p) u A \end{pmatrix}, \qquad H = \begin{pmatrix} 0 \\ -p \dfrac{dA}{dx} \\ 0 \end{pmatrix},$$

and ρ is density, u is velocity, $E = e + (u^2/2)$ where e is specific internal energy, and p is pressure. For simplicity, the equation of state is chosen here to be that of a perfect gas, $p = (\gamma - 1)\rho e$. In the steady-state equations (1) reduce to

$$F_x + H = 0 \tag{2}$$

which, together with the boundary conditions at $x = 0$ and $x = x_{\max}$ formulated in Section 2B, comprise the continuous problem which is to be approximated. For certain functions $A(x)$ and certain boundary specifications, a shock stands at some location $x = s$ in the duct. Considering s to be an unknown, the mapping

$$\begin{aligned} \xi = \xi(x, s) &= x/s, && x \leqslant s, \\ &= 1 + \frac{(k-1)(x-s)}{x_{\max} - s}, && x \geqslant s, \end{aligned} \tag{3}$$

is defined so that the physical shock location s is mapped into the computational coordinate $\xi = 1$. Likewise, the inflow location $x = 0$ and outflow location $x = x_{\max}$

are mapped into $\xi = 0$ and $\xi = k$ respectively (see Fig. 1). The constant k is defined in Section 2C. Equation (2) transforms to

$$\xi_x F_\xi + H = 0 \tag{4}$$

with appropriate transformed boundary conditions. The unknown s appears both in ξ_x and also through $A = A(x(\xi, s))$ since A must now be evaluated at $\xi =$ constant in the computational space.

B. *Boundary Conditions*

Boundary conditions must be specified at the inflow $\xi = 0$, the outflow $\xi = k$, and also at both sides of the shock $\xi = 1$, which is an internal boundary. The present approach to deriving these conditions makes use of the time-dependent formulation, especially the theory of characteristics for hyperbolic systems. From our point of view, the steady-state solution is just a special kind of unsteady solution, and the concept of information being propagated along space-time characteristics is retained. From this perspective, the characteristic compatibility equations which hold along characteristics in the time-dependent problem still hold in the steady state, only they transmit the particular information that the solution is not changing in time. This would be the situation, for example, when a steady state is reached in a time-dependent method of characteristics solution.

The characteristic slopes and the characteristic compatibility conditions for the time-dependent problem are derived in the Appendix. Assuming supersonic inflow, subsonic outflow, and a diverging duct with a single shock, the qualitative nature of these characteristics at the boundaries of the computational domain is depicted in Fig. 2 (other configurations can be treated analogously). At inflow $\xi = 0$, three

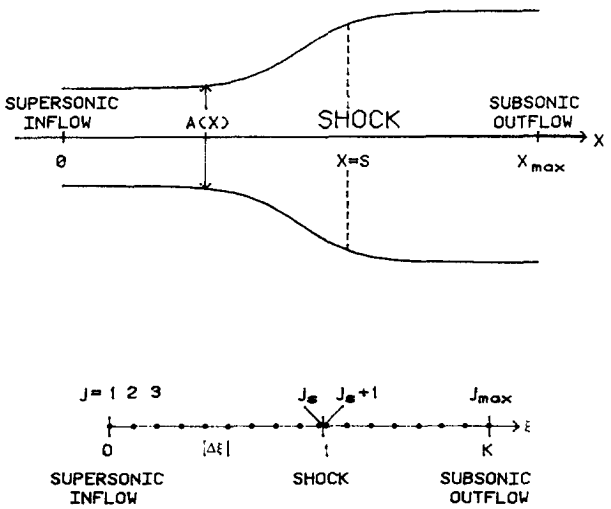

FIG. 1. Physical space x and computational space ξ with mesh point numbering indicated.

