JOURNAL OF COMPUTATIONAL PHYSICS **135,** 250–258 (1997)
ARTICLE NO. CP975705

Approximate Riemann Solvers, Parameter Vectors, and Difference Schemes

P. L. Roe

Royal Aircraft Establishment, Bedford, United Kingdom

Received August 14, 1980; revised March 30, 1981

Several numerical schemes for the solution of hyperbolic conservation laws are based on exploiting the information obtained by considering a sequence of Riemann problems. It is argued that in existing schemes much of this information is degraded and that only certain features of the exact solution are worth striving for. It is shown that these features can be obtained by constructing a matrix with a certain "Property U." Matrices having this property are exhibited for the equations of steady and unsteady gasdynamics. In order to construct them, it is found helpful to introduce "parameter vectors" which notably simplify the structure of the conservation laws.

INTRODUCTION

We consider the initial-value problem for a hyperbolic system of conservation laws, i.e., we seek a vector $\mathbf{u}(x, t)$ such that

$$\mathbf{u}_t + \mathbf{F}_x = 0 \tag{1}$$

and

$$\mathbf{u}(x, 0) = \mathbf{u}_0(x), \tag{2}$$

where $\mathbf{F}$ is some vector-valued function of $\mathbf{u}$, such that the Jacobian matrix $A = \partial\mathbf{F}/\partial\mathbf{u}$ has only real eigenvalues.

We introduce the discrete representation $x_i = x_0 + i\,\Delta x$, $t_n = t_0 + n\,\Delta t$ and suppose that $\mathbf{u}_i^n$ is some approximation to $\mathbf{u}(x_i, t_n)$.

A multitude of strategies have been devised to obtain numerical results for the discrete problem, and their relative merits are still largely unclear. We shall address in this paper some questions relating to those methods which attempt to construct the solution by solving a succession of Riemann problems. Recall that the Riemann problem is the initial-value problem obtained when the general data, Eq. (2), is specialised to

$$\mathbf{u}(x, 0) \equiv \mathbf{u}_L(x < 0), \quad \mathbf{u}(x, 0) \equiv \mathbf{u}_R(x > 0). \tag{3}$$

That is, we study the breakup of a single discontinuity. Those cases where $\mathbf{F}(\mathbf{u})$ is linear are well known and essentially trivial (see [1] for a brief discussion). Those cases where u is scalar and F is nonlinear can be surprisingly intricate, but have been thoroughly investigated [2]. If $\mathbf{u}$ is a vector and $\mathbf{F}$ is nonlinear, then the problem involves nonlinear algebraic equations together with, usually, logical conditions which express the fact that a given member of the wave system may be present either as a shockwave or as an expansion fan. A review of results relating to the general Riemann problem has been given by Lax [3]. In general, the most efficient way to solve these equations will depend on the system of conservation laws (1) from which they derive; ingenuity is required to exploit special features of each individual system. For the special case of the unsteady Euler equations in one space dimension, an algorithm was devised by Godunov [4] and is available in the books by Richtmyer and Morton [5] and Holt [6]. A variant which converges more rapidly was worked out by van Leer [7].

The usual way of incorporating the Riemann problem into the numerical solution is to take $(\mathbf{u}_i^n, \mathbf{u}_{i+1}^n)$, for each i in turn, as pairs of states defining a sequence of Riemann problems, which are then thought of as providing information about the solution within each interval $(i, i + 1)$. Various individual methods are then distinguished by the way in which this information is put to use. Briefly, we review these below.

Godunov [4] supposed that the initial data could be replaced by a piecewise constant set of states with the discontinuities at $\{x_{i+1/2}\}$. He then found the exact solution to this simplified problem. After some time step Δt (less than Δx divided by the greatest wavespeed found in the Riemann solutions) he replaced the exact solution by a new piecewise constant approximation, whilst preserving integral properties of the conserved variable $\mathbf{u}$. The first major extension to this line of approach was made by van Leer [7], who approximated the data and the solution at

Reprinted from Volume 43, Number 2, October 1981, pages 357–372.

0021-9991/97 $25.00

each subsequent time level, by piecewise linear segments, allowing discontinuities between the segments. This required the solution to an interaction problem which was more general than Riemann's, but raised the order of accuracy of the method from one to two.

A parallel line of development was initiated by Glimm [8], who followed Godunov as far as the exact solution to the simplified problem, but then obtained the new approximation by a random sampling procedure. The sampling produces solutions, which are conservative only on the average, but has the advantage that near a moving, isolated, discontinuity, the solution is incremented either by the full amount of the jump, or not at all. In this way, initially sharp discontinuities remain sharp, and for some technical problems this property is important. More refined sampling procedures have been introduced (see, e.g., Chorin [9], but so far the accuracy of the method approaches unity from below.