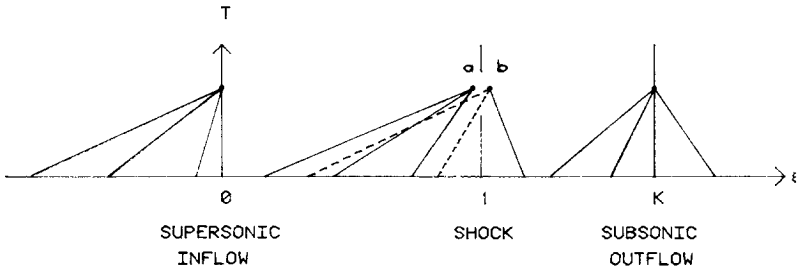

FIG. 2. Characteristics at boundaries of computational space for the time-dependent problem.

characteristics come from outside the computational interval $0 \leqslant \xi \leqslant k$, indicating that three boundary values must be specified at $\xi = 0$. These are

$$\begin{aligned} \rho(\xi = 0) &= \rho_{\text{in}}, \\ u(\xi = 0) &= u_{\text{in}}, \\ e(\xi = 0) &= e_{\text{in}}. \end{aligned} \tag{5}$$

At outflow $\xi = k$, two characteristics reach the boundary from inside the domain, and one from outside. This means that one boundary value must be specified, taken here to be

$$\rho(\xi = k) = \rho_{\text{out}}. \tag{6}$$

The two additional equations needed at $\xi = k$ are the characteristic compatibility equations (A3) and (A5) from the Appendix, with time derivatives dropped. They are:

$$-p\frac{\partial \rho}{\partial \xi} + \rho^2 \frac{\partial e}{\partial \xi} = 0, \tag{7}$$

$$\xi_x(c + u)\left[\frac{\partial p}{\partial \xi} + \rho c \frac{\partial u}{\partial \xi}\right] + \frac{dA/dx}{A} u\rho c^2 = 0, \tag{8}$$

where c is the sound speed.

The two sides of the shock, points a and b in Fig. 2, are related by the shock jump conditions

$$\begin{aligned} (\rho u)_a &= (\rho u)_b, \\ (p + \rho u^2)_a &= (p + \rho u^2)_b, \\ (e + p/\rho + u^2/2)_a &= (e + p/\rho + u^2/2)_b. \end{aligned} \tag{9}$$

Additionally, at pre-shock point a, three characteristic compatibility conditions hold; these are equivalent in the steady state to the transformed differential equations (4).

At post-shock point b, two characteristics (dashed lines) cross the shock and are thus prohibited from carrying information. The characteristic compatibility condition corresponding to the one admissible characteristic reaching point b from the interval $1 \leqslant \xi \leqslant k$ is (A4), which becomes in the steady state

$$\xi_x(c-u)\left[\frac{\partial p}{\partial \xi} - \rho c \frac{\partial u}{\partial \xi}\right] - \frac{dA/dx}{A} u\rho c^2 = 0. \tag{10}$$

While it may appear that these seven conditions (Eqs. (9), (10), and the three differential equations at point a) overspecify the shock boundary, it will be seen that the "extra" equation (10) is needed to determine the unknown physical shock system s. The use of other equations to close the system will be discussed at the end of Section 3.

C. *Finite Difference Equations and Solution by Newton's Method*

A uniform finite difference mesh is used in the computational space, except that there are two mesh points at $\xi = 1$ which represent the pre- and post-shock states (Fig. 1). The mesh points ξ_j are given by $\xi_j = (j-1)\Delta\xi$ for $j = 1, 2, \ldots, j_s$ and $\xi_j = (j-2)\Delta\xi$ for $j = j_s + 1, \ldots, j_{\max}$, where $\Delta\xi = 1/(j_s - 1)$ and j_s and $j_s + 1$ are the pre- and post-shock points, respectively. The constant k in (3) is used to select the relative computational resolutions of the physical intervals $0 \leqslant x \leqslant s$ and $s \leqslant x \leqslant x_{\max}$. The variable p is eliminated by using the equation of state, so that the unknowns at each mesh point $1 \leqslant j \leqslant j_{\max}$ are ρ_j, u_j, and e_j.

At "interior" mesh points $j = 2, \ldots, j_s - 1$ and $j = j_s + 2, \ldots, j_{\max} - 1$, Eqs. (4) are approximated by the three centered difference equations

$$(\xi_x)_j \frac{[F_{j+1} - F_{j-1}]}{2\Delta\xi} + H_j = 0. \tag{11}$$

At inflow point $j = 1$, the three inflow conditions (5), namely, $\rho_1 = \rho_{\text{in}}$, $u_1 = u_{\text{in}}$, and $e_1 = e_{\text{in}}$ are specified. At outflow point $j = j_{\max}$, $\rho_{j_{\max}} = \rho_{\text{out}}$ is specified, along with first-order accurate backward difference approximations of the two compatibility conditions (7) and (8). At the pre-shock point $j = j_s$, three backward difference equations

$$(\xi_x)_j \frac{[F_j - F_{j-1}]}{\Delta\xi} + H_j = 0 \tag{12}$$

are specified. The jump conditions (9) relate $(\rho, u, e)_{j_s}$ to $(\rho, u, e)_{j_s+1}$. Finally, at the post-shock point $j_s + 1$, a first-order forward difference approximation of (10) is used. These equations provide $N = 3j_{\max} + 1$ algebraic equations in the N unknowns $u_1, \rho_1, e_1, u_2, \rho_2, e_2, \ldots, u_{j_{\max}}, \rho_{j_{\max}}, e_{j_{\max}}, s$ (in this order).

Let $\mathbf{v}$ be the vector of unknowns, $G(\mathbf{v}) = 0$ be the system of N nonlinear equations in N unknowns, and $B(\mathbf{v})$ be the $N \times N$ Jacobian matrix for this system, i.e., $B_{ij} = \partial G_i / \partial v_j$. A modified Newton's method for finding the $(k+1)$ iterate is given by

$$B(\mathbf{v}^k) \cdot (\mathbf{v}^{k+1} - \mathbf{v}^k) = -\alpha G(\mathbf{v}^k) \tag{13}$$

where α is a "damping factor." For $\alpha = 1$ (usual Newton's method) if $G(\mathbf{v}^*) = 0$ and $B(\mathbf{v}^*)$ is nonsingular then in a neighborhood of $\mathbf{v}^*$ the iteration converges quadratically, i.e., $\varepsilon^n \leqslant d(\varepsilon^{n-1})^2$, where the iteration error ε is defined by $\varepsilon^{n+1} = \max_j |v_j^{n+1} - v_j^n|$. However, the initial guess $\mathbf{v}^0$ must often be close to $\mathbf{v}^*$ to ensure convergence. This guess $\mathbf{v}^0$ is easily supplied for one-dimensional duct flow, and a poor guess can often be improved by taking $\alpha < 1$ until the "damped Newton" iterates are close enough to the correct solution. Furthermore, although Newton's method converges in a few iterations, a linear system must be solved at each iteration. For standard Gaussian elimination and a full matrix this usually requires $O(N^3)$ operations, so that for large N the computational time becomes prohibitive. For one-dimensional duct flow the Jacobian has a special structure, being nearly block tridiagonal (with some singular diagonal blocks) except for an extra row resulting from the compatibility condition (10) at the shock, and an extra column due to the unknown shock location s.

To extend this methodology to multidimensional problems, one may try to take advantage of the fact that $B(\mathbf{v})$ is sparse and has a nearly banded structure [4]. Alternatively, one may use a quasi-Newton method [5], or use Newton's method on columns or rows in the computational grid. Since a good initial guess is difficult to obtain for multidimensional problems, we are considering a hybrid technique where the guess is supplied by a shock-capturing finite difference method. These ideas will be explored in future work.