It seems to the present author that the expense of producing an accurate solution to the Riemann problem would only be justified if the abundance of information which is thereby made available could be put to some rather sophisticated use. Indeed, it must somehow be true that the accuracy with which it is worthwhile solving the Riemann problem will be limited by the way we intend using the solution. For example, we may consider the use of less accurate solutions in existing methods. Harten and Lax [10] have devised an approximate Riemann solver particularly designed for incorporation into Godunov-type or Glimm-type difference schemes. The approximation developed herein could be (and has been) used in the same way, but in comparison with the Harten and Lax approximation it suffers the theoretical disadvantage of not having a naturally constructed entropy condition; this point will be discussed in more detail later. On the other hand, the present approximation is designed to provide the information needed to obtain high formal accuracy, following the strategy set out by Roe [1]. That paper essentially described a mechanism by which any algorithm developed for numerical solution of the linear advection equation

$$u_t + au_x = 0 \tag{4}$$

can be generalised to the case of nonlinear systems. A large body of unpublished numerical results for Burger's equation, the nonlinear shallow-water equations, and the steady and unsteady Euler equations, demonstrates that all qualitative features of each algorithm are faithfully transmitted by the mechanism. The same evidence suggests that accuracy also carries over, at least to third order. An essential stage in the mechanism is the approximate solution of a nonlinear Riemann problem.

In this paper we consider approximate solutions which are exact solutions to an approximate problem, viz.,

$$\mathbf{u}_t + \tilde{A}\mathbf{u}_x = 0,$$

where $\tilde{A}$ is a constant matrix, and the data $(\mathbf{u}_L, \mathbf{u}_R)$ are of course unaltered. $\tilde{A}$ is to be chosen so that it is representative of local conditions. Candidates which immediately come to mind are $\tilde{A} = \frac{1}{2}(A_L + A_R)$, or $\tilde{A} = A(\frac{1}{2}(\mathbf{u}_L + \mathbf{u}_R))$. We shall, however, only accept a matrix $\tilde{A}(\mathbf{u}_L, \mathbf{u}_R)$ which satisfies the following list of properties:

(i) It constitutes a linear mapping from the vector space $\mathbf{u}$ to the vector space $\mathbf{F}$.

(ii) As $\mathbf{u}_L \to \mathbf{u}_R \to \mathbf{u}$, $\tilde{A}(\mathbf{u}_L, \mathbf{u}_R) \to A(\mathbf{u})$, where $A = \partial\mathbf{F}/\partial\mathbf{u}$.

(iii) For any $\mathbf{u}_L, \mathbf{u}_R, \tilde{A}(\mathbf{u}_L, \mathbf{u}_R) \times (\mathbf{u}_L - \mathbf{u}_R) = \mathbf{F}_L - \mathbf{F}_R$.

(iv) The eigenvectors of $\tilde{A}$ are linearly independent.

In general, neither of the candidates mentioned above would satisfy condition (iii).

Once such a matrix has been constructed, then its eigenvalues can be considered as the wavespeeds of the Riemann problem, and the projections of $\mathbf{u}_R - \mathbf{u}_L$ onto its eigenvectors are the jumps which take place between intermediate states [1]. In [1], it is shown that (iii) is a sufficient condition for the algorithm produced by the mechanism to be conservative. It is also shown that (iii) and (iv) are necessary and sufficient conditions for the algorithm to "recognize" a shockwave. By this we mean that if $(\mathbf{u}_L, \mathbf{u}_R)$ satisfy the jump condition

$$(\mathbf{F}_L - \mathbf{F}_R) = S(\mathbf{u}_L - \mathbf{u}_R) \tag{5}$$

for some scalar S, then, by (iii), S is an eigenvalue of $\tilde{A}$. A projection of $(\mathbf{u}_L, \mathbf{u}_R)$ onto the eigenvectors of $\tilde{A}$ will (because of (iv)) be solely onto the eigenvector which corresponds to S. In this special case, the solution of the Riemann problem will be exact.

Evidently (ii) is a necessary condition if we are to recover smoothly the linearized algorithm from the nonlinear version.

Collectively, this list of properties has been christened Property U (since it is intended to ensure *uniform* validity across discontinuities). It is thought to specify desirable properties of a Riemann solver because of the following heuristic argument. Consider a region of the x, t plane containing $O(N^2)$ points and traversed by a finite number of discontinuities. At the majority of points, nothing very special is happening, and the choice of method is not critical. At a number of points which is $O(N)$, we are close to a single discontinuity; here Property U will allow us to recognize the situation and to deal with it appropriately. At a number of points which is $O(1)$, we are close to two or more discontinuities, and such a situation will not be resolved on a fixed grid. In accordance with the shock-

capturing philosophy, we must here put our trust in conservation, which is also assured by Property U.

If we have to deal with more space dimensions, say (x, y) as well as t, then successful shock-capturing involves additional difficulties. One of these is that there is no obvious "generalized Riemann problem" to serve as a building block. In practice, it has been found [12] that if the multidimensional operator is split into a sequence of one-dimensional operators, then the present method may be applied to each operator. This gives good results so long as the shockwaves remain aligned with the computing grid; such shockwaves are accurately recognized and appropriately treated. Problems arise when the shockwave lies obliquely across the grid, and they are particularly severe as the solutions attempt to reach a steady state. This is because neither of the split operators, by itself, recognizes the oblique shock as being in equilibrium. Under these conditions, the concept of operator splitting becomes rather dubious, but the finding of a satisfactory alternative is beyond the scope of this paper.