3. Results

In this section we present our computational results for the model problem, addressing first the question of finite difference accuracy. The problem as formulated in Section 2 was solved computationally with $j_{max} = 16$ and 32 mesh points. The specified inflow conditions were $A_{in} = 1.050$, $\rho_{in} = 0.502$, $e_{in} = 1.897$, $u_{in} = 1.299$ and the outflow conditions were $A_{out} = 1.745$, $\rho_{out} = 0.776$. The duct cross-sectional area $A(x)$ is shown in Fig. 3, and the initial guess for the Newton iteration was taken to be

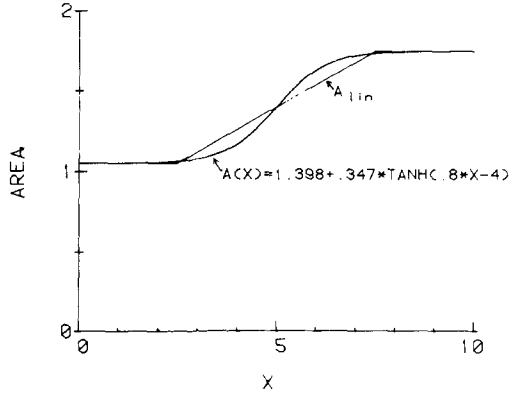

FIG. 3. Duct areas $A(x)$, A_{lin} used in computations.

the exact solution for a perturbed duct shape A_{1in}. (This problem with a duct much like A_{1in} was solved by Crocco [2] using a time-asymptotic shock-capturing method). For the present problem an exact solution can be computed, and in this exact solution a shock stands at $A = 1.347$, $x = s_{exact} = 4.816$. The computational results for density obtained using 16 and 32 mesh points are compared with the exact solution in Figs. 4a,b. The shock jump conditions are, of course, always identically satisfied by the computed results, although for a shock whose position, and hence strength, is slightly in error. This error in shock position, defined as $(s_{exact} - s_{computed})/s_{exact}$, is rather small, being 0.028 and 0.005 for $j_{max} = 16$ and 32, respectively. Indeed, the computed results with 32 mesh points are virtually indistinguishable from the exact solution. These results may be favorably compared with shock capturing results (e.g., [1] and [2]) or with time-dependent shock-tracking results [3].

The question of computational efficiency will now be discussed. Defining the iteration error as in Section 2C, the above computation with $\alpha = 1$ and $\varepsilon^1 \sim 10^{-1}$ converges quadratically to $\varepsilon^5 \sim 10^{-13}$ (machine accuracy) in five iterations. This

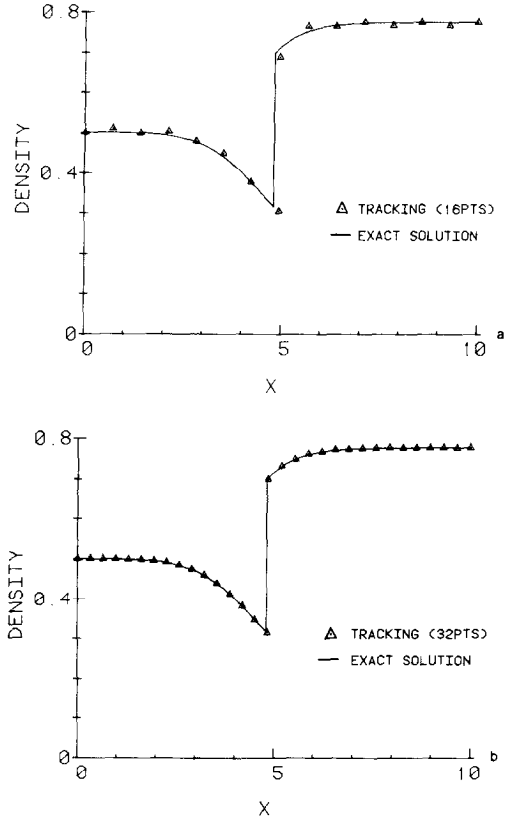

FIG. 4. Density vs. position for the exact and computed solutions with (a) 16 and (b) 32 mesh points.