CONSTRUCTION OF $\tilde{A}$

It is very easy to construct $\tilde{A}$ so that it meets conditions (i) and (ii) above. Condition (iv) is easily checked a posteriori. The difficulty lies entirely with condition (iii). The existence of an $\tilde{A}$ satisfying condition (iii) follows from the mean value theorem. Let θ be a parameter which varies linearly between 0 and 1 along a straight path connecting $\mathbf{u}_L$ to $\mathbf{u}_R$, so that

$$\mathbf{u}(\theta) = \mathbf{u}_L + \theta(\mathbf{u}_R - \mathbf{u}_L), \quad d\mathbf{u} = (\mathbf{u}_R - \mathbf{u}_L)\, d\theta.$$

Then

$$\begin{aligned}\mathbf{F}(\mathbf{u}_R) - \boldsymbol{F}(\mathbf{u}_L) &= \int_0^1 \frac{d\mathbf{F}}{d\theta}\, d\theta \\ &= \int_0^1 A(\theta)\, \frac{d\mathbf{u}}{d\theta}\, d\theta \\ &= \int_0^1 A(\theta)\, d\theta \cdot (\mathbf{u}_R - \mathbf{u}_L),\end{aligned}$$

whence

$$\tilde{A} = \int_0^1 A(\theta)\, d\theta.$$

However, the integrals involved may not emerge in closed form, or the closed form might be expensive to compute. By a more subtle choice of path, candidates for $\tilde{A}$ can be found which are integrable, and an approach similar to this has been taken by Osher and Solomon [12].

Since computational speed is a major requirement, it is interesting to remark that in the common case where each component of $\mathbf{F}$ is a rational function of the components of $\mathbf{u}$, an $\tilde{A}$ can be found whose entries are likewise rational functions of the components of $\mathbf{u}$. The following identities are obviously true for arbitrarily large jumps of any scalar quantities, where $\Delta(\cdot) = (\cdot)_L - (\cdot)_R$:

$$\Delta(p + q) \equiv \Delta p + \Delta q, \tag{6a}$$

$$\Delta(pq) \equiv \overline{p}\, \Delta q + \overline{q}\, \Delta p, \tag{6b}$$

$$\Delta(1/q) \equiv -\Delta q / q^{*2}, \tag{6c}$$

where the overbar denotes an arithmetic, and the asterisk a geometric, mean value. In this way we shall be led to a formula of the form

$$F_j(\mathbf{u}_L) - F_j(\mathbf{u}_R) = \sum_i \tilde{a}_{ij}[(u_i)_L - (u_i)_R], \tag{7}$$

where each $\tilde{a}_{ij}$ depends on $\mathbf{u}_L$ and $\mathbf{u}_R$. The matrix whose entries are $\tilde{a}_{ij}$ satisfies conditions (i) to (iii) of Property U.

Matrices for the unsteady Euler equations were constructed in this way and are given in [11]. However, there are many disadvantages to this construction.

In the first place, it is far from unique, as may easily be seen by applying the second of the above formulae in an obvious way to $\Delta(pqr)$ and noting that the outcome depends on the order of doing the multiplications. Second, the formulae obtained tend to be rather cumbersome, especially bearing in mind that we actually want to obtain analytical expressions for the eigenvalues and eigenvectors of $\tilde{A}$. Not only does the algebra become almost impossible to carry through without error, but the whole object of creating a neat and efficient approximate solution is being defeated. We present in the following section a simple device which has so far worked each time that it has been tried. Indeed, in the case of the Euler equations, it simplifies the structure so much that it may well have applications outside the present context.

THE PARAMETER VECTOR

The inspiration for this section is taken from the common experience in analytic geometry that a plane curve $y(x)$ may in some cases be much more easily described by a parametric form $y = y(t)$, $x = x(t)$. We may therefore expect that a multidimensional manifold such as $\mathbf{F}(\mathbf{u})$ may sometimes be more amenable if represented as $\mathbf{F} = \mathbf{F}(\mathbf{w})$, $\mathbf{u} = \mathbf{u}(\mathbf{w})$, where we may speak of $\mathbf{w}$ as a parameter vector. We now exhibit a very useful parameter vector for the Euler equations, which we write as

$$\mathbf{u}_t + \mathbf{F}_x + \mathbf{G}_y + \mathbf{H}_z = 0, \tag{8}$$

where

$$\mathbf{u} = \begin{pmatrix} \rho \\ \rho u \\ \rho v \\ \rho w \\ e \end{pmatrix}, \qquad \mathbf{F} = \begin{pmatrix} \rho u \\ p + \rho u^2 \\ \rho uv \\ \rho uw \\ u(p+e) \end{pmatrix},$$