computation takes about 6 sec of CDC 6500 computer time for $j_{\max} = 32$ when Gaussian elimination with partial pivoting is used to solve (13). Like others (e.g., [6]) we observe that the number of iterations needed to converge to a specified error tolerance is largely independent of the number of mesh points. While it is generally expected that solving an $N \times N$ linear system will take $O(N^3)$ time, we observe $O(N^2)$ for our particular system, probably due to the sparsity and structure of the particular Jacobian matrix $B(\mathbf{v})$.

In order to make a *very* rough comparison between the efficiencies of the time asymptotic method and the present approach, the explicit finite difference methods of MacCormack [7] and Lax [8] were used to obtain shock-capturing, time-asymptotic solutions of this problem using the same initial guess. These time-asymptotic computations used 32 mesh points and a time step equal to 0.9 of that allowed by the CFL stability criterion (i.e., $\max(|u| + c)\,\Delta t/\Delta x = 0.9$). At the downstream boundary ρ_{out} was specified and compatibility conditions (A3) and (A5) were used. We believe that the computational efficiencies for these two methods, the first being mildly dissipative and the second highly dissipative, are likely to bracket those for most explicit time-asymptotic methods, including time-dependent shock tracking. In Fig. 5 the iteration error ε is plotted versus CDC 6500 computer time for both time-asymptotic methods and for the present tracking method. Computation time was chosen as the standard of comparison because the time asymptotic method requires many iterations with little time per iteration, while Newton's method requires few iterations but takes more time per iteration. The mildly dissipative MacCormack scheme converges only when a small amount of artificial viscosity is added (like that in [1, p. 20]), and gives a solution with a slightly oscillatory captured shock at approximately the correct location. The highly dissipative Lax scheme converges fairly rapidly but yields a very smooth solution without recognizable shock structure (see [1]). Steady shock tracking with Newton's method is more efficient than either of these time-asymptotic methods and gives the very good results of Fig. 4b.

Next we examine the convergence of the Newton iterations when the initial guess is less accurate than above. In this case the method is found to converge to *a* solution of the steady difference equations by taking α small (~ 0.1) for the first few iterations (giving linear convergence) and then switching to $\alpha = 1$. However, when the initial guess was taken to be that used by Crocco [2], namely a shock-free solution with the downstream condition ρ_{out} suddenly imposed, the Newton iterates with $j_{\max} = 16$ converged to the "spurious" solution shown in Fig. 6. This computed solution is a bona fide solution of the difference equations, but is one which we believe violates the entropy condition [9] at the spurious jump in the region upstream of the tracked shock. Many time asymptotic methods which do not specifically enforce the entropy condition can also converge to such spurious solutions (e.g., [10]). In the present context, these spurious solutions are readily identifiable. When the converged spurious solution is perturbed to nearly agree with the correct difference solution, and this is taken as an initial guess, Newton's method converges to the correct solution (i.e., the one satisfying the entropy condition, Fig. 4a).

Finally we consider the choice of Eq. (10) as the Nth equation which closes the

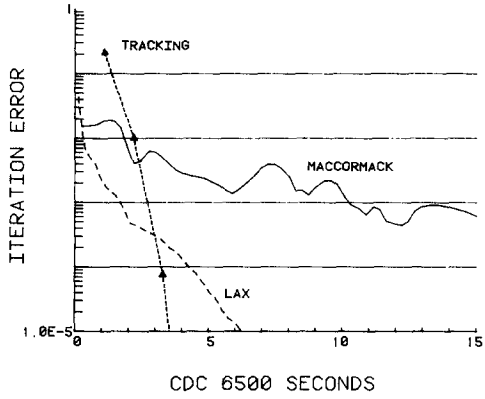

FIG. 5. Iteration error ε versus CDC 6500 computer time for 32 mesh point solutions obtained with the steady-tracking method and the time-asymptotic shock-capturing methods of MacCormack and Lax.

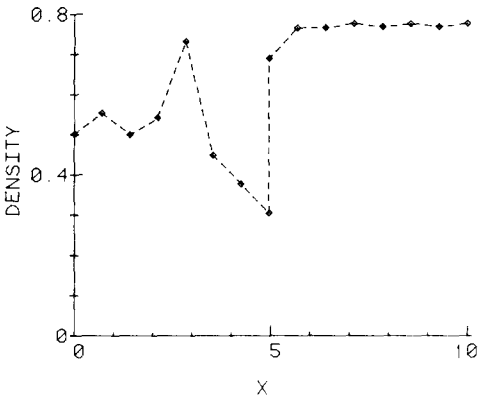

FIG. 6. Density versus position for the "spurious" solution obtained with the initial guess of Crocco [2].

algebraic system. This choice was based on the time-dependent shock tracking formulation, where only one of the characteristics is admissible (the others cross the shock; dashed lines in Fig. 2) and where use of inadmissible characteristic information usually leads to instability. While according to the philosophy outlined at the beginning of Section 2B the time-dependent characteristics have meaning for a steady solution, it is also possible to look directly at the steady-state formulation (without reference to the time-dependent problem) in which such space-time characteristics are meaningless. While it is difficult to resolve this paradox, it is certainly possible that some other condition applied at the post-shock point will work equally well in the steady formulation. Indeed, we have obtained results of accuracy comparable to those presented above by replacing Eq. (10) with the mass conservation equation (first component of Eq. (4)) using first-order accurate forward differencing at $j = j_s + 1$ (and with other formulations as well).

APPENDIX: Derivation of Characteristic Compatibility Conditions

Some background on characteristic compatibility equations can be found in [11]. Consider Eqs. (1) and the time-dependent transformation $\xi = \xi(x, t)$, $T = t$. Equations (1) transform to

$$U_T + \xi_t U_\xi + \xi_x F_\xi + H = 0. \tag{A1}$$

Define $Q = (\rho, u, e)^t$. Then Eq. (A1) is equivalant to

$$Q_T + (\xi_t I + \xi_x D_0)\, Q_\xi + E_0 = 0, \tag{A2}$$

where $C = [\partial U/\partial Q]$, $D = [\partial F/\partial Q]$, $D_0 = C^{-1}D$, $E_0 = C^{-1}\{[(dA/dx)/A]\, F + H\}$. Specifically,

$$D_0 = \begin{pmatrix} u & \rho & 0 \\ k_1/\rho & u & k_2/\rho \\ 0 & p/\rho & u \end{pmatrix}, \qquad E_0 = \frac{dA/dx}{A}\, u \begin{pmatrix} \rho \\ 0 \\ p/\rho \end{pmatrix},$$

where $k_1 = (\partial p/\partial \rho)_e$, $k_2 = (\partial p/\partial e)_\rho$.

The characteristic matrix for Eq. (A2) is

$$A^*(\lambda_1, \lambda_2) = \lambda_1 I + \lambda_2(\xi_t I + \xi_x D_0) = \begin{pmatrix} \sigma & \Lambda_2 \rho & 0 \\ k_1 \Lambda_2/\rho & \sigma & k_2 \Lambda_2/\rho \\ 0 & \Lambda_2 p/\rho & \sigma \end{pmatrix},$$