$$\mathbf{G} = \begin{pmatrix} \rho v \\ \rho uv \\ p + \rho v^2 \\ \rho vw \\ v(p+e) \end{pmatrix}, \quad \mathbf{H} = \begin{pmatrix} \rho w \\ \rho uw \\ \rho vw \\ p + \rho w^2 \\ w(p+e) \end{pmatrix}, \tag{9}$$

in which ρ = density, p = static pressure, (u, v, w) = velocity in cartesian coordinates (x, y, z), and e is the total energy, related to the other variables by an equation of state which, for a perfect gas, is

$$e = \frac{p}{(\gamma - 1)} + \tfrac{1}{2}\rho(u^2 + v^2 + w^2). \tag{10}$$

Various special cases (two-dimensional, steady, etc.) follow obviously by striking out the irrelevant terms. We assert that every component of **u**, **F**, **G**, **H** is merely quadratic in the components of

$$\mathbf{w} = \rho^{1/2}(1, u, v, w, H)^{\mathrm{T}}, \tag{11}$$

where total enthalpy H is related to previously defined quantities by $\rho H = e + p$. Other choices are possible for this fifth component, but lead to marginally more complicated algebra.

The truth of this is obvious in most cases. For example, $u_1 = w_1^2$, $G_4 = w_3 w_4$, etc. Some of the less obvious ones are

$$u_5 = \frac{w_1 w_5}{\gamma} + \frac{(\gamma - 1)}{2\gamma}(w_2^2 + w_3^2 + w_4^2), \tag{12a}$$

$$F_2 = \frac{\gamma - 1}{\gamma} w_1 w_5 + \frac{\gamma + 1}{2\gamma} w_2^2 - \frac{\gamma - 1}{2\gamma}(w_3^2 + w_4^2), \tag{12b}$$

and G_3, H_4 follow by symmetry.

It is now very easy to represent any jump in the spaces **u**, **F**, **G**, **H** in terms of its image in the space **w**, merely by use of (6b). For example, given any pair of states $(\mathbf{u}_L, \mathbf{u}_R)$ and their images $(\mathbf{w}_L, \mathbf{w}_R)$ we can write

$$(\mathbf{u}_L - \mathbf{u}_R) \equiv \tilde{B}(\mathbf{w}_L - \mathbf{w}_R), \tag{13}$$

where

$$\tilde{B} = \begin{pmatrix} 2\overline{w}_1 & 0 & 0 & 0 & 0 \\ \overline{w}_2 & \overline{w}_1 & 0 & 0 & 0 \\ \overline{w}_3 & 0 & \overline{w}_1 & 0 & 0 \\ \overline{w}_4 & 0 & 0 & \overline{w}_1 & 0 \\ \dfrac{\overline{w}_5}{\gamma} & \dfrac{\gamma - 1}{\gamma}\overline{w}_2 & \dfrac{\gamma - 1}{\gamma}\overline{w}_3 & \dfrac{\gamma - 1}{\gamma}\overline{w}_4 & \dfrac{\overline{w}_1}{\gamma} \end{pmatrix} \tag{14}$$

and all the overbars denote arithmetic means. Likewise we can write

$$(\mathbf{F}_L - \mathbf{F}_R) \equiv \tilde{C}(\mathbf{w}_L - \mathbf{W}_R), \tag{15}$$

where

$$\tilde{C} = \begin{pmatrix} \overline{w}_2 & \overline{w}_1 & 0 & 0 & 0 \\ \dfrac{\gamma - 1}{\gamma}\overline{w}_5 & \dfrac{\gamma - 1}{\gamma}\overline{w}_2 & -\dfrac{\gamma - 1}{\gamma}\overline{w}_3 & -\dfrac{\gamma - 1}{\gamma}\overline{w}_4 & \dfrac{\gamma - 1}{\gamma}\overline{w}_1 \\ 0 & \overline{w}_3 & \overline{w}_2 & 0 & 0 \\ 0 & \overline{w}_4 & 0 & \overline{w}_2 & 0 \\ 0 & \overline{w}_5 & 0 & 0 & \overline{w}_2 \end{pmatrix}. \tag{16}$$

Obviously these very simple formulae are closely related to the homogeneous property of the Euler equations. The vectors **u**, **F**, **G**, **H** are each homogeneous of order one with respect to any of the others; also each of them is homogeneous of order two with respect to **w**. However, the homogeneous property is not essential for the existence of a parameter vector. The reader may experiment with parameter vectors for the shallow-water equations, which do not have the homogeneous property.

EIGENVECTORS AND EIGENVALUES FOR THE EULER EQUATIONS

Now suppose that we wish to analyse (by operator splitting) some problem of unsteady three-dimensional flow. We will wish to construct the eigenvalues and eigenvectors of some matrix $\tilde{A}$ which maps $\Delta\mathbf{u}$ onto $\Delta\mathbf{F}$ with Property U. (The matrices which map $\Delta\mathbf{u}$ onto $\Delta\mathbf{G}$, $\Delta\mathbf{H}$ will follow from symmetry.) Evidently we may choose $\tilde{A} = (\tilde{C})(\tilde{B})^{-1}$. To find the eigenvalues of this mapping we may solve

$$((\tilde{C})(\tilde{B})^{-1} - \lambda I)\,\Delta\mathbf{u} = 0,$$

i.e.,

$$(\tilde{C} - \lambda\tilde{B})\,\Delta\mathbf{w} = 0, \tag{17}$$

i.e.,

$$\det(\tilde{C} - \lambda\tilde{B}) = 0. \tag{18}$$

At this stage it is convenient to divide through by $\overline{w}_1$ and then to adopt a convention that for the remainder of this paper u, for example, means

$$u = \frac{\rho_L^{1/2}u_L + \rho_R^{1/2}u_R}{\rho_L^{1/2} + \rho_R^{1/2}} = \frac{\overline{w}_2}{\overline{w}_1}$$

with a similar interpretation for v, w, H, i.e.,

$$v = \frac{\overline{w}_3}{\overline{w}_1}, \quad w = \frac{\overline{w}_4}{\overline{w}_1}, \quad H = \frac{\overline{w}_5}{\overline{w}_1}.$$

Then (18) reduces to

$$(\lambda - u)^3[(\lambda - u)^2 - (\gamma - 1)\{H - \tfrac{1}{2}(u^2 + v^2 + w^2)\}] = 0. \tag{19}$$

To find the eigenvectors, it is easiest to begin by finding their images in $\overline{w}$-space (by solving (17) with λ set equal to a root of (19)), and then mapping into $\mathbf{u}$-space. The results are

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ u - a \\ v \\ w \\ H - ua \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 0 \\ v \\ 0 \\ v^2 \end{pmatrix}, \quad \mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ w \\ w^2 \end{pmatrix},$$
$$\mathbf{e}_4 = \begin{pmatrix} 1 \\ u \\ v \\ w \\ \frac{1}{2}q^2 \end{pmatrix}, \quad \mathbf{e}_5 = \begin{pmatrix} 1 \\ u + a \\ v \\ w \\ H + ua \end{pmatrix} \tag{20}$$

(where $q^2 = u^2 + v^2 + w^2$ and $a^2 = (\gamma - 1)[H - \frac{1}{2}q^2]$), corresponding to the eigenvalues

$$\lambda_1 = u - a, \quad \lambda_2 = u, \quad \lambda_3 = u, \quad \lambda_4 = u, \quad \lambda_5 = u + a. \tag{21}$$

To complete the analysis we must show how to project an arbitrary $\Delta\mathbf{u}$ onto the eigenvectors as basis, i.e., how to find the coefficients a_i in

$$\Delta\mathbf{u} = \sum_1^5 a_i\mathbf{e}_i.$$

A routine calculation yields

$$\frac{a^2}{\gamma - 1}a_4 = (H - q^2)\,\Delta u_1 + u\,\Delta u_2 + v\,\Delta u_3 + w\,\Delta u_4 - \Delta u_5, \tag{22a}$$
$$wa_3 = \Delta u_4 - w\,\Delta u_2, \tag{22b}$$
$$va_2 = \Delta u_3 - v\,\Delta u_1, \tag{22c}$$
$$a_1 + a_5 = \Delta u_1 - a_4, \tag{22d}$$
$$a(a_5 - a_1) = \Delta u_2 - u\,\Delta u_1. \tag{22e}$$

For computational purposes, it is better to extract factors v, w from $\mathbf{e}_2$, $\mathbf{e}_3$, so that a_2, a_3 are never indeterminate. If this is done, we carry out the a posteriori check that the eigenvectors form a linearly independent set by arranging them into a matrix and finding the determinant; this comes out to be $2a^3/(\gamma - 1)$, which is never small unless the Mach number is very large.

If we wish to solve problems of wholly supersonic flow by marching in the x-direction, then we need a similar analysis for the mapping $\Delta\mathbf{F} \to \Delta\mathbf{G}$. This follows an identical pattern, but the results are slightly more complicated. The equation for the eigenvalues is

$$(\lambda u - v)^3[(\lambda u - v)^2 - a^2(1 + \lambda^2)] = 0 \tag{23}$$

from which we obtain eigenvectors

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ u + v/R \\ v - u/R \\ w \\ H \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ -H \end{pmatrix}, \quad \mathbf{e}_3 = \begin{pmatrix} w \\ 0 \\ 0 \\ 2H \\ wH \end{pmatrix},$$
$$\mathbf{e}_4 = \begin{pmatrix} 1 + \frac{1}{2}q^2/H \\ 2u \\ 2v \\ 2w \\ H + \frac{1}{2}q^2 \end{pmatrix}, \quad \mathbf{e}_5 = \begin{pmatrix} 1 \\ u - v/R \\ v + u/R \\ w \\ H \end{pmatrix} \tag{24}$$

corresponding to the eigenvalues

$$\lambda_1 = \frac{v - u/R}{u + v/R}, \quad \lambda_2 = \frac{v}{u}, \quad \lambda_3 = \frac{v}{u},$$
$$\lambda_4 = \frac{v}{u}, \quad \lambda_5 = \frac{v + u/R}{u - v/R}, \tag{25}$$