where $\sigma = \Lambda_1 + \Lambda_2 u$, $\Lambda_1 = \lambda_1 + \lambda_2 \xi_t$, $\Lambda_2 = \lambda_2 \xi_x$. The characteristic condition is $\det A^* = \sigma(\sigma^2 - \Lambda_2^2 c^2) = 0$ where $c^2 = k_1 + k_2 p/\rho^2$ is the square of the speed of sound. Corresponding to the three characteristic conditions $\sigma_1 = 0$, $\sigma_{2,3} = \pm \Lambda_2 c$ are the three left null vectors (defined by $\mathbf{l} A^* = 0$) $\mathbf{l}_1 = (-p, 0, \rho^2)$, $\mathbf{l}_{2,3} = (-k_1 \sigma_{2,3}, \rho \Lambda_2 c^2, -k_2 \sigma_{2,3})$. The three characteristic compatibility conditions are obtained by left multiplying Eqs. (A2) by $\mathbf{l}_1$, $\mathbf{l}_2$, and $\mathbf{l}_3$. They are:

$$-p \frac{\partial \rho}{\partial T} + \rho^2 \frac{\partial e}{\partial T} + q\left(-p \frac{\partial \rho}{\partial \xi} + \rho^2 \frac{\partial e}{\partial \xi}\right) = 0, \tag{A3}$$

$$-\sigma \frac{\partial p}{\partial T} + \rho c^2 \Lambda_2 \frac{\partial u}{\partial T} + (-\sigma q + c^2 \Lambda_2 \xi_x) \frac{\partial p}{\partial \xi} + \rho c^2(-\sigma \xi_x + \Lambda_2 q) \frac{\partial u}{\partial \xi} - \frac{dA/dx}{A}\, u \rho c^2 \sigma = 0, \tag{A4), (A5}$$

where $q = \xi_t + u\xi_x$ and where $\sigma = \sigma_2$, $\sigma = \sigma_3$ for Eqs. (A4), (A5), respectively.

The slopes of the characteristics in the (ξ, T) plane are given by $d\xi/dT = q$, $q \mp \xi_x c$. These slopes play a crucial role in formulating the boundary and shock conditions.

ACKNOWLEDGMENT

This work was sponsored by the Reentry Technology Program Office and by NSWC Independent Research funds. The authors wish to thank Louis Hackerman for his programming assistance.

REFERENCES

1. W. VAN HOVE AND A. ARTS, "Comparison of Several Finite Difference Schemes for Time Marching Methods as Applied to One Dimensional Nozzle Flow," Von Karman Institute for Fluid Dynamics Technical Note 132, June 1979.
2. L. CROCCO, *AIAA J.* **3** (1965), 1824–1832.
3. G. MORETTI, "Thoughts and Afterthoughts About Shock Computations," Polytechnic Institute of Brooklyn PIBAL Report 72–37, Dec. 1972.
4. A. RIZZI, "Solution by Newton's Method to the Steady Transonic Euler Equations (Proceedings of the Sixth International Conference on Numerical Methods in Fluid Dynamics)," Lecture Notes in Physics 90, Springer-Verlag, New York, p. 460.
5. W. C. RHEINBOLDT, "Methods for Solving Systems of Nonlinear Equations," Regional Conference Series in Applied Mathematics, Vol. 14, SIAM, Philadelphia, 1974.
6. E. L. ALLGOWER AND S. F. MCCORMICK, *Num. Math.* **29** (1978), 237–260.
7. R. W. MACCORMACK, "The Effect of Viscosity in Hypervelocity Impact Cratering," AIAA Paper 69–354, 1969.
8. P. D. LAX, *Commun. Pure Appl. Math.* **7** (1954), 159–193.
9. P. D. LAX, "Hyperbolic Systems of Conservation Laws and the Mathematical Theory of Shock Waves," Regional Conference Series in Applied Mathematics, Vol. 11, SIAM, Philadelphia, 1973.
10. A. HARTEN, J. M. HYMAN, AND P. D. LAX, *Commun. Pure Appl. Math.* **29** (1976), 297–322.
11. R. COURANT AND D. HILBERT, "Methods of Mathematical Physics," Vol. 2, Interscience, New York, 1962, pp. 577–599.