where

$$R^2 = \frac{u^2 + v^2}{a^2} - 1. \tag{26}$$

As before, we complete the solution by expressing

$$\Delta \mathbf{F} = \sum_1^5 a_i \mathbf{e}_i$$

and we quote the results, for which it is convenient to introduce

$$S = a_1 + a_5 + 2a_4. \tag{27}$$

Then we have

$$2Ha_2 = H\,\Delta F_1 - \Delta F_5, \tag{28a}$$
$$q^2 S = u\,\Delta F_2 + v\,\Delta F_3, \tag{28b}$$
$$2Ha_3 = \Delta F_4 - wS, \tag{28c}$$
$$(1 - \tfrac{1}{2}q^2/H)a_4 = S + a_2 - \Delta F_1, \tag{28d}$$
$$q^2(a_5 - a_1) = R(u\,\Delta F_3 - v\,\Delta F_2), \tag{28e}$$
$$a_5 + a_1 = S - 2a_4. \tag{28f}$$

The a posteriori check on independence of the eigenvalues leads to a determinant value of $8a^2H(u^2 + v^2)/R$, confirming our expectation that the only troublesome conditions in supersonic flow will be those where the local Mach number is very large, or its x-component is close to unity. Notice that this analysis simplifies usefully for the common special case of isoenergetic flow, defined by H = const. This condition replaces the fifth equation in the Euler system, and we may delete the second eigenvector, since some simple algebra reveals that $a_2 = 0$ in isoenergetic flow. Also we can shorten the solution for the $\{a_i\}$.

Strictly, no such simplification is possible in the unsteady case, since even if a flow originates in an isoenergetic stream, it will not remain isoenergetic unless very special conditions apply. However, a fictitious flow is a legitimate device for computing towards the steady state, as discussed by Viviand [13]. There are several ways to do this, and it is not easy to see which will converge fastest to the steady state. However, from the viewpoint of the present analysis, some alternatives are much simpler than others. This can be illustrated by means of the one-dimensional unsteady equations. Consider the momentum equation

$$(\rho u)_t + (p + \rho u^2)_x = 0, \tag{29}$$

where, in order to discuss the differential equations, we have reverted to conventional notation. We will eliminate p from this equation by making some assumption which is valid in the steady limit, and the result will combine with the continuity equation to give a pair of equations which can be solved for (ρ, u).

One strategy is to assume H = const, and then to differentiate H with respect to x, so obtaining

$$p_x = \frac{a^2}{\gamma}\rho_x - \frac{(\gamma - 1)}{\gamma}\rho u u_x. \tag{30}$$

The eigenvalues which arise from this approach are the roots of

$$\gamma\lambda^2 - (\gamma + 1)u\lambda + u^2 - a^2 = 0 \tag{31}$$

in which, as previously, $a^2 = (\gamma - 1)[H - \frac{1}{2}u^2]$. These are not the eigenvalues of the real time-dependent flow, although it is interesting that in both cases $u^2 = a^2$ produces a zero root. This strategy is mentioned by Viviand [13].

Alternatively, we assume

$$p_x = a^2\rho_x \tag{32}$$

with the same definition of a^2, and then the eigenvalues are found to be the roots of

$$\lambda^2 - 2u\lambda + u^2 - a^2 = 0 \tag{33}$$

just as in the real-time flow. There is no real paradox in reaching the contradictory conditions (31) and (33) from the false assumption H = const.

To implement the assumption (32), it can be shown that we must delete the fifth component from the eigenvectors in (20) and delete the fourth eigenvector from the list, whilst replacing (22a) by $a_4 = 0$. By taking this approach, the present method leads to a pseudo-time-dependent analysis of very simple structure.

A NUMERICAL EXPERIMENT

It is doubtful whether the accuracy of an approximate Riemann solution can profitably be discussed without reference to its intended use, so that numerical evidence has only a very limited value. However, one particular experiment has led to a rather striking result which does seem worth reporting. A variety of a new and established finite-difference schemes were compared by Sod [14] on the basis

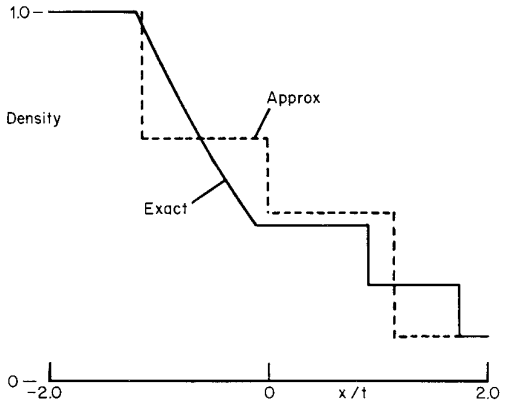

FIG. 1. Exact and approximate solutions to a Riemann problem.

TABLE I

Index (I)	Computed density	
	Godunov	Ref. [1]
72	0.2658	0.2655
73	0.2654	0.2652
74	0.2631	0.2629
75	0.2460	0.2458
76	0.1878	0.1881
77	0.1368	0.1370
78	0.1260	0.1260
79	0.1251	0.1251
80	0.1250	0.1250

of their performance on a standard shock-tube problem formulated in Eulerian coordinates. The density ratio was taken to be 1 : 8, and the pressure ratio, 1 : 10. In Fig. 1 we compare the exact solution for the density distribution with the present approximation. The comparison is not very close, but it is worth observing that the area beneath the two curves is identical (and can readily be proved so).

Now consider what happens if the exact Riemann solution is used in conjunction with Godunov's [4] finite-difference scheme to solve the same problem by advancing 35 time steps with $\Delta t/\Delta x = 0.411$ (Sod's standard comparison). Truncation error in the finite-difference scheme degrades the results, producing the somewhat smeared solution shown in Fig. 2. Now, if Godunov's scheme is applied to the linear advection equation, it reduces to first-order upwind differencing. First-order upwind differencing can be passed through the mechanism described in [1] to produce a new, nonlinear, first-order scheme, in which is embedded the present approximation. We can then try out the new scheme on the same test problem. Now by contrast, matters improve after the first time step. When compared on the standard basis (35 time steps) the differences between the two solutions are insignificant everywhere. They are largest near the progressing shockwave, but even there, Table I shows that they are very small.

The fact that the differences between the two solutions are so slight, compared with the truncation errors which exist in both, supports the arguments put forward in the Introduction. More accurate solutions to the same problem by methods involving the present analysis are shown in [1].

It may be interesting to insert here some observations concerning run times for the two methods. An attempt was made to eliminate subjective bias from the comparison by having rival programmers work on each method and the exercise formed part of a small project to improve programming technique. The algorithm for the exact solution was taken from Sod [14], but was improved in various small ways. At the time of the exercise, the version due

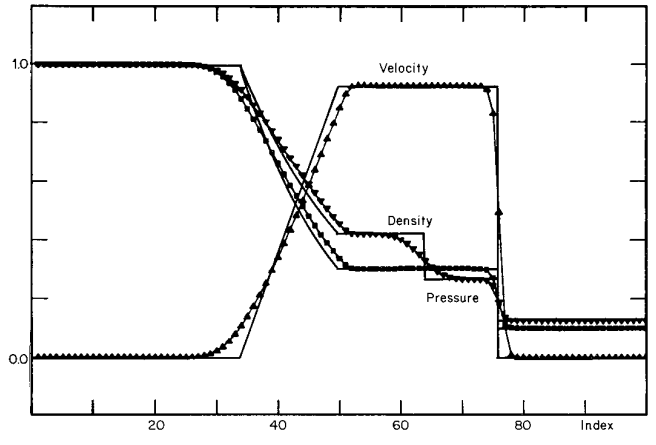

FIG. 2. Numerical solution to a shock-tube problem, incorporating either the exact or the approximate Riemann solution in a first-order upwind difference scheme.

to van Leer [7] was not available. However, this would probably not affect the results much since van Leer's main contribution is to reduce the number of iterations required to reach convergence, and on average only 1.6 iterations were needed anyway. Each of these iterations was found to take 0.20 ms on a DEC KL-10 computer using its optimised FORTRAN compiler, so that the total time spent on solving the Riemann problems was $0.20 \times 1.6 \times 99$ grid intervals $\times$ 35 time steps $\times$ 10^{-3} = 1.11 s. The total CPU time, excluding input/output operations, was 2.02 s, the remaining 0.91 s being accounted for by the Godunov differencing scheme into which the Riemann solutions were incorporated.

To find an approximate solution of each Riemann problem by the present method took 0.27 ms, so that the total time spent solving Riemann problems was 0.94 s, not in itself a very significant reduction. However, the pay-off comes from the fact that the information is immediately available in usable form, as time increments of $\mathbf{u}$ due to each wave system. This meant that the total CPU time (again excluding input/output operations) was only 1.09 s. Furthermore, this same feature could be exploited in higher-order algorithms, of the kind described in [1]. A typical second-order algorithm took about 1.20 s, and the incorporation of additional logic to make the results monotone raised the time to 1.78 s. The third-order monotone results shown in [1] took 1.90 s to produce.

A NOTE OF ENTROPY CONDITIONS

The main purpose served by introducing a Riemann solver (either exact or approximate) into a finite-difference scheme is presumably that of providing physical realism by correctly discriminating between information which should propagate with different speeds. The most basic distinction is simply between information which should propagate to the left or to the right. When the Riemann problem is solved exactly for a flux difference $\mathbf{F}_{i+1} - \mathbf{F}_i = \Delta\mathbf{F}_{i+1/2}$, this is accomplished as follows. Let $\mathbf{F}_M$ be the state obtained in the exact solution for $x = 0$, $t > 0$. Then

$$\Delta\overleftarrow{\mathbf{F}}_{i+1/2} = \mathbf{F}_M - \mathbf{F}_i,$$

$$\Delta\overrightarrow{\mathbf{F}}_{i+1/2} = \mathbf{F}_{i+1} - \mathbf{F}_M,$$

where the arrows denote contributions from left- and right-moving waves. In the present approximation we have (dropping the i subscript)

$$\Delta\mathbf{F} = \sum_j \lambda_j \alpha_j \mathbf{e}_j$$

and we identify the left- and right-moving components as

$$\Delta\overleftarrow{\mathbf{F}} = \sum_j \lambda_j \alpha_j \mathbf{e}_j \quad (\lambda_j < 0),$$

$$\Delta\overrightarrow{\mathbf{F}} = \sum_j \lambda_j \alpha_j \mathbf{e}_j \quad (\lambda_j < 0).$$

Now this identification fails, and some of the realism may be lost, if one of the waves (let us say the kth wave) should in fact be a fan which spans $x = 0$. In the exact solution part of the kth wave contribution to $\Delta\overleftarrow{F}$, and the remainder to $\Delta\overrightarrow{F}$, but in the approximate solution all of the kth wave will contribute to one or the other, depending on the sign of λ_k. An extreme example of this failure arises if we consider an interval for which $\mathbf{F}_i = \mathbf{F}_{i+1}$, $\mathbf{u}_i \neq \mathbf{u}_{i+1}$, and conditions are such that an expansion should occur, rather than a shockwave. In such a case, although $\Delta\mathbf{F}$ is zero, $\Delta\overleftarrow{\mathbf{F}}$ and $\Delta\overrightarrow{\mathbf{F}}$, as determined by the exact Riemann solution, will be equal and opposite. A differencing scheme which uses this information can break the discontinuity into the correct fan-like structure.

The present approximate analysis, if applied straightforwardly, would yield $\Delta\overleftarrow{\mathbf{F}} = \Delta\overrightarrow{\mathbf{F}} = 0$ and so provide no motive for breaking up the unrealistic discontinuity. An empirical cure for the problem is not hard to devise. For each interval which contains a sonic point one creates artificially an equal and opposite contribution to $\Delta\overleftarrow{\mathbf{F}}$ and $\Delta\overrightarrow{\mathbf{F}}$. Work is in hand to provide a fuller theoretical justification for this process and to find its neatest implementation.

It is quite possible that these precautions may sometimes be dispensed with. When the present methods have been applied to analyse the transonic flow over aerofoils [11] no special treatment was found necessary near the sonic line. However, it would be unwise to rely on such computations in any case where the physical correctness of the solution was less apparent. The prospective user would be advised to employ either an empirical device of the kind described above, or some form of artificial viscosity. A general theory which combines approximate Riemann solvers and artificial viscosities has been presented by Harten and Lax [10].

CONCLUDING REMARKS

In [1] we described a strategy for obtaining numerical solutions to hyperbolic initial-value problems. In the present paper we have enlarged the details of one element in that strategy, but feel that our results may perhaps find wider application. Our investigations into the Euler equations have revealed some very tidy structures. Also the existence of a fast approximate Riemann solver may be found more generally useful. Our programming experience is that the present direct method is about as time-consuming as one cycle of the iterative procedures mentioned in the Introduction. In the majority of cases which arise in a finite-difference calculation, the direct solution is already

very accurate and might be used to reduce the number of iterations needed to obtain any more exact solutions which may be required. However, this approach has yet to be explored. Finally we remark that a variety of interesting formulae may be obtained by translating the results of this paper from conservative variables into physical variables, such as (ρ, u, v, w, p).

REFERENCES

1. P. L. Roe, in *Proceedings, Seventh Int. Conf. Num. Meth. Fluid Dyn.* (Springer-Verlag, New York/Berlin, 1981).
2. O. A. Oleinik, *Trans. Amer. Math. Soc. Sect.* 2 **33,** 285 (1963).
3. P. D. Lax, *Hyperbolic Systems of Conservation Laws and the Mathematical Theory of Shock Waves* (SIAM, Philadelphia, 1972).
4. S. K. Godunov, *Mat. Sb.* **47,** 271 (1959). [As U.S. JPS translation 7226 (1960)]
5. R. D. Richtmyer and K. W. Morton, *Difference Method for Initial-Value Problems* (Interscience, New York, 1967).
6. M. Holt, *Numerical Methods in Fluid Dynamics* (Springer-Verlag, New York/Berlin, 1977).
7. B. van Leer, *J. Comput. Phys.* **32,** 101 (1979).
8. J. Glimm, *Comm. Pure Appl. Math.* **18,** 697 (1965).
9. A. J. Chorin, *J. Comput. Phys.* **22,** 517 (1976).
10. A. Harten and P. D. Lax, NYU Report 20E/ER/03077-167 (1980). [*SIAM J. Numer. Anal.* **18,** 289 (1981)]
11. C. C. L. Sells, RAE TR 80065 (1980).
12. S. Osher and F. Solomon, *Math. Comp*. **38,** 339 (1982).
13. H. Viviand, in *Proceedings, Seventh Int. Conf. Num. Meth. Fluid Dyn.* (Springer-Verlag, New York/Berlin, 1981).
14. G. A. Sod, *J. Comput. Phys.* **27,** 1 (1978).